



THE QUESTION OF NATURALIZING PHENOMENOLOGY

James Mensch (Charles University, Prague)

The attempt to use the results of phenomenology in cognitive and neural science has in the past decade become increasingly widespread. It is, however, open to the objection that phenomenology does not concern itself with the embodied, empirical subject, but rather with the non-causally determined “transcendental” subject. If this is true, then the attempt to employ its results is bound to come to grief on the opposition of two different accounts of consciousness: the non-causal, transcendental paradigm put forward by phenomenology and the causal paradigm assumed by cognitive and neural science. In what follows, I shall analyze this objection in terms of the conception of subjectivity the objection presupposes. By employing a different conception, I shall then show how it can be met. My aim will be to explain how we can empirically use the insights of phenomenology without denaturing the consciousness it studies.

One of the most remarkable developments of the past decade has been the attempt to marry phenomenology to cognitive science. Perhaps nothing else has so revitalized phenomenology, making it a topic of interest in wider philosophical and scientific communities.¹

¹ Representative works in this area include Francisco Varela, Evan T. Thompson and Eleanor Rosch, *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge: MIT Press, 1991); Francisco Varela, “Neurophenomenology: A Methodological Remedy to the Hard Problem,” *Journal of Consciousness Studies*, vol. 3, no. 4 (1996), 330–50; Francisco Varela, “The Naturalization of Phenomenology as the Transcendence of Nature. Searching for Generative Mutual Constraints,” *Alter*, vol. 5, no. 4 (1997), 355–81; Shaun Gallagher, “Mutual Enlightenment: Recent Phenomenology in Cognitive Science,” *Journal of Consciousness Studies*, vol. 4, no. 3 (1997), 195–214; J. Petitot, F. Varela, B. Pachoud and J.-M. Roy, eds., *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science* (Stanford: Stanford University Press, 1999); Shaun Gallagher, “Phenomenology and Experimental Design: Toward a Phenomenologically Enlightened Experimental Science,” *Journal of Consciousness Studies*, vol. 10, no. 9–10 (2003), 85–99; Antoine Lutz and Evan Thompson, “Neurophenomenology: Integrating Subjective Experience and Brain Dynamics in the Neuroscience



The reasoning behind this initiative is relatively straightforward. Cognitive science studies artificial and brain-based intelligence. But before we can speak of artificial intelligence, we must have some knowledge of natural intelligence; that is, we must understand our own cognitive functioning. Similarly, to understand how the brain functions, we need to grasp the cognitive processes that such functioning realizes. This is precisely what phenomenology provides: it studies the cognitive acts through which we apprehend the world, observes the constitutive buildup of such acts, and attends to the temporal constitution at work in the genesis of every act, every intentional relation we have to the world. Its results, which have been accumulating since the beginning of the last century, thus offer cognitive science a trove of information for its projects.

As obvious as this conclusion appears, it is not immune to some fundamental objections. The chief of these is that phenomenology does not concern itself with the real, psychological subject, but rather with the “transcendental” subject. By virtue of the *epoché* that reveals it, this subject, as Husserl writes, “loses that which gives it the value of something real in the naïvely experienced, pre-given world; it loses its sense of being a soul of an animal organism which exists in a pre-given, spatial-temporal nature.”² As such, the transcendental subject no longer has its sense of being causally determined by this spatial-temporal nature. Given this, how can such a subject serve as a paradigm for understanding either artificial or organic, brain-based intelligence? As part of the world, the latter are causally determined structures, but the transcendental subject, as Husserl asserts, has to be “considered as absolute in itself and as existing for itself ‘before’ all worldly being.”³ If this is true, then the attempt to marry phenomenology with cognitive science is bound to come to grief on the opposition of different accounts of consciousness: the non-causal, transcendental paradigm put forward by phenomenology as opposed to the causal one assumed by cognitive science. Any attempt to employ phenomenology in cognitive science must, then, reinterpret its results in terms of the causal account

of Consciousness,” *Journal of Consciousness Studies*, vol. 10, no. 9–10 (2003), 31–52; and Shaun Gallagher and Dan Zahavi, *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science* (London and New York: Routledge, 2008).

² Edmund Husserl, “Nachwort,” in *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Drittes Buch*, (ed.) M. Biemel (The Hague: Martinus Nijhoff, 1971), 145. Except where noted, all translations from German to English are the author’s.

³ *Ibid.*, 146.

assumed by the latter. In its attempt to use transcendental consciousness as a paradigm for understanding artificial or brain-based intelligence, cognitive science must transform it into a part of nature. This naturalization of consciousness is, in fact, a denaturing of it. It is a transformation that makes us lose sight of what is essential to consciousness.

In what follows, I shall examine this objection in order to reveal the concept of subjectivity that it presupposes. By employing a different conception, I shall then show how this objection can be met. My aim will be to explain how we can use the insights of phenomenology without denaturing the consciousness it studies.⁴ This will involve specifying the relation between transcendental subjectivity and its empirical, actually functioning analogue.

§1. The Non-Worldly Subject

A good way to introduce the objection is to note that phenomenology is often distinguished from the natural sciences through its essentially critical function, which it inherited from Kant. As Iso Kern notes, Husserl lectured on Kant throughout his career.⁵ Not only does Husserl equate his notion of constitution with that of Kant's synthesis⁶, he also identifies the project of the "new science" that Kant advanced with that of phenomenology.⁷ This science is critical. Thus,

⁴ Here, let me acknowledge Mark W. Brown's Ph.D. thesis, "Naturalizing Phenomenology: An Essay on the Phenomenological Limits of Naturalizing Phenomenology" (Murdoch, Australia: Murdoch University, 2008), from which I have drawn some of the formulations of this objection. My expression of them is my own, as is my attempt to meet them.

⁵ Iso Kern gives the list of Kant's lectures in his excellent study, *Husserl und Kant* (The Hague: Martinus Nijhoff, 1964), 425–27. The copies of the Critique of Pure Reason that Husserl read (as shown by the annotations and underlinings) are those edited by Hartenstein (1867–68), Kehrbach (1878), and Vorländer (1899). See Kern, *Husserl und Kant*, 428–29.

⁶ In Husserl's words: "What is called constitution, this is what Kant obviously had in mind under the rubric, 'connection as an operation of the understanding,' synthesis." (MS, B IV 12, 2–3, 1920) I am grateful to the Director of the Husserl Archives in Leuven, Belgium for permission to cite from the unpublished manuscripts.

⁷ He writes, for example, "...the revolution in the very nature of philosophical thought which Kant promoted and allowed to arise in the powerful, perhaps even violent proposal of a new science is still the challenge of the present; and this new science is our own task and a task which can never be abandoned in all the future." (Edmund Husserl, "Kant und die Idee der Transzendentalphilosophie," in *Erste Philosophie [1923/1924], Erste Teil, Kritische Ideen-*

while the natural sciences, including cognitive science, aim at the understanding and manipulation of the external world, phenomenology's focus is on the subjective performances that generate the natural sciences. It is not interested in contributing to their results, but rather in understanding how these results are achieved. This is why it examines the evidence that a science has for its basic positions and shows how the basic sense structures that characterize the science relate to such evidence. Phenomenology's critical function is that of limiting a science's claims to the sense structures that its evidence supports. Against this, one could argue that one may grant this critical focus without prejudicing the use of phenomenology by the cognitive sciences. One could assert that no matter what the focus of phenomenology is, its results stand on their own. As genuine insights into our cognitive processes, their validity does not depend upon the particular aims of phenomenology, but only on the accuracy of its accounts. This holds, in particular, for its descriptions of the subjective performances that characterize consciousness as such. Such performances are carried out by the same subjectivity that the cognitive and neurological sciences study, namely, our embodied subjectivity. Thus, given that *their subject matter is the same*, the results of phenomenology can be used by cognitive science.

This, however, is precisely what the objection will not allow. Following Husserl's lead, it asserts that what phenomenology studies is not part of the world. It does not investigate the empirical subject, but rather the non-worldly, transcendental subject. As prior to the world, this subject cannot be described in worldly terms. To deny this is to deny the point of Husserl's *epoché*. This *epoché* is what first allows us to do phenomenology. We perform it when we suspend our belief in the natural world. Such a suspension is not a denial, but rather an attempt to examine with unprejudiced eyes the evidence we have for it. This means that we cannot avail ourselves of any

geschichte, [ed.] R. Boehm [The Hague: Martinus Nijhoff, 1956], 240.) This revolution is Kant's proposal of "a transcendental, scientific theory of the essential possibility of the constitution of a true objectivity in transcendental subjectivity" ("Kants Kopernikanische Umdrehung," in *ibid.*, 227). As Husserl elsewhere expresses this, Kant "brought about the recognition that the world, which is for us, only exists for us in our cognition and that the world for us is nothing but that which, under the title of objective knowledge, takes shape in our experiences and thought." (MS, F I 32, 114a, 1927, "Natur und Geist") This affinity does not mean that Husserl did not have sharp disagreements with Kant. For an account of these, see James Mensch, *Intersubjectivity and Transcendental Idealism* (Albany: State University of New York Press, 1988). 106–25; and Kern, *Husserl und Kant*, 55–134.

thesis that presupposes the existence of this world, and this includes all the theses of natural science. Rather than employing these, phenomenology's focus is on the evidence we have for them.⁸ In pursuing such evidence, phenomenology discovers the transcendental subject, the subject that constitutes the sense of the world from such evidence. We thus have the aforementioned distinction between the transcendental and the empirical subject. The empirical subject presupposes the world; the transcendental subject does not. Descriptions of the former employ causal terms; those of the latter avoid them since their use presupposes the causal intertwining of subjectivity and the world and, hence, assumes the thesis of the natural world. Given this, we cannot explain the transcendental subject by referring to our embodied empirical subject. In fact, the explanatory relation is the reverse. Phenomenology reveals the empirical subject as a sense structure constituted by the transcendental subject from the evidence available to it.

The conception of transcendental subjectivity at work here is that of a non-worldly agent that is somehow prior to and, hence, independent of the world. This agent is both conscious of the world and constitutes it as the object of its consciousness. There are certainly passages in Husserl that would seem to support this interpretation. Husserl, in the *Cartesianische Meditationen*, asserts that "transcendental subjectivity...constitutes both sense and being."⁹ He writes that the project of phenomenology is to make "every being itself, be it real or ideal,...understandable as a constituted product [*Gebilde*] of transcendental subjectivity."¹⁰ As for the world, it is described in *Die Krisis* as "a world whose being is being from subjective performances, and this with such evidence that another world is not thinkable at all."¹¹ Taken literally, such remarks would position transcendental subjectivity as a God-like agent, one that created the world from nothing. Against such a view is a fact that Husserl maintains

⁸ As Roman Ingarden observes, the logical point of the *epoché* is to avoid the fallacy of the *petitio principii*, i.e., of assuming as part of the evidence for a thesis something that presupposes this thesis. To do so is to "beg a principle" and assume what one was trying to prove. See Roman Ingarden, *On the Motives which Led Husserl to Transcendental Idealism*, (tr.) A. Hannibalsson (The Hague: Martinus Nijhoff, 1975), 12.

⁹ Edmund Husserl, *Cartesianische Meditationen*, (ed.) S. Strasser (The Hague: Martinus Nijhoff, 1963), 118.

¹⁰ *Ibid.*

¹¹ Edmund Husserl, *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*, (ed.) W. Biemel (The Hague: Martinus Nijhoff, 1962), 100.

throughout his career, namely, that consciousness is dependent in its functioning on externally provided data. Such data, he asserts, “are nothing produced by consciousness. They are...that which has come into being alien to consciousness, that which has been received, as opposed to what has been produced through consciousness’s own spontaneity.” This means that “consciousness is nothing without impression.”¹² Externally provided impressions are the source of all its contents. As for the subject or ego of consciousness, it is, apart from such impressions, “an empty form,” one that is individualized through the contents provided by the impressions.¹³

§2. Consciousness as a Function

This dependence points to a second, non-idealistic view of Husserl’s position. Not only can it be supported by a large number of texts, but it also allows us to use his insights in cognitive science. Its conception of consciousness appears, for example, in the second book of the *Ideen*, where Husserl shows how our embodiment determines our consciousness.¹⁴ In this conception, it is the real embodied subject that is conscious of the world. As for transcendental subjectivity, this is not to be considered an agent. It, rather, designates a set of functions that the real embodied subject engages in. This means that we have to qualify the assertions of the *Cartesianische Meditationen* and take them as referring to sense and being *for us*. What is at issue is not being itself, but rather our action of positing being from the evidence it provides us. We do this by making sense out of a given material. This making (or constituting) sense is a many-layered process. Essentially, it is a matter of identifying a sense as a one-in-many and positing this unity as a common referent for an originally apprehended multiplicity, be this a multiplicity of impressions, perceptions, perceptual objects, or states of affairs composed of these. The actual, embodied subject engages in this activity. As for transcendental subjectivity, it is not a particular entity either within

¹² Edmund Husserl, *Zur Phänomenologie des inneren Zeitbewusstseins (1893–1917)*, (ed.) R. Boehm (The Hague: Martinus Nijhoff, 1966), 100.

¹³ “One can say that the ego of the cogito is completely devoid of a material, specific essence, comparable, indeed, with another ego, but comparable only as an empty form that is “individualized” through the stream: this, in the sense of its uniqueness.” (MS, E III 2, 18, 1921, Husserl Archives, Leuven, Belgium)

¹⁴ See, e.g., Edmund Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Zweites Buch*, (ed.) W. Biemel (The Hague: Martinus Nijhoff, 1952), §§35–42.

the world or prior to it, but rather, it refers to the structure of this synthetic, interpretive function of identifying unities in multiplicities.

This function, we should note, is responsible for the transcendental ego understood as an “empty form.” As James Edie expresses this view of this ego, it is not an actual experiencing ego. It is, rather, “an impersonal, necessary, universal, eidetic structure,” one that “is lived in and through each unique consciousness, each ego-life.”¹⁵ The *C Manuscripts* of the 1930s take this structure as the centring of our conscious life. In Husserl’s words, “The ego is the ‘subject’ of consciousness; subject, here, is only another word for the centering which all life possesses as an egological life, i.e., as a living in order to experience something, to be conscious of it.”¹⁶ As he elsewhere expresses this: “I am I, the center of the egological [*Ichlichkeiten*].”¹⁷ The ego, in other words, is the centre of the centring of our consciousness.¹⁸ As such, it exists as the zero-point of our life. Thus, each of us, when we regard ourselves in terms of our experience, always find that we are “here.” This means that we cannot depart from ourselves, but always find ourselves at the spatial centre or zero-point of our environments. This is the point from which the “near” and the “far” are measured. Phenomenologically regarded, this “here” is defined by the perspectival unfolding of the objects that surround us as we move through the world. The sides the objects show all point to us as a centre. We experience the different rates of their perspectival unfolding as exhibiting their different distances from us. As the familiar experience of gazing from a moving car window shows, objects we take as close by have a higher angular rate of turning than those that we apprehend as further away. This sense of space with its correlative zero-point depends, of course, on our apprehending time. The unfolding perspectives of our surrounding objects cannot vanish the moment after their apprehension. Retention is required to grasp the rate of their unfolding and protention (or anticipation) is needed if we are to make use of what we

¹⁵ James Edie, “The Question of the Transcendental Ego: Sartre’s Critique of Husserl,” in *Husserl in His Contemporary Radiance: Proceedings of the 24th Annual Meeting of the Husserl Circle* (Waterloo, 1992), 271–72.

¹⁶ Edmund Husserl, MS, C 3, 26a, March 1931, in *Späte Texte über Zeitkonstitution (1929-1934): Die C-Manuskripte*, (ed.) D. Lohmar (Dordrecht: Springer Verlag, 2006), 35.

¹⁷ Edmund Husserl, MS, C 7, 9b, June–July 1932, in *Späte Texte über Zeitkonstitution*, 122.

¹⁸ As such a centre, it is also called a *pole*. In Husserl’s words, “The *central ego* is the necessary *ego pole* of all experience.” (MS, M III 3, XI, 21, September 1921, Husserl Archives, Leuven, Belgium) See also Husserl, *Ideen, Zweites Buch*, 105.

retain to make our way in the world. Retaining and protending their relative rates of unfolding, we locate ourselves in the world. This locating is not just spatial, but also temporal. Situated between the retained past and the anticipated future, we find ourselves at a temporal zero-point. Given that the content that we retain and protend positions us spatially, the “primal now” of this zero-point is always accompanied by a “here.” We thus constantly take ourselves as a spatial-temporal centre. This is our “empty form” as an ego.

To see how the temporal aspect of this form is set by transcendental subjectivity, we have to turn to Husserl’s account of the constitution of “a lasting and remaining now.” He writes that “it is constituted as a fixed form for a content which streams through it and as the source point for all constituted modifications.” This constitution of “the fixed form of the primally welling primal now” occurs along with the constitution of “a two-sided continuity of forms that are just as fixed...the continuum of what is just past and that of futurities.”¹⁹ These situate our temporal zero-point between them. As a result, the now appears as a “fixed form,” *through which* time appears to flow and *in which* its moments appear to well up as present and actual. It *appears*, in other words, as a “primal welling middle point” of time. I stress the word “appear” since the source of time, for Husserl, is not the ego, but rather externally received impressions.²⁰ The ego’s constitution as a temporal zero-point between the retained past and the protended future makes the registering of these impressions *appear* as the welling up of moments from this zero-point. Thus, as Husserl immediately adds, the real result of the process of retention and protention is the constitution of “a stationary and remaining form-continuity [*Formkontinuität*] for what streams through it, which is always co-constituted as streaming.”²¹ This form-continuity

¹⁹ “In diesem Strömen ist ein stehendes und bleibendes Ur-Jetzt als starre Form für einen durchströmenden Gehalt konstituiert und als Urquellpunkt aller konstituierten Modifikationen. Konstituiert ist aber in eins mit der starren Form des Urquellenden, Ur-Jetzt eine zweiseitige Kontinuität von ebenso starren Formen; also im Ganzen ist konstituiert ein starres Kontinuum der Form, in dem das Ur-Jetzt urquellender Mittelpunkt ist für zwei Kontinua als Zweige der Abwandlungsmodi: das Kontinuum der Soebengewesenheiten und das der Zukünftigkeiten.” (MS, C 2, 11a, September–October 1931 in Husserl, *Späte Texte über Zeitkonstitution*, 8)

²⁰ Husserl, in fact, never abandons this point. As he writes in 1931: “Die Uraffektion...muß schon strömend im Gange sein, damit der Einsatzpunkt als prätemporaler Punkt der passiv konstituierten prätemporalen Zeitstrecke konstituiert sein kann.” (MS, B III 3, 4a, 1931, Husserl Archives, Leuven, Belgium)

²¹ The passage continues, “Dies aber ist eine stehende und bleibende Formkontinuität für das sie Durchströmende, als durchströmend immerzu

is simply that of the centring of experience about the now. The central ego, considered as an eidetic structure, is the centre of this form-continuity. Since the form-continuity is such for a temporally streaming material, this centre's constitution always occurs together with the constitution of this material—a constitution that involves placing the ego in time through retention and protention and, hence, making it a centre relative to this streaming material. As the material streams through this centre on its way from the anticipated future to the retained past, the material appears to well up from this centre as if from its source.

We need not go into the constitution of the spatial aspect of the form of an ego—a constitution that essentially involves the perspectival situating of ourselves at a visual zero-point—to see the relation between the transcendental ego and transcendental subjectivity. The form that designates this ego is a result of the functioning whose structure is given by transcendental subjectivity. Now, what functions is the actual embodied human being. He or she is the agent of temporalization. To state the obvious, a human being is not, *qua* embodied, a spatial-temporal zero-point. He or she is not a Cartesian non-extended ego. Our constitution of space and time, however, is such that we always locate this functioning in the zero-point of the here and the now. To use a Kantian phrase, this is a necessary transcendental illusion. We only become trapped by it, however, when we assume that such a non-extended, non-worldly subject is a *res cogitans*, *i.e.*, is something actually functioning, actually conscious. When we avoid this illusion, we can see the import of Husserl's question, "is not consciousness function...?" He continues, "What is necessary?...We have to examine [intentional] experiences as functions.... We have to ask ourselves: What is 'accomplished' in them? What kind of sense is present in them, what kind of sense is progressively forming itself in them?...How do functions synthetically, teleologically unite into the unity of a function, etc.?"²² The call, here, is to

Mitkonstituiertes; und im Durchströmen dieser Form ist eine wundersame Synthesis in beständigem strömendem Gang, in der sich als individuelles Sein konstituiert, was jetzt urquellend auftritt, was, das Formensystem der Soeben durchlaufend, immerfort dasselbe verbleibt, aber dasselbe in kontinuierlich anderen Modis des Soeben." (MS, C 2, 11a, September–October 1931 in Husserl, Späte Texte über Zeitkonstitution, 8)

²² The extended quote is: "...ist nicht Bewußtsein Funktion...? Was ist also notwendig? Es sind intentionale Erlebnisse, Erlebnisse als Funktionen, als relativ geschlossene Funktionen betrachten, sie betrachtend nachleben, neu durchleben, Akte vollziehen und sie wiederholend nachvollziehen und sich dabei befragen, was darin 'geleistet' wird, was für Sinne darin liegt und sich fortgestaltet, was man

examine the functions that allow us to form and progressively unite senses into greater and greater wholes—greater and greater unities in multiplicities. The intentional experience considered as such a function is that of taking our experiences as experiences *of* some common referent, a referent that stands as a unity for their multiplicity. Of course, at the basis of all such functions, as the Bernau and C-manuscripts show, is that of temporalization. While it goes on within the depths of our organic being, it can, as a function, be formally described.

§3. Formalization and Scientific Reductionism

For an example of this description, we can turn to the process of retention. According to Husserl, this short-term memory, which lasts barely a minute, is built up of a chain of retentions of retentions of...some original content.²³ Thus, first there is an impressional consciousness, then this is retained, then this retention is itself retained, and so on for up to a minute. We can mathematically model this process by using parentheses to express retention. The use of parentheses surrounding parentheses would then express the retention of a retention. Thus a simple model of the retentive process would be given by the series: *i*, (*i*), ((*i*))..., each later member being taken as a retention of the earlier. A computational algorithm can be written for this model.²⁴ A corresponding algorithm can also be

dabei tut und was dadurch für Sinnesleistung geleistet wird im Übergang zu den umfassenden Zusammenhängen in der Einheit des Lebens, wie Funktionen mit Funktionen sich zur Einheit einer Funktion synthetisch teleologisch einigen, usw." (MS, A VI 31, 19a, 1920, Husserl Archives, Leuven, Belgium)

²³ In Husserl's words, the retention of the original content "changes into retention of retention and does so continuously." The result is that "a fixed continuum of retention arises in such a way that each later point is a retention for every earlier point." (Husserl, *Zur Phänomenologie des inneren Zeitbewusstseins*, 29)

²⁴ The function is:

```
(Defun Retention (X Impression)
  (cond ((= X 0) Impression)
        (T (Retention (- X 1) (list Impression)))).
```

"Defun" means define the function. "Retention" is the name of the function. Its variables are "X" and "Impression." "X" stands for the number of retentions the impression is to undergo. "Impression" stands for the impression to be retained. The second line states a condition for the computation. If X is equal to 0, *i.e.*, if the number of retentions required is zero, the function returns the impression and the computation ceases. Otherwise, it proceeds to the third line. The "T" tells it to perform the computation which follows to the right. First, 1 is subtracted from X—*i.e.*, the number of required retentions is reduced by one through the

written for the process by which we retain successive contents in the order we experience them. Feeding into it the successive contents, A B C D E, it would express their order through sets of increasing parentheses: (E (D (C (B (A)))))).²⁵ Now, to employ this model in a naturalistic account of consciousness would be to use these increasing parentheses as temporal tags for the successive contents. In terms of artificial intelligence, this would involve accumulating the data provided by the machine's transponders into defined temporal phrases for its processing. As for brain-based intelligence, one would use this model to investigate the sequences of neuronal firing. Generally, whatever the model, the focus would be on the implementation of a given task: in this case, the task of responding to the world's temporal givenness.

The formalization involved in the above example is typical of the procedure employed by cognitive scientists in their attempt to naturalize phenomenology. An analysis of the invariant structures of these performances is followed by the development of a mathemati-

instruction “(- X 1).” Then, the impression to be retained is surrounded by parentheses through the instruction “(list Impression).” Finally, the original function is called again through the instruction “(Retention (- X 1) (list Impression))”. The variables of this function, however, *have been transformed through the first two operations just specified*. For example, if “X” was originally given the value 3, the first operation reduces it to 2. If “Impression” was given the value i, the second operation gives it the value (i). Thus, the call to the original function, “Retention,” is a call for it to carry out its computation on a set of values arrived at through the results of its previous computation. This iterative process continues with 2 being reduced to 1 and then to 0 and (i) being transformed to ((i)) and then to (((i))). When X is 0, then the second line tells it to stop and return the value that “Impression” now has—that is, (((i))). Thus, (Retention 3 ‘i) yields (((i))). This signifies that “i,” the impression, has sunk back to a retention of a retention of a retention of “i.”

²⁵ The function's arguments are: “phrase”—*e.g.*, A B C D E—a given “initial element”—*e.g.*, A—and “X” which signifies the number of retentions. The function is:

```
(defun phrase-retention (phrase initial-element X)
  (cond ((equal nil (cdr phrase)) (Retention (- X 1) initial-element))
        (T ( phrase-retention (cdr phrase) (cons (cadr phrase) (list initial-element)
          (- X 1))))).
```

Here, (phrase-retention ‘(A B C D E) ‘(A) 10) yields ((((((E (D (C (B (A)))))))))). Of the 10 retentions, five are used to retain the phrase and result in (E (D (C (B (A))))). Five more occasion the sinking down of the phrase as a whole five further degrees of pastness. The third line of the algorithm calls up the initial function “phrase-retention” which means that the function processes the results of its previous operation. Within this reprocessing there is also a call to the function “Retention,” given in footnote 24, which also reprocesses the results of its previous operation.

cal model for such constructions. Expressing this in mathematical algorithms, the cognitive scientist employs the latter in a naturalistic account of consciousness.²⁶ Now, in assuming that consciousness is a function, we preclude an objection that can be brought against the attempt to capture it mathematically. This is that consciousness, taken as a concrete entity, is not itself mathematical. The experiences that compose it are not mathematical idealities, but rather the concrete qualitative contents—the sights, sounds, tastes, smells, and textures—of our daily lives. This point is undeniable. But equally certain is the fact that we can describe a function in mathematical terms without asserting that what performs this function is itself mathematical. To assert this would be to ontologize the description, making it the “true being” or reality of the thing described. This is the error that *Die Krisis* denounces in its account of modern science. Post-Galilean science takes its equations as expressing the reality of the world. At bottom, its error is the simple one of substituting the description for the thing described. Just as the law of gravitation is not the gravitating bodies whose relations it describes, so a mathematical relation is not itself the things it relates. This point, however, does not hold with regard to functions. Such functions are not things but rather formal relations and processes. As such, their reality can be caught mathematically. To call consciousness a function, as Husserl does, is to assert that what is essential to it is not the concrete qualitative content of the experiences composing it, but rather the performances an actual subject engages in with regard to such material.

There is a second, related objection against the use of mathematics to describe consciousness. It is based on a specific interpretation of the scientific method, one involving science’s conception of causality. The conception arises from Galileo’s separation of the primary from the secondary qualities of matter. The primary are its quantifiable aspects, such as its measurable distances, speeds, wavelengths, weights, and so on. The secondary are its qualitative aspects, *i.e.*, the tastes, sounds, colours, and so on that our five senses reveal. Natural science, following Galileo, takes causality as pertaining to the quantifiable aspects of the material world. This means that before we can causally describe the world, we must first reduce its secondary qualities to its primary. Sound has to be understood as the frequency

²⁶ For a detailed discussion of this procedure, see David Marr, *Vision* (San Francisco: W. H. Freeman, 1982), 25ff.; and Terence Horgan and John Tienson, *Connectionism and the Philosophy of Psychology* (Cambridge: MIT Press, 1996), 24–45.

of the pressure waves reaching our ear, colour to the wavelengths of light, and so forth. It is only in terms of such measurable qualities that we can mathematically formulate causal relations. The advantage we gain from this is not just greater precision. It is also the ability to make objectively verifiable (third-person) claims based on objectively measurable data. The same cannot be said of the secondary qualities of matter. We can experience them only through our own senses. Not directly quantifiable, they can claim only a subjective (first-person) validity. Natural science gets beyond this by reducing the secondary to the primary. Concretely, this means explaining the sensuously appearing world in terms of its underlying physical processes. Given this, natural science stands unalterably opposed to phenomenology since phenomenology, by definition, studies appearances while science explains them away by reducing them to physical processes. The causality science employs presupposes this reduction, and so do its mathematical descriptions. Given this, we conflate irreconcilable paradigms when we mathematize phenomenological descriptions and employ them in causal accounts of consciousness. The cognitive science that attempts this forgets that the causality employed by modern science does not imply the formalization of subjective experience, but rather the reduction of it to non-subjective processes.

As with the initial objection we considered, what is at issue here is consciousness as a worldly reality. The objection we began with asserted that we have to suspend this in order to regard the functioning of the transcendental subject, considered as an entity outside of and prior to the world. As part of this, we must also suspend the methods and assumptions of natural science. The objection we just considered buttresses this demand by asserting that if we accept the assumptions of science, consciousness disappears. Thus, as it correctly notes, the contents of our embodied consciousness are the tastes, touches, smells, sights, and textures provided by our bodily senses. To reduce the secondary qualities presented by such contents to primary ones is to take them as the effects of the world on our embodied consciousness. It is to see their reality in terms of non-subjective processes—this, by relating these qualities to the physical features of the world that produces them. The objection thus treats consciousness as a natural entity causally related to other natural entities, these being the objects affecting our bodily senses. Such a natural entity, consisting of non-subjective brain processes, is then asserted to be neither subjective nor conscious. This interpretation

may be favoured by those who, like Daniel Dennett, seek to deny consciousness on scientific grounds.²⁷ What is overlooked, however, is the crucial point that although such sensuous, qualitative contents are interpreted as the effects of the primary contents of the world, they themselves are not explained away or reduced to such primary contents. The secondary contents themselves do not suffer reduction. What is denied is simply their claim to *directly represent* the features of the external world. As is obvious, were we to dispense with these contents, science would lose its observational (empirical) basis. In point of fact, what is at issue is not such contents, but rather their reference. As Descartes formulates this, the attempt of science is to move from variations in the secondary contents, for example those of a change in colour, to the “corresponding variations” in the material world.²⁸ Now, when we turn from our embodied, empirical consciousness to transcendental consciousness taken as a function, what is at issue is not the physical reference of these individual contents, but rather those synthetic, constitutive performances that place these contents in intentional relations—relations that direct them to appearing referents. As was noted above, such performances, rather than the particular contents composing it, are the defining elements of transcendental consciousness. What cognitive science attempts to draw from phenomenology is an account of these performances.

Implicit in the above is the fact that the notion of causality, considered by itself, has nothing to do with the reduction of appearances to physical processes. As Hume and Kant showed, it is simply a formal relation, one involving necessity in the sequence of appearances. To say that “A causes B” is simply to assert that the experience of A is necessarily followed by the experience of B. One may either give such necessity a psychological force as Hume did or affirm it as an *a priori* category as Kant chose to do. In either case, however, the concept is inherently silent on the relation of primary to secondary

²⁷ According to Daniel Dennett, qualia or the secondary qualities of the world are “mere complexes of mechanically accomplished dispositions to react.” (Daniel Dennett, *Consciousness Explained* [Boston: Little, Brown and Company, 1991], 386) Accordingly, he writes, “[a] philosopher’s zombie, you will recall, is behaviorally indistinguishable from a normal human being, but is not conscious.” (*ibid.*, 405) For Dennett, “[w]e’re all zombies. Nobody is conscious.” (*ibid.*, 406)

²⁸ In Descartes’ words, “from the fact that I perceive different kinds of colors, odors, tastes, sounds, heat, hardness and so on, I very readily conclude that in the objects from which these various sense perceptions proceed there are some corresponding variations.” (Rene Descartes, *Meditations on First Philosophy*, VI, [tr.] L. LaFleur [New York: Macmillan, 1990], 77)

qualities. It only concerns conditions for validly drawing causal relations. Its strictly formal character does not mean that we cannot use it to guide us in discovering the causal relations of physical processes. Such relations, insofar as they involve necessary sequences, can be taken as applications of this formal law. These applications, as concerning physical processes, do concern the primary qualities of matter. This, however, does not mean that the concept of causality inherently involves the reduction of subjective experience to physical processes.

§4. The Relation of Consciousness Considered as a Function to Our Embodied, Empirical Consciousness

In his long career, Husserl embraced a number of positions regarding the nature of transcendental consciousness or subjectivity.²⁹ In choosing to follow the one that sees it as a function, we have, as I noted, embraced a position that allows us to employ his insights in cognitive science. Such employment takes this function as a set of formal relations (such as those of retention and protention) and processes structured by these relations (such as that of time consciousness). To see how this function relates to our actual, embodied consciousness, we have to turn to the passages of Husserl's *Logische Untersuchungen* where he describes the relation between the laws of logic and the actual functioning of our thought. On the one hand, such laws are silent on this functioning. To assert that "p implies q is equivalent to its not being the case that p holds and q does not hold" is not to make a reference to our actual thought processes. Similarly, as Husserl writes, "[t]he proposition *A is true*...says nothing about anyone's judgment, not even about judgments of anyone in general." In this, logic is like mathematics, where, as Husserl writes, "the statement that $a + b = b + a$...says nothing about anyone's counting or addition."³⁰ On the other hand, such statements, which have a pure theoretical content, can be turned into norms for our activities. They can be read as guides for what we can grasp with inner evidence. Thus, the "principles of contradiction and excluded middle" can be

²⁹ For an account of his idealistic position, see James Mensch, *Intersubjectivity and Transcendental Phenomenology* (Albany: State University of New York Press, 1988).

³⁰ Edmund Husserl, *Logische Untersuchungen, Erster Band*, in *Edmund Husserl, Gesammelte Schriften*, (ed.) U. Panzer (Hamburg: Felix Meiner Verlag, 1992), vol. 2, 187.

transformed into the proposition that “one and only one of two mutually contradictory judgments can manifest inner evidence.”³¹ Similarly, “A is true” can be transformed into the proposition “it is possible for anyone and everyone to judge that A is the case.”³² Positively, then, such laws state what we can grasp with inner evidence. Negatively, they exclude certain things from being objects of such evidence. We cannot, for example, authentically think—in the sense of “evidently grasp”—the concept of a square circle. Neither can we apprehend, say, a colour without an extension. As Husserl expresses this:

It is, accordingly, at once clear how far the logical laws and, in the first instance, the ideal laws of “authentic” thinking also claim a *psychological* meaning and also regulate the course of factual mental events. Each genuine “pure” law, which expresses a compatibility or incompatibility grounded in the nature of particular species, limits, when it relates to a species of mentally realizable contents, the empirical possibilities of psychological (phenomenological) coexistence and succession. What is seen to be incompatible in *specie* cannot be united or made compatible in empirical instances.³³

As is obvious, these laws of authentic thinking do not themselves think. Yet, insofar as they state the conditions for valid thought, they are applicable to our thinking.

Implicit, here, is the distinction that Husserl draws between conditions of validity (*Geltung*) and those of applicability (*Anwendung*).³⁴ As Husserl observes in the *Logische Untersuchungen*, the two involve very different laws. The formal laws of arithmetic, for example, give us the conditions under which additions are valid. Calculations which violate them are invalid. Quite different laws are at work when we make these laws applicable to adding machines. A mechanical adding machine uses the laws of the gear and lever, while a modern calculator uses those of electronics. Yet both instan-

³¹ *Ibid.*, vol. 2, 186.

³² *Ibid.*, vol. 2, 189.

³³ *Ibid.*, vol. 4, 727.

³⁴ Husserl makes this distinction with regard to the logical laws. For such laws to be applicable to us, we have to be able to keep propositional meanings stable. Children, before the “age of reason,” cannot do this. If we fail to distinguish the validity from the applicability of this law, we would have to call the law of non-contradiction *invalid* whenever, through age, illness, or infirmity, we could not fulfill the condition of holding meanings stable. See *ibid.*, vol. 2, 107–108.

tiate the same mathematical laws. A similar argument can be made with regard to our consciousness understood as a function. Such a consciousness represents a set of performances. For example, we apprehend objects by identifying perspectival patterns of appearing and assigning them referents. Doing so, we interpret the perceptions of a given pattern as perceptions of a given object—for example, a box that we turn in our hands, viewing it first from one side and then from another. Given that this is actually how we see objects, this process must be one that is instantiated in our embodied being, that is, in our actual empirical subjectivity. The laws of applicability for such interpretive functions are thus biological ones—those having to do with our brains. These, however, are not the laws of consciousness as a function. They do not apply to the transcendental subjectivity that is defined by such functions. Only if we equate the two sorts of laws could we be accused of ignoring the special nature of transcendental consciousness. Given that such a consciousness is not a thing but a function, this would be to commit the same kind of fallacy as equating the laws of arithmetic with those, say, of a mechanical adding machine. The relation of our embodied, empirical subjectivity to transcendental subjectivity is the same as that between this adding machine and the laws that specify the validity of its operations. Our empirical subjectivity embodies the transcendental by instantiating through its own organic processes the functions that characterize the transcendental.

§5. The Nature of Our Claim

In conclusion, we should note that the claim we are advancing is an empirical one. As such, it is open to being empirically verified or falsified. Thus, just as the laws of mathematics can help us understand the particular arrangement of a calculating machine, so the corresponding laws that describe consciousness as a function should serve us as clues to the biological basis of our perceptual, interpretive activity. Of course, the case is much simpler with regard to a machine. The basic logical operators *is*, *not*, *or*, and *and* can be instantiated in the logical gates of a computer. If a current is allowed to pass through a wire, we instantiate the operator *is*. If it is blocked from passing, we instantiate the *not*. As for the *and*, this is instantiated when, in order for the current to pass through a wire, it must pass through at least two of the wires leading into it. Similarly, the *or* is instantiated when we require that the current pass through only one of the wires leading into a wire in order for the current to pass through the latter wire itself. Given that very complex logical func-

tions can be written using these operators, they can be instantiated in a computer through the appropriate choice of logical gates. The possibility of such instantiation is simply the technical one of building a machine where the operators and the logical functions constructed from them have a field of applicability.

This, of course, is not possible with our organically based intelligence. Nevertheless, since our claim is empirical, it should allow us some predictive scope. For example, if, as Husserl asserts, the fundamental activity of perceptual consciousness is synthetic, *i.e.*, consists of performances joining together the various elements that our senses provide, we should see such synthetic activity in our neural processes. As an example of such activity, we might look to the process of neural binding, where different groups of neurons fire in synchrony to unite the different representations, *e.g.* those of colour and shape, processed by each group. Similarly, Husserl's position that all perception involves the interpretation of sense data should also have predictive force.³⁵ On the level of consciousness, this position claims that when the same data are subject to different perceptual interpretations, we should see different objects. This is something that can be verified by regarding the optical illusions where first one figure and then another alternately appear. What we experience, in the shift between one and the other—say, a young girl and an old woman—is the shift in our interpretation of what we see. A corresponding neural activity should lie behind this. To take one final example, Husserl asserts that we grasp our body as our own through the phenomena of double touch. Touching ourselves, we feel ourselves being touched. This is not the case when we touch other objects. Although we feel them, we do not feel their being touched. If this is correct, then someone losing feeling in a bodily limb would, in the absence of visual data, not recognize the limb as his own. Again, this is something that can be verified. On the neurological level, this lack of recognition would point to the sense of double touch as

³⁵ As Husserl expresses this, "*Zur Wahrnehmung gehört, dass etwas in ihr erscheine; aber die Interpretation macht aus, was wir Erscheinen nennen, mag sie unrichtig sein oder nicht, mag sie sich getreu und adäquat an den Rahmen des unmittelbar Gegebenen halten oder ihn, künftige Wahrnehmung gleichsam antizipierend, überschreiten. Das Haus erscheint mir — wodurch anders, als dass ich die wirklich erlebten Sinnesinhalt in gewisser Weise interpretiere. Ich höre einen Leierkasten — die empfundenen Töne deute ich eben als Leierkasten-töne. Ebenso nehme ich interpretierend meine psychischen Erscheinungen wahr, die mich 'durchschauende Seligkeit, den Kummer im Herzen' usw. Sie heißen 'Erscheinungen,' oder besser erscheinende Inhalte, eben als Inhalte der perzeptiver Interpretation.*" (*ibid.*, vol. 4, 762)

essential to the neural binding that allows us to recognize our bodies as our own. As these examples indicate, the demonstration that the functions characterizing transcendental consciousness are applicable to our organic, actually perceiving consciousness is an empirical one. What is needed is not a philosophical proof of this hypothesis, but actual empirical work. The same holds when we attempt to apply the insights of phenomenology to the construction of machine-based intelligence.

james.mensch@gmail.com