How is Free Will Possible?

This note shows how a self-model approach can be used to describe the genesis of consciousness and the relationship between consciousness and free will.[1]

In doing so the note describes a hypothetical information processing arrangement called the *action cycle*, which is considered central to the moment-by-moment delivery of a person's subjective sense of self as a physical body acting on a physical world. [2]

Introductory material

To allow this explanation to be free standing – and to provide a clear common basis with other material presented on this website – this note begins with an expanded, mildly modified version of the material provided at the beginning of the Introductory Summary[1]:

1) Assume there exists independent of any observing subject, a single universe, U, which contains all things, including all human beings.

2) If it is accepted U exists in its own right, then a *component* of U – call it the *noumenal world*, W[r] – can be defined, where W[r] is the observer-independent component of U.

   2a) W[r] will be the world underlying appearances but not including appearances. Thus, W[r] will be where information capable of sensory perception comes from, regardless of whether there are observers present to perceive that information.

3) The component of a person which is manifest at the level of W[r] can be called their *noumenal body*, B[r].

   3a) Similarly, there will be a component of a person's brain manifest at the level of W[r], call it brain[r]. Also, there will be trees[r], rocks[r], beetles[r] and so on. (These examples introduce use of the suffix [r] to denote the *noumenal component* of any entity or process.)

4) A person's conscious experience of themselves as they act in the world will be generated by processes[r] in their brain[r] which *inter alia* use real-time sensory input flowing from W[r] into B[r]. Their brain[r] must use that sensory input to generate, for its internal purposes – including conscious sensory perception, and the conscious planning and execution of physical acts – a dynamic *real-time* representation[3] of B[r] in W[r] which can be called B[i] in W[i]. Call W[i] the *world image* and call B[i] the *body image*.

   4a) Your B[i] will be your body as you consciously perceive it, and W[i] will be the world you perceive as containing your body and all other physical things. Thus, *importantly* W[i] will be what you refer to as "the physical world". W[i] will contain your B[i], and within it your brain[i], and also – beyond B[i] – trees[i], rocks[i], beetles[i] and so on.[4]

---

[1] This note is an element of a wider discussion, which can be accessed via http://teleodyne.com.

[2] The action cycle is described in a wider context in the Main Essay (page 56-59) (see http://teleodyne.com).

[3] Here the term 'representation' can be used interchangeably with the term 'model', as applied in the concept of a *phenomenal self-model* after Metzinger, T. *Being No One* MIT Press, Cambridge, Massachusetts USA 2003.

[4] These examples introduce use of the suffix [i] to denote the *phenomenal component* of any entity or process. Note that any entity[r] which is capable of being perceived will lead to a subjectively perceived entity[i] that is unique to each observer. Thus, in the presence of a table[r], what you see as the table will be your table[i] and what I see will be my table[i].

The idea that conscious experience of the self-in-the-world must rely on brain[r] generating a real-time representation of B[r] in W[r] is logically necessary. Consider the following example:

> View the ceiling above your head. You know that you are able to see the ceiling only because light travels from it into your eyes, where incoming light patterns are then translated into neural impulses. But if that is true, how can the ceiling appear subjectively to you to be outside and above your head, rather than inside your brain, where those neural impulses are being processed?
>
> The reason can only be that information[r] is passing from ceiling[r] via your eyes[r], into your brain[r], where it is being processed such that you subjectively can experience that information as ceiling[i] which you perceive as being some distance outside and above your head[i]. The only efficient explanation can be that, the whole time, head[i] and ceiling[i] are inside W[i] which, *at the level of information processing*, must be inside your brain[r]. W[i] is *inside* your brain[r] by virtue of having been generated by your brain[r], *within* your brain[r], through brain[r] processing of sensory input[r].

Figures 1 and 2 illustrate key aspects of the ideas expressed above:

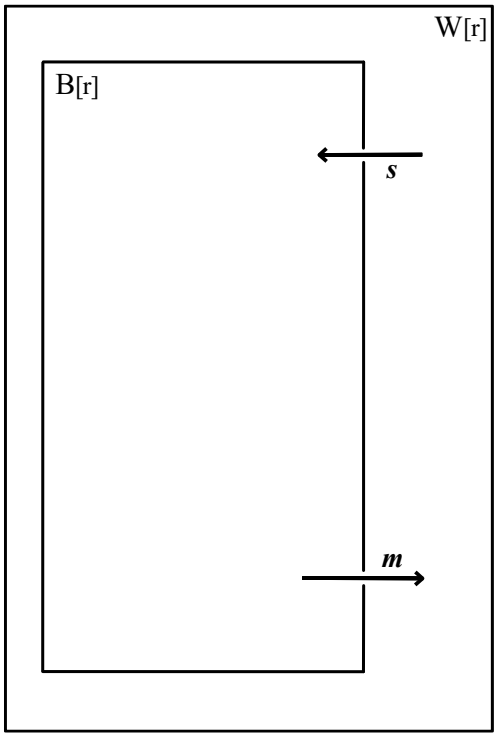

Figure 1

Figure 1 illustrates the boundaries and information flows discussed above at the noumenal level, where W[r] represents the *noumenal world*, B[r] represents the component of W[r] that is the *noumenal body*, ***s*** represents the information stream that is sensory input from W[r] into B[r] and ***m*** represents the information stream that is *voluntary* motor output from B[r] into W[r].

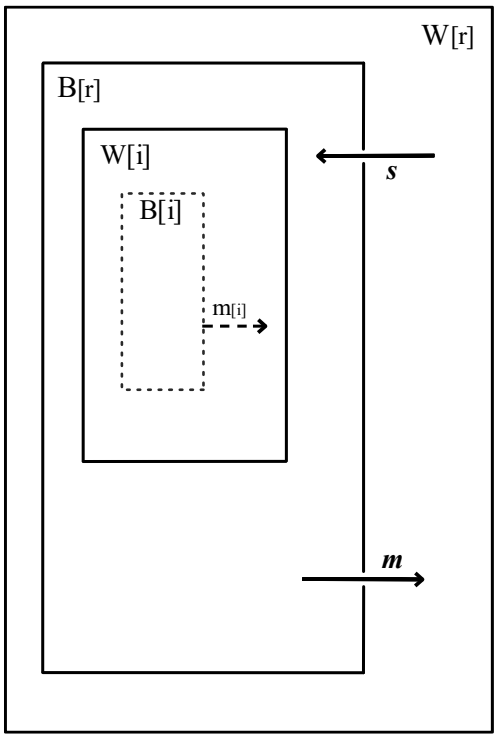

Figure 2

Figure 2 extends Figure 1 to illustrate the presence – within B[r] at the information processing level – of a *real-time* representation of B[r] in W[r], shown as B[i] in W[i], where W[i] is the *world image* and B[i] is the *body image*, as described from (1) to (4a) in the text. The boundary around B[i] is dotted simply to illustrate that it is a boundary at the phenomenal level. It denotes the surface of a person's body as they consciously perceive it. Likewise, the dotted arrow denoted m[i] is a person's motor output – i.e. their actions on the world and changes of body posture – as they perceive these, as opposed to ***m***, which is the actual voluntary motor output they are making at the noumenal level of B[r] in W[r]. (We do not *directly* perceive the noumenal level sensory input ***s***, except in the very crucial sense of its use by information processing systems within B[r] to generate B[i] in W[i], so no dotted arrow marked 's[i]' is shown in the figure.)

In relation to consciousness, we can speculate that:

5) If a person consciously experiences themselves in real time as B[i] in W[i] then there must be a set of processes in brain[r] which maintain *in real time* a high level of *fidelity* between the phenomenal level of B[i] in W[i] and the noumenal level of B[r] in W[r].

5a) That must be so, or a person's capacity for normal, effective sensory-motor coordination would fail. Why would it fail? Because each physical *action* we make must first occur at the noumenal level of B[r] acting upon W[r], and then be followed, on a moment-by-moment basis, with monitoring at the level of B[i] in W[i] for confirmation that the action is proceeding as planned. If something is seen to be going wrong with the action as it is reflected in B[i] in W[i], then conscious correction of that action can take place. Thus, if in real time the dynamic representation B[i] in W[i] did not with adequate fidelity reflect the dynamic situation of B[r] in W[r] a person would be unable to practice effective sensory-motor coordination.

6) If there is a set of processes in brain[r] responsible for maintaining real-time fidelity between B[i] in W[i] and B[r] in W[r], then they may have a frequency. That is, there may be a frequency with which B[i] in W[i] is updated to reflect changes in B[r] in W[r]. Let us assume for the sake of discussion that there is such a refresh rate and that its frequency is about four times a second (4 Hertz).[5]

How might all this lead to an account of the processes by which consciousness can be generated? Consider the following:

7) A subjective, moment-by-moment sense of self-in-the-world may be generated in the course of a process in brain[r] which at a rate of say about 4 Hertz *realigns* a person's B[i] in W[i] (the physical world as they perceive it) to update it back to high fidelity with their B[r] in W[r] (the physical world as they act upon it).

   7a) The moment-by-moment instantiation of consciousness might then take place at about 4 Hertz, when each realignment to bring B[i] in W[i] back to high fidelity with B[r] in W[r] enables an information processing arrangement in brain[r] to operate on an absolute presumption that B[i] in W[i] is *in identity with* B[r] in W[r].

In other words, consider that, on a moment-by-moment basis, consciousness might arise through brain[r] processes that manipulate information on a strict operational premise that the self as it perceives itself, B[i] in W[i], *is one and the same* as the self as it acts, B[r] in W[r].

Definition and operation of the action cycle

To better understand how this can work – and how free will can arise from it – requires a more specific description of the brain[r] information processing arrangements referred to above. In this connection a detailed description of a *hypothetical* information processing arrangement called the *action cycle* has been made in the Main Essay[2]. The following description is a complementary account of the same idea.

What needs to be considered closely is how a conscious adult human being might interact with their physical environment in real time, on a moment-by-moment basis. The proposal made here is that this interaction may involve a six step (at least) information processing cycle operating in the following sequence[6], that runs at about 4 times a second, that is operating continuously and end-to-end whenever a person is conscious, and that is central to delivering moment-by-moment consciousness:

Step 1 At the noumenal level a person's noumenal body, B[r], will transfer information into the wider noumenal world, W[r]. Define the component of that information that is transferred intentionally – i.e. that drives *voluntary* movement – as being *motor output[r]*, ***m***, where ***m*** will have been designed by brain[r] within B[r] to effect a desired change in the person's physical environment, and/or in their location and/or posture in that environment. Say that this information is generated within B[r] over time as a series of segments, or quanta, of motor output[7] – ***m*** *quanta* – which will then be expressed by B[r] into W[r]. Call this expression by B[r] of ***m*** output Step 1 (S1) of the action cycle. Say that as a result of Step 1 of the action cycle the resulting momentary state of B[r] in W[r] is {B[r] in W[r]}^{a} at time (t), t = a.

[5] This rate is a working estimate based roughly on the minimum time interval – about 500 msec – that people appear able consciously to mark (and see footnote 22 below).

[6] How much of the proposed action cycle should ultimately be considered to be sequential, and what segments of it may run in parallel or even as subcycles, will be discussed elsewhere. For now, for the sake of ease of introduction of concepts, assume as a first approximation that the action cycle is fully sequential.

[7] See Appendix 1 to this note for discussion of the use of quantization to describe processes handling motor output and sensory input. Appendix 1 also explains that saying motor output ***m*** and sensory input ***s*** are quantized, as is done here, only approximates what is described in the Main Essay and is used above only to aid easier initial understanding. From a phenomenal perspective, sensory experience seems quite continuous and voluntary physical actions seem quite smooth.

Step 2 Also at the noumenal level, a person's noumenal body, B[r] will receive information from W[r]. Say that the component of that information that will subsequently be perceived consciously – i.e. at the phenomenal level – will be made up of *sensory input[r]*, ***s***, that will be processed inside the person's brain[r] to update B[r]'s model of itself in the world, B[i] in W[i] (consistent with (4), above). Say also that this information, ***s***, is accepted by a person's brain[r] over time as a series of segments, or quanta, of sensory input, ***s*** *quanta*. Call the input of an ***s*** quantum into a person's brain[r] Step 2 (S2) of the action cycle.

Step 3 Consistent with the proposition that a person's brain[r] is running in real-time a model of B[r] in W[r] in the form of B[i] in W[i] – as per (4) to (6) above – say that brain[r] processes the quantum of ***s*** received at S2 to update the model B[i] in W[i]. Call this Step 3 (S3) of the action cycle and call this updated, momentary state of B[i] in W[i], {B[i] in W[i]}^{a}. Under this arrangement the contents of {B[i] in W[i]}^{a} will be causally informed by {B[r] in W[r]}^{a} via ***s***.[8]

Step 4 Now, say that before S1 – at what will be defined below as Step 6 (S6) – the design and creation of an ***m*** output quantum will have taken place through an *unconscious* brain[r] process that *inductively predicts*[9] the impact that this ***m*** output quantum will have on the state of B[i] in W[i]: in other words, through a brain[r] process that will generate an ***m*** output quantum that – when expressed at S1 – will cause a change at the noumenal level, as it drives expression of information from B[r] into W[r], but that will nevertheless have been generated by brain[r] within B[r] on the basis of a predicted impact at the phenomenal level, B[i] in W[i].

To allow this to happen, say that there is an arrangement within brain[r] that has accumulated, and 'learned' – as a person grows from being a physically uncoordinated baby, then a toddler, and then through play as a child on into adulthood – an operationally effective library of *associations* that will allow a quite accurate inductive prediction to be made of the impact that any one of a range of possible ***m*** output quanta is likely to have on B[i] in W[i]. Call the arrangement within brain[r] that contains this accumulated learning the *association matrix*.[10]

Say then, that at the previous S6, an inductively predicted state of B[i] in W[i] – call this predicted state {B[i] in W[i]}^{a*} – will have been generated as the predicted outcome of the ***m*** then expressed at S1 whilst also being held within brain[r] in memory.

Now define Step 4 (S4) as a process where {B[i] in W[i]}^{a}, which has been derived at S3, is compared to {B[i] in W[i]}^{a*}, which was predicted at the previous S6. Say this comparison yields the difference – that is, the error in prediction – between {B[i] in W[i]}^{a} and {B[i] in W[i]}^{a*}; call that error {B[i] in W[i]}^{aE}.

---

[8] In other words, the relationship between {B[r] in W[r]}^{a} as a source of sensory information and {B[i] in W[i]}^{a} will be causal, with ***s*** coming from B[r] in W[r] into brain[r] in B[r] to determine what is reflected in B[i] in W[i], consistent with B[i] in W[i] being a high-fidelity model of B[r] in W[r].

[9] 'Predicts' in the same mechanistic, unconscious sense by which forms of man-made artificial intelligence – for example neural network based machines – can predict the effect an output will have on a system; N.b. In this connection the process described here can be considered a form of *active inference* within the *predictive processing* paradigm: e.g. see Hohwy, J. (2020) New Directions in Predictive Processing. *Mind & Language 35*, 209-223 and references therein.

[10] The proposed concept, properties and operations of association matrices are described in the Main Essay and form a crucial and central part of that essay. This material does not lend itself to a simple summary. For a clearer idea of what an association matrix is, and does, readers are referred to the Main Essay (via http://teleodyne.com) and the supporting Summary Notes (http://teleodyne.com/summary_notes.pdf). Note also that an association matrix can be considered to be a *generative model* within the predictive processing paradigm: e.g. as per the introductory paragraph and references in Parr, T. and Friston, K. (2018) The Anatomy of Inference: Generative Models and Brain Structure *Frontiers in Computational Neuroscience 12*:90. One further point: to the extent that the brain[r] architecture allowing formation of an association matrix[r] is the product of natural selection under Darwinian evolutionary pressure, the associative 'knowledge' held in an association matrix will in fact be considerably more than just what has been 'learned' by any given individual as they interact with W[r] from infancy. Indeed, at the genetically driven level of overall neurological architecture[r] the capacity to form an association matrix will represent a deeper level of 'learning' about the relationship between B[i] in W[i] and B[r] in W[r], with this level of learning having accumulated species-wide and on an evolutionary timescale.

Step 5 Then say that, at Step 5 (S5), brain[r] uses $\{B[i] \text{ in } W[i]\}^{aE}$ to apply an enduring correction to the 'association matrix' described at S4 above. So S5 will be the step where the association matrix 'learns' to further minimize error between all future $\{B[i] \text{ in } W[i]\}^{n*}$ and $\{B[i] \text{ in } W[i]\}^{n}$, where the divergence found in each $\{B[i] \text{ in } W[i]\}^{nE}$ will be between the predictive model of B[r] in W[r] held in the association matrix and the reality of B[r] in W[r] as it acts at the noumenal level.

Now say that steady learning, based on corrections made over successive S5, will incrementally allow the association matrix to be used at subsequent S6 inductively to deliver ever more accurate predictions of how one or another ***m*** output quantum will change the state of B[i] in W[i]. Suppose these predictions then become so accurate that $\{B[i] \text{ in } W[i]\}^{nE}$ approaches nil for extended periods, such as where a person performs familiar sensory-motor tasks in familiar environments.[11] For these series of S5 moments, B[r] information processing systems will be able to sustain that $\{B[i] \text{ in } W[i]\}^{n}$ will remain, to within practical tolerances, *equivalent* to $\{B[i] \text{ in } W[i]\}^{n*}$.

Step 6 As described above, S6 will be the process within brain[r] where a new ***m*** output quantum is designed and created – in line for expression at the subsequent S1. Here a brain[r] process will draw upon the accumulated 'knowledge' held in the association matrix to formulate an ***m*** output quantum that it inductively *predicts* will lead to a specific state of B[i] in W[i], namely $\{B[i] \text{ in } W[i]\}^{(a+1)}$. Say that the brain[r] processing involved in formulating the next ***m*** quantum will need to refer, as a key starting point, to $\{B[i] \text{ in } W[i]\}^{a}$ and then use the association matrix to formulate the required ***m*** output quantum and its associated predicted state $\{B[i] \text{ in } W[i]\}^{(a+1)*}$. [12]

It is proposed that the action cycle will be run through continuously – with each S1 to S6 followed immediately by a subsequent S1 to S6 – at around four times a second, for any period over which the person remains continuously conscious. Call a single iteration of the action cycle from S1 to S6 one *beat* of the action cycle.

For the purposes of discussion, say we can count out as follows from a given arbitrary starting point, t = a, in relation to the sequence of beats (a, a+1, a+2…n) of the action cycle:

$\{B[r] \text{ in } W[r]\}^{a}$, $\{B[r] \text{ in } W[r]\}^{(a+1)}$, $\{B[r] \text{ in } W[r]\}^{(a+2)}$, ... ,$\{B[r] \text{ in } W[r]\}^{n}$

And, correspondingly:

$\{B[i] \text{ in } W[i]\}^{a}$, $\{B[i] \text{ in } W[i]\}^{(a+1)}$, $\{B[i] \text{ in } W[i]\}^{(a+2)}$, ... ,$\{B[i] \text{ in } W[i]\}^{n}$
$\{B[i] \text{ in } W[i]\}^{a*}$, $\{B[i] \text{ in } W[i]\}^{(a+1)*}$, $\{B[i] \text{ in } W[i]\}^{(a+2)*}$, ... ,$\{B[i] \text{ in } W[i]\}^{n*}$
$\{B[i] \text{ in } W[i]\}^{aE}$, $\{B[i] \text{ in } W[i]\}^{(a+1)E}$, $\{B[i] \text{ in } W[i]\}^{(a+2)E}$, ... ,$\{B[i] \text{ in } W[i]\}^{nE}$

A tabulation showing the action cycle is provided at http://teleodyne.com/action_cycleF.pdf

[11] Note that going from infancy – through learning to walk, and to tie shoe laces – to the point where a person becomes competent in a full range of well-coordinated physical actions will involve hundreds of millions of iterations of the action cycle, each providing some level of error correction data to the developing association matrix until its ability for use in prediction is extremely good. This type of learning, albeit at a less sophisticated level, is already being accomplished in artificial intelligence systems involving, for example, neural networks and can be considered a kind of Bayesian learning as this is understood within the predictive processing paradigm; for example, see: Friston, K. (2003). Learning and Inference in the Brain. *Neural Networks 16*, 1325-1352. And more broadly in relation to this see Friston, K. (2010). The Free Energy Principle: A Unified Brain Theory? *Nature Reviews Neuroscience 11*, 127-138, Hohwy, J. *The Predictive Mind* Oxford University Press, Oxford UK 2013 and Clark, A. *Surfing Uncertainty: Prediction, Action and the Embodied Mind* Oxford University Press, New York, USA 2016.

[12] Any $\{B[i] \text{ in } W[i]\}^{n}$ will be a phenomenal state already informed by ***s***, whereas $\{B[i] \text{ in } W[i]\}^{n*}$ will be the correlate – potentially more erroneous – earlier prediction of that state generated within brain[r] along with a correlate ***m*** at the preceding S6. Again, this step corresponds to *active inference* within the predictive processing paradigm, with the association matrix serving as a *generative model*.

Consciousness in relation to the action cycle

It has been held here, and elsewhere[13], that consciousness can arise through the operation of a phenomenal self-model (PSM). For the type of PSM described on this website, and above under 'Introductory material', it has been further proposed that:

7) A subjective, moment-by-moment sense of self-in-the-world may be generated in the course of a process in brain[r] which at a rate of about 4 Hertz realigns a person's B[i] in W[i] (the physical world as they perceive it) to update it back to high fidelity with their B[r] in W[r] (the physical world as they act upon it).

Now consider this assertion in relation to the action cycle. It is likely that for a physically competent person[14], whose association matrix is mature and operationally effective, that the degree of error correction required at S5 will be *very small* for almost all of the physical actions they undertake, particularly those of a routine and well-rehearsed nature.[15]

This means that, for a physically competent person, it will emerge through development and operation of smooth and effective sensory-motor coordination, that $\{B[i] \text{ in } W[i]\}^{n*}$ will – *for all practical purposes* – achieve *convergence* with $\{B[i] \text{ in } W[i]\}^{n}$. In other words, it will be the case that error correction at S5 will show an *operationally negligible* level of discrepancy between $\{B[i] \text{ in } W[i]\}^{n*}$ and $\{B[i] \text{ in } W[i]\}^{n}$ over very extended periods of time.[16]

Importantly, this should allow a range of brain[r] information processing systems to evolve and operate on an *inductive presumption* that $\{B[i] \text{ in } W[i]\}^{n*}$ will *always* be shown at S5 to be *fully converged with* $\{B[i] \text{ in } W[i]\}^{n}$; i.e. on a presumption that $\{B[i] \text{ in } W[i]\}^{nE}$ will always be nil.

In other words information processing systems in B[r] – at the noumenal level of B[r] in W[r] – will be able to use the association matrix highly reliably to 'guess' what the state of B[i] – at the phenomenal level of B[i] in W[i] – will come to be if a given motor output, ***m***, is made at the noumenal level by B[r] into W[r].[17]

Then, over extended periods of time, as B[r] serially acts on W[r] by outputting such ***m***, it will be possible for the ***m*** formulating information processing systems within B[r] inductively to operate on the basis that they – and all of B[r] – *are in identity with* B[i]; that B[i] *is* B[r].

In real time it will happen as follows: at each successive S5 of the action cycle – where the observed $\{B[i] \text{ in } W[i]\}^{n}$ will be shown repeatedly to be virtually identical to $\{B[i] \text{ in } W[i]\}^{n*}$ – B[r] information processing systems will be able repeatedly to refer to the observed B[i] as if it is B[r] since ***m*** output formulated and executed through the *actual agency* of B[r] will repeatedly be manifest – exactly as predicted – through the *apparent agency* of B[i].[18]

Put less formally, at each successive S5 – as a series of actions unfold where B[r] information processing systems can repeatedly refer to B[i] as being B[r] – a series of momentary subjective states will arise along the lines of a subjectively pervasive, felt *foundational assertion*: "that was me that did that". The "me" that is actually making the action is the noumenal B[r], and the "that" which the action has resulted in will be manifest to the relevant B[r] information processing systems as the consequent state of B[i] in W[i]. And when that consequent state of B[i] in W[i] is revealed to be exactly as was inductively predicted by those B[r] information processing systems, the "me" in the felt assertion "that was me that did that" will represent B[r] to itself as actually having been the phenomenal entity B[i].

---

[13] In particular, see Metzinger, T. *Being No One* MIT Press, Cambridge, Massachusetts USA 2003.

[14] A 'physically competent person' is defined here as a person with normal adult – i.e. everyday highly effective – sensory-motor coordination skills.

[15] That is, $\{B[i] \text{ in } W[i]\}^{nE}$ will be so small as to be operationally insignificant during almost all physical activities being undertaken.

[16] Indeed, over many hours, and tens of thousands of iterations of the action cycle.

[17] In terms of the predictive processing paradigm this is to say that, through ongoing implementation of active inference, refinement of the generative model embodied in the association matrix will have reached an extremely high level of predictive proficiency.

[18] The relevant 'equivalences' are provided in a concise form in the note to http://teleodyne.com/action_cycleF.pdf

Thus, as the action cycle proceeds there will be an ongoing series of subjectively experienced "that was me that did that" moments, coinciding with B[r] information processing systems being able to operate for sustained periods – on the basis *inductively* shown at each S5 – that B[i] *is* B[r] and, moreover, that all B[r] information processing systems must therefore actually reside *within* B[i], the phenomenal self. (And we *do* on a sustained basis subjectively locate *all* aspects of ourselves – and not just physical aspects – as being 'within' our 'physical' phenomenal selves, B[i].)[19]

In this way sustained convergence of $\{B[i] \text{ in } W[i]\}^{n*}$ with $\{B[i] \text{ in } W[i]\}^{n}$ will enable brain[r] information processing systems within a given B[r] to *define* B[r] – the person's noumenal body, their body as it acts on the noumenal world – as *being* B[i], the person's phenomenal body as it 'acts on' and exists within their phenomenal world, W[i]. All this despite the fact that – from an information processing perspective and relative to the noumenal level of W[r] – W[i], and *within it* B[i], will actually reside *within* brain[r], which will itself reside *within* B[r], as shown in Figure 2.

This process will provide for the essence of a person's momentary subjectivity and self-perception through enabling their B[r] to operate as if it is both 'inside and outside' itself[20] – to be both itself, B[r], as it predicts and acts, but also serially to experience, observe and monitor itself undertaking those acts, as B[i] – on the basis that B[i] is *one and the same* as B[r].[21]

Specifically, in this way a series of 'subjective moments' will arise at around four times a second, at each S5 of the ongoing action cycle, when it is reaffirmed that $\{B[i] \text{ in } W[i]\}^{n*}$ is in identity with $\{B[i] \text{ in } W[i]\}^{n}$. Successive 'subjective moments' may then, as they arise through successive beats of the action cycle, merge into an experienced continuous stream of subjective self-realization – that is, an experienced continuous stream of *consciousness* – by B[r] of being a physical self-in-the-world in the form of B[i] in W[i].[22,23,24]

Free will

One formulation of the problem of free will is to ask how sensory input can lead to anything other than motor output that is causally determined by the information in that sensory input.

Now consider the set of ideas provided above, which describe how – through operation of the action cycle – moment-by-moment consciousness of the self as a physical body acting on the physical world can arise. The key to understanding how free will can apply lies in clearly understanding that operation of the action cycle does not only involve causal processes but that, at its heart, it also involves an acausal, *semantic* process.[25]

---

[19] This creates the circumstances that lead to the Mind-Body Problem and explains how that problem arises.

[20] For a deeper account of how B[r] can operate as if it is both 'inside and outside' itself see *Dynamics* in Appendix 1 to *The Construction of Phenomenal Time*, the sequel to this note, at https://teleodyne.com/time.pdf.

[21] When processes in brain[r] begin to operate on the inductive basis that B[i] actually *is* B[r] this generates what Thomas Metzinger and others have called the 'transparency' property of phenomenal self-models able to deliver consciousness. Put another way – and crudely – it is the ongoing maintenance of a near-perfect 'illusion' that B[i] is exactly the same thing as B[r] that allows a moment-by-moment sense of consciousness to arise in a brain[r] information processing system that would otherwise be unconscious, whilst hiding from us that our phenomenal experience is of a model, not of raw physical reality.

[22] Subjectively, consciousness of the self as a physical agent seems continuous. But the idea that subjective experience has a rate and may in some ways be semi-continuous is consistent with several observations. For example, if the action cycle ran a lot faster than four times a second, people who make films would have to make the frame rate faster to retain the impression of continuous motion as the film is viewed. Another example is that we cannot consciously perceive the slow movement of the big hand on a clock or see the fast flight of a bullet. But we can certainly perceive the movement of a clock's second hand quite well and, if it has a stepped movement, we can subjectively mark each second as it passes. Using an analogue stopwatch with a stepped movement though, it is impossible consciously to mark the passage of each tenth second and extremely hard consciously to mark a quarter second movement. This rough 'titration' of capacity for subjective awareness of shorter and shorter frequencies of movement, supports the idea that the action cycle will have rate of the order of around four times a second – and see the note under the table at http://teleodyne.com/action_cycleF.pdf. See also Appendix 1 below; and VanRullen, R. (2016). Perceptual Cycles. *Trends in Cognitive Sciences 20*, 723-735 and Benedetto, A., Morrone, M. C. and Tomassini, A. (2019) The Common Rhythm of Action and Perception. *Journal of Cognitive Neuroscience 32*, 187-200.

[23] It is proposed that this is the *process* that generates Thomas Metzinger's 'appearance of a world'. Moreover, the rate of the action cycle will be what anchors our subjective experience in time by defining the operative width (duration) of what we experience as 'now' in that 'appearance of a world', which is the phenomenal world of $\{B[i] \text{ in } W[i]\}^{a}$, $\{B[i] \text{ in } W[i]\}^{a+1}$, $\{B[i] \text{ in } W[i]\}^{a+2}$, …

[24] Note that considerable complementary detail on the information processing hypothesis presented here – above this footnote marker – is provided in the Main Essay. Most of what follows goes beyond what is presented in the Main Essay and is new.

[25] Here a *semantic process* is intended to mean a process based on *semantic relationships* – such as the relationship between the word "boot" and the object it refers to. Note that such relationships are not causal, but are *arbitrarily* assigned purely by association, by some type of information processing system. So, for example, speaking about the world does not and cannot *directly physically cause* any change in the world related to the *meaning* of what is said, since the relationship between a language and the world it refers to is semantic not causal.

Specifically, the formulation of motor output, ***m***, at S6 of the action cycle will be based upon:

- {B[i] in W[i]}^{a}, which will be manifest to consciousness as the phenomenal self-in-the-world at some given moment, t = a;
- the presence of some form of 'motivation' to change {B[i] in W[i]}^{a} into {B[i] in W[i]}$^{(a+1)*}$ (and thence {B[i] in W[i]}$^{(a+2)*}$, {B[i] in W[i]}$^{(a+3)*}$, …) that is designed to get B[r] to a *goal state*[26] – call such a state {B[i] in W[i]}^{G}; and
- the 'knowledge base' described under S4 above as the association matrix.

Say that based on {B[i] in W[i]}^{a}, an optimum {B[i] in W[i]}$^{(a+1)*}$ to aim for is selected by the motivational process in B[r]. Under S6 of the action cycle this selection will yield two products:

- an ***m*** output quantum ready to be expressed at the level of B[r] in W[r] at S1; and
- a record of the predicted state {B[i] in W[i]}$^{(a+1)*}$ to compare at S4 with {B[i] in W[i]}$^{(a+1)}$, the state of B[i] in W[i] that is then actually observed following execution of the ***m*** output quantum.

As discussed earlier, for a physically competent person, it will emerge through operation of smooth and effective sensory-motor coordination, that {B[i] in W[i]}^{n*} will, for all practical purposes, remain converged with {B[i] in W[i]}^{n} over long periods of operation of the action cycle. This will only be possible because – over hundreds of millions of beats of the action cycle – the association matrix will have learned[27] so full and refined a vocabulary of associations between potential ***m*** output quanta and the potential states of {B[i] in W[i]} that each of those ***m*** quanta would deliver if expressed, that for virtually every ***m*** output quantum designed at S6 and then expressed at S1, {B[i] in W[i]}$^{(a+1)*}$ (the predicted outcome of the motor output) will turn out to be practically identical to {B[i] in W[i]}$^{(a+1)}$ (the observed outcome of that motor output).

What is crucial here is to recognize that the 'knowledge' held in the association matrix of the relationship between any given ***m*** output quantum, and what expression of such an output quantum will lead to at the phenomenal level, is *semantic* knowledge.

In other words, what the association matrix will have learned, *through a process of trial and error*, as the infant grows into a toddler – then into a child, and then ultimately to an adult level of competence in sensory-motor coordination – essentially is a *language*. The association matrix will have learned – through trial and error, over years of 'practice' at S5[28] – that if a certain series of ***m*** output quanta are 'said' from B[r] into W[r] at the noumenal level, a certain series of predicted results will reliably be observed – i.e. 'heard' through the senses – at the phenomenal level, consistent with virtually complete convergence between {B[i] in W[i]}^{n} and {B[i] in W[i]}^{n*}.

---

[26] For the purposes of this note, drinking water can serve as an example of a goal, where the motivational input referred to above would arise in brain[r] due to neurological[r] detection of low blood hydration[r] levels. At the point of drinking, the phenomenal state achieved can be defined as a *goal state*, {B[i] in W[i]}^{G}. The process of finding/getting water would therefore entail a series of 'steps in the right direction' amounting to a full phenomenal series, {B[i] in W[i]}^{a}, {B[i] in W[i]}$^{(a+1)}$, {B[i] in W[i]}$^{(a+2)}$, {B[i] in W[i]}$^{(a+3)}$, … {B[i] in W[i]}^{G}. (A deeper description of the information processing arrangements by which such motivational input could arise, could inform the operations of the action cycle, and could be experienced subjectively (in this example, as thirst), is provided in the Main Essay (from page 28 onwards) and there is a supporting, less technical description at http://teleodyne.com/summary_notes.pdf).

[27] That is, 'learned' through repeated error correction adjustments using S5.

[28] And under species-wide Darwinian natural selection, the emergence at the neurophysiological[r] level of the appropriate brain[r] information processing architecture – including a capacity to form and operate an association matrix – will also have been a form of trial and error 'learning'.

So what will be happening during a sequence of competent sensory-motor coordination is that at the noumenal level brain[r], through B[r], will be 'saying' a stream of actions into W[r] that brain[r] accurately predicts – relying on semantic knowledge held in the association matrix – will have a corresponding sequence of specific outcomes at the phenomenal level. Moreover, the outcomes being sought will be due to a 'motivational' input into S6 of the action cycle from elsewhere within B[r] associated with getting B[r] to a goal state. Such motivational input will *not* be part of the sensory input ***s*** but will have arisen from a separate source through some separate process and channel within B[r].[29]

Let us now look at the order of events described above and relate it to the formulation of the problem of free will provided earlier; which was to ask how sensory input into a person's brain can lead to anything other than motor output that is causally determined by the information in that sensory input.

First, there is no reason to assume that what is happening when a person consciously interacts with their environment is a chain of events that *begins* at some arbitrarily chosen moment with a set of sensory input. The action cycle does not have any beginning or end – it operates as an ongoing cycle.[30]

So, let us begin by considering the formulation of a motor output quantum, ***m***, at S6 and its subsequent expression at S1.

This formulation and expression clearly will not be *causally driven* by the ***s*** input quantum received at S2, although it will be *informed* by it. As described above, formulation of an ***m*** output will be made with reference to $\{B[i] \text{ in } W[i]\}^{a}$ – which is determined by ***s*** – but determining what brain[r] 'wants to see' next as $\{B[i] \text{ in } W[i]\}^{(a+1)*}$, $\{B[i] \text{ in } W[i]\}^{(a+2)*}$, $\{B[i] \text{ in } W[i]\}^{(a+3)*}$, … , $\{B[i] \text{ in } W[i]\}^{G}$ – and the relationship between this desired series of phenomenal outcomes and formulation of the ***m*** output phrases[31] needed to get them – will rely upon other factors as well.

In particular, and most importantly, the formulation of these ***m*** output phrases will rely upon the association matrix and all of the *learned associations* that it holds. This knowledge will be *semantic knowledge*, in the form of a full vocabulary of ***m*** output phrases to match any brain[r] motivational impulse to achieve a specific $\{B[i] \text{ in } W[i]\}^{(a+1)*}$, $\{B[i] \text{ in } W[i]\}^{(a+2)*}$, $\{B[i] \text{ in } W[i]\}^{(a+3)*}$, … , $\{B[i] \text{ in } W[i]\}^{G}$. Moreover, these brain[r] 'motivational impulses' will almost always – for routine conscious behaviour – be about brain[r] seeking to drive specific changes in W[r] to allow realization of some goal state or another, $\{B[i] \text{ in } W[i]\}^{G}$.

From this perspective it can be seen that the role of ***s*** input – as $\{B[i] \text{ in } W[i]\}^{a}$ – in the overall ongoing operation of the action cycle will be to provide brain[r] with an ability to *monitor* – and where necessary *correct* – the results of its bid to deliver $\{B[i] \text{ in } W[i]\}^{(a+1)}$, $\{B[i] \text{ in } W[i]\}^{(a+2)}$, $\{B[i] \text{ in } W[i]\}^{(a+3)}$, … through to a goal state, $\{B[i] \text{ in } W[i]\}^{G}$. As such, ***s*** input will not cause or determine ***m*** output, because that output will have been generated semantically.

---

[29] For depth on how such motivational input might arise and be processed into ***m*** output see the Main Essay (page 24-60) at http://teleodyne.com

[30] The action cycle will only begin and end its operation with waking to consciousness and loss of consciousness, respectively. The latter is not relevant to the current discussion. The former is almost always a tapered process of coming to awareness where, again, it is unclear which step of the cycle comes first.

[31] Define an ***m*** output phrase as a short end-on-end sequence of ***m*** output quanta, several-to-many beats of the action cycle long. This is a reasonable idea given that we know that even extended motor output sequences can become quite automatic, similar to the way words are used to make up sentences, and in keeping with the idea of motor output taking the form of a *language* which is, in a sense *spoken* by B[r] into W[r] and which has a semantic relationship to the states of B[i] in W[i] it is crafted to deliver.

To conclude then, immediate personal experience and 'common sense' tell us that in normal circumstances adult human beings – ourselves – drive and shape our physical environment to an overwhelmingly greater extent than we are driven or shaped by it, including by any agency of momentary sensory input.

We drive and shape our physical environment through consciously initiating and guiding actions in order to reach goals. This observation is entirely consistent with the technical explanation provided above, which reveals a semantic – *not causal* – linkage between the formulation of goal seeking motor output and the sensory input that accompanies it. In this important, foundational sense we have free will.

Brendon Hammer
Aug 2011
Rev3. May 2020
Rev4. Oct 2020[32]
Rev5. Apr 2022[33]

Appendix 1 – Contemporary research and the action cycle

Since this note was first posted in 2011 much work has been done in the cognitive and neurosciences that bears on the hypothetical process described above and in the Main Essay as the action cycle.

In particular in mid-1997 when the Main Essay was completed there was virtually no research to support or refute its key illustrative conjectures regarding the idea of an action cycle; namely:

- that sensory input or motor output might neurologically be processed as series of segments or 'quanta'[34], and
- that the sensory-motor system might form part of a *cycle* that could have a rate corresponding to a rate of real-time subjective awareness of the self as a physical being in a physical world.

The overall plan in 1997, as now, has been to get onto the page a description of a hypothetical, overarching information processing architecture and mechanisms able to illustrate how certain key processes might be brought together to deliver in real time a subjective, conscious awareness of the self as a physical being in a physical world.

The concern here has not been to get everything right, especially at the level of specifics. This is currently beyond the reach of overarching projects of the type presented here due to the relatively undeveloped state of knowledge and scientifically demonstrated theory in the cognitive and neurosciences. A more predominant aim has been to provide an encompassing, if conjectural, construct that might help in the overall task of trying to understand what new evidence and ideas in the philosophy, cognitive science and neuroscience of consciousness might mean as they emerge.

Apropos, given the weight of relevant research published in the last decade relevant to the idea of an action cycle, it is useful to look at what impact these recent research findings have on this part of the overall scheme, as presented here and in the Main Essay.

---

[32] This revision simply footnotes more clearly the way the approach described in this note has intersected with the predictive processing paradigm. The revision also makes some small formatting adjustments.

[33] This revision is simply to flag by adding this footnote, and adding to footnote 35, that new material at Working Note A – Part 1 (Section 9.4.1) proposes recasting the action cycle into what is called the *recognition cycle*. Importantly for the purposes of the current note, the recognition cycle plays virtually the same net role as the action cycle and can be substituted for it without undermining any of the key concepts presented here. On top of this, new material at Working Note A – Part 3 (Section 16) provides a more fundamental approach to the question of free will by addressing head on whether free will can be compatible with a deterministic universe. The relevant notes can be accessed at https://teleodyne.com/working_note_A_1.pdf and https://teleodyne.com/working_note_A_3.pdf.

[34] These were respectively labeled in the Main Essay as 'phrases' or 'quanta' in ϕ information and κ information. Pages 56-57 of the Main Essay provide a description of ϕ information content (ϕ incon) and κ information content (κ incon). Importantly, ϕ incon quanta and κ incon quanta are different to what is described in the above note on Free Will as ***m*** quanta and ***s*** quanta, the latter being an approximation to help get the larger ideas presented here across as simply as possible. In reality ***m*** output is smooth and ***s*** input is continuous as it falls upon the B[r] interface with W[r]. This difference can be handled through the processing architecture proposed in the Main Essay, specifically by the processors labelled P_0 and P_1, which would *inter alia* respectively perform both smoothing and bundling functions. Evidence for smoothing of conscious perception of moving objects where the visual input is shown to be sampled through discrete time windows is provided in Schneider, K. A. (2018) The Flash-Lag, Frölich and Related Motion Illusions Are Natural Consequences of Discrete Sampling in the Visual System. *Frontiers in Psychology 9* 1227.

Impact of recent research findings

In 1997 – without benefit of experimental evidence – it in any case seemed necessary to segment motor and sensory information flows in order to construct a stepwise action cycle. Describing error correction and updating processes able to ensure that the phenomenally experienced world of B[i] in W[i] could serially be reset in real time to keep up with the noumenal world of B[r] in W[r] seemed analytically too difficult using a continuous process model. Moreover, it was apparent that approaches based on quanta could often *subsequently* be brought to approximate continuous behaviour or be transformed into continuous approaches.[35]

Regardless, there is now a large body of evidence to show that the neurological systems handling both sensory and motor processes involve segmented information flow.

Much of this work has been summarized in reviews by VanRullen (2016) and Benedetto et al. (2019).[36]

In relation to the idea of an action cycle, the most interesting observations to come out of these reviews are that:

- Neural rhythms at various frequencies – measurable in normal brains using single-neuron studies through to whole-brain techniques such as EEG and MEG – play a crucial role in sensory, cognitive and motor processing, and probably correspond to what VanRullen has called 'perceptual cycles'
    - In at least some sensory processing, temporal parsing serially separates sensory information into discrete epochs.
- For vision on its own, and perhaps for other senses on their own – when no sensory-motor coordination or motor action is being undertaken – a subject's experience of phenomenal (consciously experienced and reportable) effects indicates a ~10 Hz processing cycle.
    - But when motor output is undertaken at the same time as sensory processing the observed frequency of 'sensory-motor synchronization' is around 3-7 Hz.
- Benedetto et. al (2019) point to experimental evidence that visual performance shows rhythmic 3-7 Hz periodicity 'time-locked' to hand movement, and that this action-locked perceptual rhythmicity emerges well before movement onset, suggesting an 'automatic' coupling between visual processing and motor planning.
    - They observe that this coupling is established independently of any intrinsic link between the motor and sensory systems and is present during arbitrary sensory-motor contingencies; that is, concurrently performed but functionally unrelated visual and motor tasks.
    - They also observe that, "the rhythmic [3-7 Hz] coupling between the visual and motor system is not invariant to the current motor state, suggesting that it may structure dynamically the functional interplay between the two systems, enabling timely incorporation of sensory information within the ongoing motor plan".

These results and insights support the idea of a cyclic process operating at around 4 Hz that recruits sensory, motor and possibly wider cognitive processes and which entails segmented information flows broadly consistent with the idea of an action cycle.

---

[35] For example, in relating quantum mechanics to classical mechanics. (And now see Section 9.4.1 in Working Note A – Part 1 at https://teleodyne.com/working_note_A_1.pdf.)

[36] VanRullen, R. (2016). Perceptual Cycles. *Trends in Cognitive Sciences 20*, 723-735, and Benedetto, A., Morrone, M. C. and Tomassini, A. (2019) The Common Rhythm of Action and Perception. *Journal of Cognitive Neuroscience 32*, 187-200.

VanRullen (2016) suggests that the occurrence of brain rhythms at multiple frequencies refutes the idea of a central sampling rhythm simultaneously affecting all aspects of perceptual experience. As an alternative he proposes coexistence of multiple perceptual cycles in distinct brain networks with different periodicities. This seems entirely reasonable but raises the question of how these can then be brought together into the arena of a unified phenomenal experience, and an overall conscious experience of being and acting in the physical world.

The observations of Benedetto et.al (2019) may prove helpful through indicating that when the fast visual processing rhythm is recruited into a process that also recruits motor processing rhythms – and possibly wider brain rhythms – a lower dominant frequency is observed. This would be consistent with the idea of a 4 Hz action-cycle like process forming part of the system.

In relation to people's everyday experience of a continuous rather than punctuated phenomenal world, VanRullen (2016) explores a number of possibilities, suggesting that at least for vision such continuous experience – despite segmentation of sensory information flows during relevant neural processing – may be explained by invoking a number of possible mechanisms, including 'apparent motion' and 'echoing'. The ideas around 'dynamic perceptual completion' discussed in Gruber et. al (2019) have relevance.[37] Such approaches can be applied to explaining continuous perception and motor expression in circumstances where an action-cycle like process might be present. Of particular relevance here is the work of Lee (2014) and Grush (2005, 2016)[38], which is more closely considered in the sequel to this note, *The Construction of Phenomenal Time*.[39]

None of this is to argue that an action cycle in the specific form proposed in the note above, or in the Main Essay, is what is actually happening at the level of real-world detail rather than at the level of a general concept. But what this recent neuropsychological research does show is that the idea of an action-cycle-like process operating at around 4 Hz might remain a helpful bearing point in the hunt for better specified information processing architectures able to deliver subjective experience of being a physical self in a physical world.

[37] Gruber, P.R., Smith, P.R. and Block, R.A. (2019) Dynamic perceptual completion and the dynamic snapshot view to help solve the 'two times' problem. *Phenomenology and the Cognitive Sciences*.

[38] Lee, G. (2014) Temporal Experience and the Temporal Structure of Experience. *Philosophers' Imprint 14* (3), Grush, R., (2016) On the Temporal Character of Temporal Experience, its Scale Non-invariance, and its Small Scale Structure *Manuscript. doi: 10.21224/P4WC73* and Grush, R., (2005) Internal Models and the Construction of Time: Generalizing from State Estimation to Trajectory Estimation to Address Temporal Features of Perception, Including Temporal Illusions. *Journal of Neural Engineering 2*, 209-218.

[39] *The Construction of Phenomenal Time* can be found at https://teleodyne.com/time.pdf.