Models, maps and metaphors: Why the brain is *still* not a computer, via Searle and Kant

Yasemin J. Erden [1]

Abstract. This paper builds on my AISB paper from 2013, which offered a defence of Searle's argument (that computation is observer relative) via discussion of Kant. The premise of that paper was that observer relativity is a necessary requirement for applications of concepts to objects, with a critique of, what I called, the *metaphor* of brain as digital computer. By describing such concepts as metaphors, I further claimed that they also lack epistemic content. One response to these ideas was that arguments for *mind as computer* employ *models* rather than metaphors. On this account, it is claimed that 'nature (including the brain) can be (and is) modelled computationally' [1]. This paper will explore that claim, and in so doing expand the defence of Searle's original account of computation by further developing the Kantian position on the necessary unity of consciousness in relation to objects and concepts. To this paper I add an additional Wittgensteinian view of metaphor (simile) and language to show why models offer no alternative by which to avoid the necessity of observer relativity for computation.

1 INTRODUCTION

In my recent AISB paper [2], I defended Searle's claim that the brain is not a digital computer [3], by exploring his ideas of the *observer standpoint* with regards to computation. I did this with reference to Kant, and primarily to show why it is impossible to identify ideas of computation without locating the possibility of an *identifier*. I claimed that the analogy between brain and computer is in fact *metaphorical,* while yet acknowledging both its appeal and its potential for value and creativity. Nevertheless, I stated that such a metaphor relies on the existence of those who create the metaphors, ask the questions and interpret the relations, and that (I thought) was sufficient. The argument was simply that without an observer to identify patterns and systems, there could not be any judgements of any sort, whether about computation or anything else.

The primary response to this argument concerned the question of *models*. A model is not a metaphor, or so the claim goes [1]. In this paper I will explore this counter view, and use this as a springboard to further my defence of the original position, but with the inclusion of *models*. A model is, I will argue, a manner of analogy in the same vein as a metaphor, and follows the same literary structure. It does not therefore offer a way to remove the need for the observer, but in fact further strengthens the argument that I offered at AISB 2013.

2 METAPHOR, MODELS AND MAPS

In [2], I explained that the idea of the brain as computer is problematic, by focussing on the idea of *simulation*, and accepting the very simple claim that operations can indeed be *simulated*. In such instances, simulations stand in place of events we either believe or predict have happened or could happen (likely or otherwise). I was then able to compare this function to *metaphors*, where words, ideas, concepts, and images stand in place of things in either a predictive manner, or by way of explanation. As I detailed then, the key difference is our *expectations* of a simulation as compared to a metaphor, whereby the former seemingly has claim to a higher degree of accuracy than the latter, or at the very least presents itself to us as such. Though I had challenged such assumptions, including about metaphors, I yet neglected to discuss the role of the *model* in this analysis. I turn to this now, and in so doing show that the model, as analogy or technical metaphor, is both useful and yet equally problematic. While on the one hand it may offer new paths to understanding or prediction, on the other, it gives false hopes toward objectivity.

The term model connotes something scientific, because it presents the idea of something verifiable (as advocated in logical positivism). A model presents itself as calculable, because it suggests the possibility of ideas (or facts) that are quantifiable, objective, unambiguous, and ascertainable. Conversely the rejection of a model *as* a form of metaphor would likely point to the status of the latter as ambiguous, uncertain, and almost certainly unscientific. Despite this, however, a model is itself analogous, or what we might call a *technical* metaphor. It is technical because it presents the possibility for testing and experimentation, but it remains the latter because even in this state it remains a process by which one thing *is used in place of another*. Rather like a map. A map does not claim to *be* the territory that it represents, but stands instead as a model for that which it more or less accurately depicts. Nevertheless, a map can never be said to be as accurate as the real picture (whatever we understand by accuracy) because it is dependent on scaling, interpretation, (often political) boundaries, and the selection or omission of data deemed to be more or less useful or valuable, or pertinent to a particular need or situation.

In this vein, a model succumbs to the same boundaries and thereby limitations as its analogous kin. A model cannot posit *all* the possible ways to view an event or idea, in the same way that a map cannot show all that it (in more or less broad strokes) depicts. Instead, models, maps and metaphors identify areas as *identified by observers* to be key to a description or event, or to highlight relations, ideas, similarities and difference arising. Within all of this, however, remains the need for interpretation and judgment, without which the possibility of either map-making (and reading), model construction, or metaphor writing, would be both impossible, and worthless.

To consider this in more detail, I turn to ideas on this topic by Dodig-Crnkovic [4, p.2]:

> Naturalism is the view that *nature is the only reality*. It describes nature through its structures, processes and relationships using a scientific approach. Naturalism

[1] St Mary's University, London, UK
✉ yj.erden@smuc.ac.uk

> studies the evolution of the entire natural world, including the life and development of humanity as a part of nature. Computational naturalism (pancomputationalism, naturalist computationalism) is the view that the nature is a huge network of computational processes which, according to physical laws, computes (dynamically develops) its own next state from the current one.'

This view stands in opposition to the ideas I posit above because it presupposes the possibility of an objective reality that we can access or model. Yet, and before we come to this as an argument in its own right, there is the concept of *nature* itself to consider, because the first question we need to answer is what this term depicts. Which reality? Reality as mediated by the interpreter, or an objective reality we believe exists because we experience it? If the former then this can be understood (following Kant's account) as *phenomena;* but if the latter, then as *noumena*. While I have already explained and discussed these terms in [2], a brief recount here might prove fruitful. Kant [5, A258/B313-314] states:

> When, therefore, we say that the senses represent objects *as they appear*, and the understanding objects *as they are*, the latter statement is to be taken, not in the transcendental, but in the merely empirical meaning of the terms, namely as meaning that the objects must be represented as objects of experience, that is, as appearances in thoroughgoing interconnection with one another, and not as they may be apart from their relation to possible experience (and consequently to any senses), as objects of the pure understanding. Such objects of pure understanding will always remain unknown to us; we can never even know whether such a transcendental or exceptional knowledge is possible under any conditions—at least not if it is to be the same kind of knowledge as that which stands under our ordinary categories.

If we accept this—and I suggest that to do otherwise would presume a level of objectivity unavailable to us as subjects or as *beings-in-the-world*—then we also accept that our use of the term *nature* here stands as yet another model, metaphor or map for interpreting and making sense of what we see around us. This is not to say it is not a very useful model, but then again this is no more than saying that a metaphor may also be useful. As I discuss in [6], Wittgenstein's approach to metaphor and simile was to view these as exceptionally useful in philosophy. As he states here [7, p.18]:

> In notes, examples and similes are always useful. If I could give you enough of them, that would be all that would be necessary. Usually we think of similes as second-best things, but in philosophy they are the best thing of all.
>
> Part of our subject is that we must jump about and make connections.

Models, just as metaphors and maps, offer the opportunity to explore relations between things, and to draw more or less successful conclusions and predictions. As explained in [2], metaphors (to which we now add models and maps) offer *additional* information or information reinterpreted. This is different to a view that identifies this kind of information (specific) as *intrinsic* to the properties of an object or concept (general).

For example, if we model the life-cycle of the butterfly, and identify the processes by which an egg is laid, hatches as a caterpillar, and (via pupa stage) becomes a butterfly and eventually dies, what we do is identify those key features that we predict, test, and believe (with varying degrees of the latter at each stage) are most crucial to the development of this creature. Perception is of course key, because this standard view of the development of a butterfly emerges from what we see in the first instance. There are of course other ways to model this development. Genetically or hormonally, for instance, as we offer an account that instead maps the division of genes, and other internal (unobservable to the naked eye) changes. The response to this might be that a second model of this kind would not in fact contradict the first, and this is perhaps true in the circumstances, but it is not in fact a given. There are plenty of co-existing models, which both contradict each other yet are not mutually exclusive. One such example can be taken from nanotechnology (or –science).

Properties of matter at the atomic scale compared with at the level of human perception sometimes differ radically. As I discuss elsewhere [8],

> physico-chemical properties can change at the nanoscale, nanomaterials will contain properties that behave differently from normal materials. The small size offers a large surface area (thereby offering more of the same), such that just small quantities of particles can offer large reactivity and functionality.

At the nanoscopic scale (nanoscale) gold can appear *red*, and has active properties [9], while yet being inert at standard perceptual levels. Our world *as we understand it* does not function in the same way at the nanoscale, and yet this has not been a cause to completely rethink the models by which we map our world of perceptual phenomena. Though, of course, that could still change, it is just as possible that it will not. Gold as *we generally perceive it* is gold in colour, and inert. That the model of gold at the nanoscale differs (true too for silver—which is also active at the nanoscale) does not engender a contradiction. Instead, we recognise that 'active antimicrobials (such as metal oxide nanoparticles in silver or gold, among others' [8] which are perceptible to us at one level, yet not another, make up the complex properties of this clearly little-understood material. This can of course be said of any material—though some may turn out to be less surprising. The purpose of identifying this here is to show why our perceptions of the world *as phenomena*, cannot offer accounts of *noumena*, nor guarantee unquestionable accuracy even of the phenomena as we perceive it. Against what would we measure this?

Uncertainty about nanotechnologies offers another useful tool (or model) for mapping what we mean by observer relativity above. As I describe in [8], the novel properties of some particles at the nanoscale offer both the potential for exciting new developments, but also for recognising how little we have already understood about biological organisms. More specifically, there are uncertainties that arise as a consequence of

the complexity of living systems in their responses to nanoscale entities. Despite general uncertainty about the overall functioning of our bodies as a complex organism. The assumption that changes occurring within a body do so discretely is continuously challenged, and yet the presumption of this persists through the ways in which we both treat and perceive illness, health and our physical relationship with the world.

For instance, if we offer a model of health that focuses on nutrition only, we might think that this model is limited. If we add to this information about exercise, it may be that this is considered more holistic. But we might also need to think about health in terms of mental health, or happiness, or even with regard to occupational health (including the ability to function in everyday life). Any individual model that encompasses more or less of these themes, and/or others omitted here, may be considered as more or less complete. This process by which we determine the utility of such models does not end, since at any given point new information might challenge, contradict, support, or enhance (or even a combination of these or other responses) such a model. There is no reason to assume that any particular model of nature, or of the brain, or of the claim to wandering *lonely as a cloud* could ever be otherwise.

Metaphors, models and maps exist in conceptual space between more or less disparate ideas or objects. It is for these reasons that they can prove so fruitful, but also for these reasons that they can seduce us to think they offer more than they can. It is to this that I now turn.

3 SEEING *AS*

Searle states that the identification of a process *as* computational 'does not identify an intrinsic feature of the physics, it is essentially an observer relative characterization.' I take this to mean that for something to function *as* something requires that we use the *as* in a comparative sense. In the previous incarnation of this paper I suggested that this includes be a *metaphorical* sense to the comparison. To this I now add that it also functions as a model, and in this way, the model is akin to the metaphor (as explained in the sections above). It is for this reason, or so I argue, we can follow Searle's further claim that 'nothing is intrinsically a digital computer solely in virtue of its physical properties' [3].

Just as when I place my portable computer on my lap, and thereby say that I use my lap as a table (where in the previous paper I used my hand *as a plate*) I do little more than recognise that in these instances (i.e. resting object on, or eating food over) the lap and hand function in similar ways to tables or plates. The context and circumstances help define these functions, and it would be difficult to know how to understand terms like *plate* or *table* without such context or application. It is for these sorts of reasons that I argue in [2], as indeed I do now, that it is in the making of these observations that a comparison can emerge. The *comparison* is thus relative to my (or any other who can or will draw such connections) having observed it. As I also explained in [2], the reason this does not engender multiplicity of inaccessible first-person subjective experiences of the world, is that when I utter such comparisons (e.g. about laps as tables, hands as plates, or indeed the loneliness of clouds and what it would be to wander like one), it would be easy enough for others to understand what I mean by this. Or at least, to imaginatively engage with what I might mean. For similar reasons the accusation of *relativity* in a far stronger philosophical terms can in fact be avoided, but I will return to this later.

It is at this juncture that I wish to repeat some more substantial portions of the first paper, and to explain how they apply, at every step, to the *brain as computer* analogy as both a model and as a metaphor, and how this does not change its identity as a form of *comparison,* with all the benefits and limitations that engenders.

I begin by turning once more to Searle's comment that 'we could not discover objects in nature which were functioning as chairs, except relative to some agents who regarded them or used them as chairs' [3], and to point out (once again) the relationship of this with Kant's [5, B137/138] explanation that,

> The first pure knowledge of understanding, then, upon which all the rest of its employment is based, and which also at the same time is completely independent of all conditions of sensible intuition, is the principle of the original *synthetic* unity of apperception. Thus the mere form of outer sensible intuition, space, is not yet [by itself] knowledge; it supplies only the manifold of *a priori* intuition for a possible knowledge. To know anything in space (for instance, a line), I must *draw* it, and thus synthetically bring into being a determinate combination of the given manifold, so that the unity of this act is at the same time the unity of consciousness (as in the concept of a line); and it is through this unity of consciousness that an object (a determinate space) is first known. The synthetic unity of consciousness is, therefore, an objective condition of all knowledge. It is not merely a condition that I myself require in knowing an object, but is a condition under which every intuition must stand in order *to become an object for me*. For otherwise, in the absence of this synthesis, the manifold would *not* be united in one consciousness.

As explained in [2] it is the observer who is active in both the *drawing* of the line, and in the *recognition* of it as such, and this is a condition for any knowledge we can have of lines. This applies even where this knowledge is mapped into a model and even where this model allows us to predict with repeated accuracy our understanding and manipulation of lines. But again, this does not remove the aspect of *phenomena,* whereby without this experience there can be no line to which to refer. Without any human experience, there is no knowledge of *lines* (whether as concept, or as represented in space) in any way that we could make sense of. Furthermore, our understanding of lines is far more basic than our (until now) understanding of gold, which as discussed above, is recently expanded. For all this, the map for understanding gold or lines can remain accurate (to greater or lesser degrees), while yet allowing that this experience of the world is necessarily contingent on the relation between observer and object. In the example of gold, for example, we can see that it also remains open to change.

As covered in [2] Searle expounds a similar point when he observes that 'Computational states are not discovered within the physics, they are assigned to the physics' [3], and that 'There is no way you could discover that something is intrinsically a digital computer because the characterization of it as a digital

computer is always relative to an observer who assigns a syntactical interpretation to the purely physical features of the system' [3]. If we expand the example to include that of the properties of gold, at both the nanoscale and higher, we can see why this might occur.

The example offered by Searle, of the impossibility of an 'unknown sentence' in ones head, remains key to this. As he rightly points out, a sentence requires active construction, use, and recognition. Rather like the drawing of a line. Repetition can certainly embed sentences within our languages, such that they require little apparent thought in their uttering (for example, 'how are you?' can elicit a 'fine, thanks, how are you?' with minimal genuine consideration to the question), but this shows little more than the habitual aspects of language use. Even if it sometimes seems otherwise, sentences are created, not found. This is true also for patterns, which, if not created, are identified as such by observers:

> the only sense in which the specification of the pattern by itself provides a causal explanation is that if you know that a certain pattern exists in a system you know that some cause or other is responsible for the pattern. So you can, for example, predict later stages from earlier stages [3].

What tricks us into confusing how the relation lies, is the *meaning* of the words,

> We are blinded to this difference by the fact that the same sentence, "I see a car coming toward me", can be used to record both the visual intentionality and the output of the computational model of vision. But this should not obscure from us the fact that the visual experience is a concrete event and is produced in the brain by specific electro-chemical biological processes. To confuse these events and processes with formal symbol manipulation is to confuse the reality with the model.

We can be further tricked if we assume the model gives us access to noumena unmediated by the relation with the observer (as phenomena).

Thus, as I claimed in [2], I restate here: the question with which Searle engaged in [3] is one to discard as 'ill defined', rather than refute. The model of brain as computer *can* be useful, but it can also serve to limit how we think about what we understand and even expect from both brains and computers. Particularly if we discard those elements that do not fit the model we construct. For instance, and building on the example above, if we were to privilege one model of the properties of gold over another (including regarding nanoscale properties as more important than its socio-ecomonic status). The fact is that *brain as computer* has been a dominant model for a number of years, which has had ramifications, not least for how we understand mind and mental activity. As discussed above, we might think deficient a model of health which focussed on nutrition alone (and apart from context), yet there are substantial claims made about illnesses such as depression as illnesses of the brain alone. Attempts to cure them in this fashion follow accordingly, and sometimes with disregard for context and other factors, as discussed in [10].

Kant [5, A244/B302] once again proves instrumental here:

> So long as the definition of possibility, existence, and necessity is sought solely in pure understanding, they cannot be explained save through an obvious tautology. For to substitute the logical possibility of the concept (namely, that the concept does not contradict itself) for the transcendental possibility of things (namely, that an object corresponds to the concept) can deceive and leave satisfied only the simple-minded.

4 THE CONTINGENCY OF OBSERVER AND OBJECT IS *NECESSARILY* NECESSARY

The previous section was titled 'seeing *as*', by which I partly alluded to Wittgenstein. In this section I return to Searle's account of computation in relation to syntax, but I will also expand this to include comment by Wittgenstein on drawing connections.

In [3] Searle claims that 'syntax is essentially an observer relative notion', to which he adds, 'The ascription of syntactical properties is always relative to an agent or observer who treats certain physical phenomena as syntactical' [3]. Furthermore,

> The multiple realizability of computationally equivalent processes in different physical media was not just a sign that the processes were abstract, but that they were not intrinsic to the system at all. They depended on an interpretation from outside. We were looking for some facts of the matter which would make brain processes computational; but given the way we have defined computation, there never could be any such facts of the matter. We can't, on the one hand, say that anything is a digital computer if we can assign a syntax to it and then suppose there is a factual question intrinsic to its physical operation whether or not a natural system such as the brain is a digital computer. [3]

This idea of the *requirement of interpretation from outside* is one that I touched on in [2], but which I will now expand. The first point remains that the definition of computation can only be applied to systems that we *recognise* as computational. I had stated this to be 'insubstantial' as a statement, and one that might not be controversial. I had underestimated this. To clarify therefore, I still surmise that disagreements hinge on how we understand *recognition*, or its cousin *observation*. Opponents of Searle sometimes take his claim about the contingency of observation to be naive. As I noted in [2] Endicott [11, p. 104], for instance, rejects Searle's account of computation as too simplistic. In its place he claims that 'a system is a genuine computational device when there is a correspondence between its physical states and its formal states such that the causal structure of the physical system is isomorphic to the formal structure of the computational operations' [11, p. 104]. His refutation of Searle seems to hinge on a seemingly Platonic account of computation, whereby the system or computation has an identity as a thing-in-itself, which we then *discover*. As already noted above, this is problematic for a number of reasons. Before I explore this, I wish to include further claims for the

argument from models. Dodig-Crnkovic [4, p.3] writes that 'within the info-computational framework computation is defined as information processing,' and later [p.4] explains that Darwin's *ideas* are 'supported by computational models'. To this she adds [p.4], 'Computation in nature can be described as a self-generating system consisting of networks of programs, a model inspired by […] self-modifying'

While it would seem that part of the general criticism of Searle's argument hinges on an epistemological claim about an objective truth about computation (that there *is* something that is *in-itself* and objectively, computation), it is interesting to note that the terms by which such claims are defended rely on the same comparative model. Terms like *defined*, *supported,* and *described* do not suggest objective certainty, but instead belie the subjective relativity which may prove more or less accurate or useful (as discussed above), but do not bring us closer to *noumena*. It is not a failure of our language, but rather a failure to acknowledge that our language betrays our necessarily uncertain and fuzzy descriptions of the world.

To say that something *can be described as* does not indicate a *failure to be accurate,* but rather shows the limitations of a model. The fact is that if it *can be* described that way, we take the meaning to be it *can also* be described otherwise. In the same way that a person 'can be' described as tall relative to the standard of tallness in one culture, whereas they may be described as of average height according to the standards of another. This sort of epistemic uncertainty does not always or necessarily bother us in our everyday lives (we may not mind being called tall at one time, or not at another—then again, we may), but it seems to bother us when what we seek is more certainty.

Wittgenstein is useful at this juncture, since his ideas about *seeing as* (referred to above) and language-games highlight exactly this issue of fuzziness. A quotation that draws this out is where he suggests [12, pp. 170-1]:

> The concept of 'seeing' makes a tangled impression. Well, it is tangled.—I look at the landscape, my gaze ranges over it, I see all sorts of distinct and indistinct movements; this impresses itself sharply on me, that is quite hazy. After all, how completely ragged what we see can appear! And now look at all that can be meant by 'description of what is seen'.—But this just is what is called description of what is seen. There is not one genuine proper case of such description—the rest being just vague, something which awaits clarification, or which must just be swept aside as rubbish.

As explained in [13] this neither means that all humans make the *same* connections—to which we can add *comparisons* (maps, models, metaphors)—nor that such connections are *completely* arbitrary. Wittgenstein's point it that by sharing a common *form of life*, we are apt to make similar sorts of connections and comparisons, or at the very least be capable of understanding even those that may appear radically different to ours, and vice versa. None of which could lead to the objectivity sought by some critics of Searle's position.

At this point we can return to Searle, because (I suggest) his claim in this regard is more complex than some of his critics seem to have recognised. Indeed, as I argued in [2], it points to the very core of the meaning that we understand by a term such as computation. The Platonic search for objective meaning or identity, of the sort espoused above, is flawed precisely because the expectation is that meaning is there to be *found*, rather than *determined.* For instance, and as explained in [2] where we recognise the pattern of some migrating bird formation as akin to a 'v' shape. In such instances, we are not saying that there are shapes in the world that *are* necessarily 'v', which await our discovery. Instead, the comparison, or the metaphor, map or model, is drawn between the formation of birds in flight, and the letter 'v' that we use in language. It is clear that v-*ness* (where v is a symbol) is not somehow inherent to what it is to migrate, but rather that in our interpretative understanding of the world, we make sense of what we understand based on what we have already learned. The model as comparison is extremely useful in this regard, but its position as model *that could be otherwise* cannot be ignored. Kant's [5, A258/B313-314] account of phenomena and noumena as discussed above remains pertinent here.

For these reasons, and as described in [2] it seems unfair to attribute to Searle a position from which is offered an *irrefutable* definition of computation. By which I mean his point seems not to be centred on a declaration of what computation *is* but only to show what it *isn't*, i.e. an independently and objectively verifiable system that exists *independent* from our system of its interpretation as such. The added criticism offered by Endicott [11, p. 107] that there are 'multiply realised types within the domain of physical or natural science', which are not observer relative, is also refutable. Once again, the way in which we understand such systems to operate requires that we understand there to be a system as such. On this account there is no noumenal objective reality about there being a system, and even if there were, we could know nothing about it. There is no way in which we can remove ourselves (with our interpretation of things, including our understanding of systems) such that we can *know* that a system simply *is* in any objective sense of that term. Thus, simply claiming that there *are* these systems, which are somehow separate to our interpretation of them as such, makes no sense at all.

Nor would it make sense to imagine a model in objective terms. If, for example, we recognise the system of flowers and bees and pollination (or even recognise this as a model), where do we recognise the system or model to lie? Do we recognise the system or model from the viewpoint of the flower, or that of the bee, or of the two as a symbiosis? Which part do we model, and what do we select for inclusion and consequently exclusion in so doing? Could this be seen in another way? Again, I point to the example of gold to aid our thinking here.

Yet none of this need engender accusations of relativism or anti-realism. This is partly for the reasons already explained above (regarding a shared form of life), but also because of Kant's [5, B276-277] account that it is only *by means* of 'outer experience' of objective qualities (such as mass, identity, shape) that 'inner experience' is possible. Rejecting claims of foundational Cartesian subjectivity, he notes [5, B277] that while 'the representation' *I am* may include the '*existence* of a subject' (emphasis added) it includes no '*knowledge* of that subject, and therefore also no empirical knowledge, that is, no experience of it'. To which he adds:

> For this we require, in addition to the thought of something existing, also intuition, and in this case

> inner intuition, in respect of which, that is, of time, the subject must be determined. But in order so to determine it, outer objects are quite indispensable; and it therefore follows that inner experience is itself possible only mediately, and only through outer experience.

Elsewhere he expands on the point with respect to our understanding [5, A820/B848]:

> The holding of a thing to be true is an occurrence in our understanding which, though it may rest on objective grounds, also requires subjective causes in the mind of the individual who makes the judgment. If the judgment is valid for everyone, provided only he is in possession of reason, its ground is objectively sufficient, and the holding of it to be true is entitled *conviction*…
>
> The touchstone whereby we decide whether our holding a thing to be true is conviction or mere persuasion is therefore external, namely, the possibility of communicating it and of finding it to be valid for all human reason.

To put this another way, if there is nothing about which your judgements *are*, then there is no particular *reason* to make one judgement over another [14]. Were there to be no objective grounds, explains McDowell [15, p. 67], all we would be left with would be a 'frictionless spinning in the void'. Another highly instructive metaphor, but not one which anyone would likely demand we take to be true in any apparent *objective* sense. As the paper has argued, we can quite reasonably add the concept of model to this claim, with little lost in the doing.

5 CONCLUSION

This paper sought to add even more weight to Searle's arguments as previously explored in [2], regarding the necessary relation between observer and computation. It has not spent much time engaging with the standard objections (e.g. [16] and [17]), since the ideas offered from Kant, cited throughout, should hopefully circumvent some of these more typical objections. It has further added ideas like maps and models to the mix, and attempted to show why it might be that these terms suggest more than a metaphor, but ultimately rely on the same modes of comparison, and thereby engender the same reliance on interpretation and judgement. With the failure to recognise the metaphorical or comparative nature of terms like *computation* comes a broader failure to recognise how language of this sort relies on context for meaning (as I discuss elsewhere [13]). The origins of the term *computation* as applied to humans, and its later application to machines should have given some indication of this fluency of meaning, and employment of metaphor, as well as the limitations of a model that identifies and describes behaviour. A person who can act in the manner of a computer, does not as a consequence of this *become one above all else*. There would end the value of such a metaphor.

Where this has not been understood, I have employed Kant and Wittgenstein. As explained in [2] it should by now be abundantly clear that the very possibility of my recognition of these comparisons relies on my skills for recognising metaphors, systems, patterns, maps, and models, and in offering judgements and interpretations that—I remain hopeful—may be agreeable to others. Whether by others who recognise the same patterns, or by those who can at least imaginatively explore these ideas as explained by [12] above. Copeland [18] says 'Searle is telling us no more than that if the brain is a computer, then it is so only in the sense in which all other computers are computers. This is hardly interesting'. On this I maintain disagreement. Once again, it is in the showing of *limitations*—of what is clearly a captivating metaphor or, for that matter, model—that Searle's ideas are themselves most interesting indeed. And we would do well to heed the warning that the limitations impress upon us.

ACKNOWLEDGEMENTS

I remain grateful to Stephen Rainey for his continued assistance in making sense of Kant; to Mark Bishop, whose ideas and suggestions sparked the inception of an earlier draft of this work, and whose constructive feedback on that draft helped shape its direction. In addition, sincere thanks to Gordana Dodig-Crnkovic, with whom began a very interesting discussion about models that has proven most fruitful. She was also kind enough to engage in correspondence about the topic, which was enormously useful. Finally, thanks are due to reviewers of this paper who were highly instructive in their comments. It is perhaps obvious, but worth stating anyway, that any and all faults and failures continue to be entirely my own.

REFERENCES

[1] G. Dodig-Crnkovic, *in private correspondence.*
[2] Y. J. Erden, The 'simple-minded' metaphor: Why the brain is *not* a computer, via a defence of Searle, AISB Proceedings, 2013.
[3] J. R. Searle, Is the brain a digital computer?, *Proceedings and Addresses of the American Philosophical Association*, 64: 21–37, 1990. Cited version online (accessed 05/02/2013): https://mywebspace.wisc.edu/lshapiro/web/Phil554_files/SEARLE-BDC.HTM
[4] G. Dodig-Crnkovic, Modeling Life as Cognitive Info-Computation, Computability in Europe CiE 2014 Proceedings, 2013.
[5] I. Kant, *Critique of Pure Reason* (trans. N. K. Smith), London: Macmillan press, 1933 (2nd ed.).
[6] Y. J. Erden, Wittgenstein on simile as the 'best thing' in philosophy, *Philosophical Investigations*, 35(2): 127-137, 2012.
[7] L. Wittgenstein, [*Dictating Philosophy* manuscript]. ed. A. Gibson.
[8] Y. J. Erden, Ethics, communication and safety in the use of PNFP in *Ecosustainable Polymer Nanomaterials for Food Packaging*, CRC Press, 2013.
[9] R. Elghanian *et al*. Selective Colorimetric Detection of Polynucleotides Based on the Distance-Dependent Optical Properties of Gold Nanoparticles. Science 277(1078), 1997.
[10] Y. J. Erden, Neural implants, human identity, and perceptions of illness, RiCT: Responsible Innovation Project, 2013. URL: http://responsible-innovation.org.uk/torrii/page-resource/Ethical%20issues
[11] R. P. Endicott, Searle, syntax, and observer relativity, *Canadian Journal of Philosophy,* 26(1): 101-122, 1996.
[12] L. Wittgenstein, *Philosophical Investigations*. Oxford: Blackwell (Trans. G. E. M. Anscombe), 2001.
[13] Y. J. Erden, Could a created being ever be creative? Some philosophical remarks on creativity and AI development, *Minds and Machines*, 20(3): 349-362, 2010.
[14] S. Rainey, *A Pragmatic Conception of Rationality*, Doctoral Thesis, QUB, 2008 (unpublished).

[15] J. McDowell, *Mind and World*, Camb. Mass.: Harvard Uni. Press, 1994.
[16] D. J. Chalmers, Does a rock implement every finite-state automaton?, *Synthese*, 108(3): 309–333, 1996.
[17] R. L. Chrisley, Why everything doesn't realize every computation, *Minds and Machines*, 4(4): 403–420, 1995.
[18] B. J. Copeland, What is computation?, *Synthese*, 108(3): 335–359, 1996.