The Psychology and Philosophy of Natural Numbers[1]

Oliver R. Marshall[2]

In this paper I argue against both neuropsychological and cognitive accounts of our grasp of numbers. I show that despite the points of divergence between these two accounts, they face analogous problems. Both presuppose too much about what they purport to explain to be informative, and also characterize our grasp of numbers in a way that is absurd in the light of what we already know from the point of view of mathematical practice. Then I offer a positive methodological proposal about the role that cognitive science should play in the philosophy of mathematics.

1. Introduction

Some branches of mathematics concern a given subject matter*:* a subject matter of which most of us have an intuitive grasp, prior to receiving mathematical training. For example, the axioms of Euclidean geometry concern space, or at least how space appears to us. This is a subject matter about which we have a stock of intuitions. The axioms of a ring, by contrast, do not concern a given subject matter, but instead concern the family of algebraic structures that they define. Unlike geometry, arithmetic was not originally developed from axioms. However, it too

---

[1] This paper takes its cue from the research and teaching of Haim Gaifman and Saul Kripke. I am extremely grateful to them both. I also thank Jody Azzouni, Michael Levin, Jesse Prinz, Nathan Salmon, Mark Steiner, and especially Gary Ostertag, as well as an audience at the University of Texas, Austin, for their helpful suggestions.

[2] The Saul Kripke Center, The Graduate Center, CUNY, New York, New York 10016-4309, U.S.A. OMarshall@gc.cuny.edu

concerns a given subject matter, namely the natural numbers, so there is a question regarding how we are able to grasp these intuitively. Since this seems to be an empirical question, one is led to ask whether our intuitive grasp of numbers can be explained using the resources of cognitive science.

Neuropsychologists who work on this topic typically argue that our concepts of numbers are among our core, innate representations. For example, Stanislas Dehaene writes that ‘the ultimate foundations of mathematics rests [*sic*] on core representations that have been internalized in our brains through evolution’ [2001a, p. 16].[3] Along with other leading neuropsychologists, he is optimistic about the prospects for a detailed, mathematically rigorous theory of our intuitive grasp of numbers that will transform, or even supplant, the philosophy of mathematics.[4] He goes so far as to say that his neuropsychological approach ‘can provide a new “naturalistic” approach to the problem of the foundations of mathematics’ [2001a, p. 34]. I will refer to this approach as ‘neuropsychologism.’

Then there are theorists with more modest ambitions, who see cognitive science as necessary but not sufficient for explaining our intuitive grasp of numbers. Thinkers in this camp generally appeal to logico-philosophical considerations to tell us what numbers are and what is logically required to grasp them, and to cognitive science to explain how we actually come to do so.[5] Typically, these theorists argue that concepts of numbers are not among our core, innate representations, and so try to explain how we learn these concepts during development.

---

[3] See also [Gallistel, 2011, p. 310].

[4] [Dehaene, 1997; Gallistel and Gelman, 2000; Rocha et al., 2005].

[5] See [Burge, 2007; 2009; 2010; Carey, 2009; Giaquinto, 2007; Rips, 2011].

I begin this paper by focusing on neuropsychologism. First, in sections 2 – 4, I describe and motivate Dehaene's triple-code model, including its key representational component – the accumulator. Then, in section 5, I subject this model to criticism. While I do not doubt that we have some of the capacities that advocates of this sort of view are trying to model, and do not wish to take issue with the claim that these capacities are innate, I do insist that they are insufficient to represent numbers. Further, I argue that the insufficiency of Dehaene's model renders his theory not only false but also subject to a fallacy that, curiously enough, has not been noticed in the literature.

Having despaired of neuropsychologism, I turn, in section 6, to Tyler Burge's attempt to explain our grasp of numbers in terms of our acquisition of understood numerical concepts. I argue that Burge errs in trying to characterize our grasp of numbers in a way that seems absurd in the light of what we already know by reflection on arithmetical reasoning, and that his account of acquisition is either insufficient to explain our grasp of numbers, or presupposes too much about what it purports to explain, with the result that it is uninformative. I believe that similar problems afflict the views of the other theorists in the modest camp, but arguing for this is beyond the scope of this paper.[6]

Despite all this, I think that it would be premature for philosophers of mathematics to ignore cognitive science completely. So, in the light of my criticisms, I make a positive methodological proposal about the limited role that it should be allowed to play. Then I illustrate this proposal, using my solution to a problem that has been posed, independently, by Saul Kripke and Marcus Giaquinto. While my solution to this problem resembles certain aspects of

[6] For more details see [Marshall, 2016].

Dehaene's proposal, I argue the problem cannot be solved using only the aforementioned resources of cognitive science.

2. The Number Sense

Most of us experience the phenomena of being able to estimate, visually, that there are between twenty and forty people in the room, and of being able to look at much smaller pluralities, such as three cows in a field, and see how many there are without counting. Further to these reflections, there are many disparate empirical studies in support of the hypothesis that humans, including pre-linguistic infants and people with a reduced numerical lexicon, have the ability to sense the cardinal size of pluralities. More specifically, it is hypothesized that even prior to learning numerical concepts we are able to:

(a) Perceptually estimate the cardinal size of a given plurality, and perceptually discriminate different pluralities in terms of approximations of their cardinal size.

(b) Perceive the exact cardinal size of pluralities of up to three or four members at a much faster rate than that required by discursive counting, an ability known as "subitizing."

There are various reasons to think that these abilities are innate. Firstly, as I will explain, they are also found in animals and infants, making it plausible, although by no means certain, that we inherited these abilities through our genes. Secondly, as we will see, numerate human adults with more accurate tools at their disposal —such as counting— still exercise the abilities described in (a) and (b). This would be somewhat surprising if these abilities were learned, since in that case one might well abandon them; however, it is easily explained on the hypothesis that these abilities are innate, since what is innate one cannot choose to abandon.[7] In any case, that we possess these abilities has come to be widely accepted in the psychological literature, in no small part due to the work of Dehaene, who refers to them jointly as our 'number sense' [1997]. According to Dehaene it is this innate sense that constitutes our ability to think about numbers intuitively.[8]

Before this claim can be assessed, it is necessary to distinguish between counting — which requires putting objects in one-to-one correspondence with discrete symbols or numbers— and *summation*, which requires only the accumulation of a continuous variable such as a physical magnitude.[9] For example, an egg timer does not count minutes discretely, but simply accumulates a quantity of sand. Likewise, the pedometer in an iPhone does not literally count one's steps, but accumulates a physical magnitude in response to hip movement. The reason that this distinction is important is that Dehaene hypothesizes that the number sense is an analog system that represents numbers —despite the fact that numbers are discrete— by summation, using what he calls 'a continuous quantitative internal representation' [1997, p. 220]. I will now describe some of the evidence for the claim that the number sense is analog.

Rats can learn to press a lever repeatedly before pressing a second lever to get a reward. Having learned to do this, they soon learn to respond with roughly the required number of presses on the first lever, before pressing the second and searching for the reward. The accuracy of their estimative capacities can then be measured by the probability of search after the wrong number of presses (the confounding quantity of duration having been controlled for). For each

[7] [Mandelbaum, 2013, p. 379].

[8] [Dehaene, 1997, p. 102].

[9] [Franks et al., 2006].

number of presses required by the experimenter, the mean of the distribution of the rat's responses is slightly higher than is required. Further, the standard deviation around the mean increases as a constant ratio of the mean, from which it follows that greater magnitudes must differ more than smaller ones in order for the rat to discriminate them. This accords with Weber's law,[10] that the discriminability of any two magnitudes is a function of their ratio. In other words, the ratio of the minimum change —required to discriminate two magnitudes— to the initial magnitude is constant. Weber's law applies to representations of continuous variables such as length, area and loudness, so conformity to it is evidence of analog summation rather than discrete counting.

Further experiments show that rats can also accumulate information concerning magnitude while ignoring other confounding properties of the stimuli in question. For example, they learn to press one lever in response to two flashes and another lever in response to four, before learning to press the first lever in response to two sounds and the second in response to four. Surprisingly, when presented with a flash synchronized with a sound, they press the lever corresponding to 2, and when presented with two flashes synchronized with two sounds they press the lever corresponding to 4. This suggests that they learn to associate different levers with different magnitudes, rather than with different perceptual modalities. There is also evidence that in addition to rats, birds, honeybees and cicadas can accumulate a variable that reliably correlates with the cardinal size of a given plurality rather than with its other properties.[11]

Both of the aforementioned experiments on rats have been replicated on humans. For example, in order to replicate the lever experiment, subjects are given a target number, and then

---

[10] [Platt and Johnson, 1971].

[11] [Butterworth, 1999; Burge, 2010; Carey, 2009; Dehaene, 1997; Giaquinto, 2007].

asked to press the required number of times on a keyboard as fast as possible. They are explicitly asked to do this "by feel," without counting, and are also required to press at a rate too fast to allow for verbal counting. The accuracy of their estimative capacities can then be measured by the probability of stopping after the wrong number of presses. Results have been obtained that conform to Weber's law in a way that is strikingly similar to those obtained in the lever experiment on rats.[12] As with the rats, the standard deviation around the mean of the distribution of responses increases as a constant ratio of the mean. So there is evidence that both rats and humans estimate by accumulating a continuous variable. According to Dehaene, the similarity between the data concerning human and animal behavior suggests that 'inasmuch as the approximate perception of numerosity is concerned, humans are no different from rats or pigeons' [1997, p. 61].

But why should we think that our number sense is still of use to numerate human adults? Because there is evidence that various other abilities depend on it. It is to this evidence that I now turn.

3. The Use of Number Sense by Numerate Adults

I begin with our ability to distinguish numbers during comparison tasks, an ability that is subject to two consequences of Weber's law: the distance and magnitude effects. The distance effect is that the smaller the difference between two inputs the longer it takes to distinguish them. The magnitude effect is that the greater the magnitude of two inputs the longer it takes to distinguish them, given a fixed difference in magnitude. In his [1997, pp. 61-64], Dehaene reports that both of these effects are manifest by human adults when they are asked to compare pairs of pluralities,

---

[12] [Whalen et. al, 1999].

as well as pairs of digits. Surprisingly, the distance effect is also manifest when subjects are asked to compare pairs of two-digit numerals. For example, it takes longer for human adults to decide whether 71 is greater than 65 that it does for them to decide whether 79 is greater than 65.

Dehaene also reports [1997, p. 219] that digital mechanisms are not ordinarily subject to these effects. For example, modern computers, which represent numbers in binary code, are not subject to the distance effect, since it actually takes longer for them to distinguish the pair $\{8_2, 10_2\}$ than it does for them to distinguish the closer pair $\{7_2, 8_2\}$. This is because distinguishing the former pair requires comparing the second to last digit of binary '1000' with that of '1010', while distinguishing the latter pair only requires comparing the first digit of '111' with that of '1000'. Neither are modern computers subject to the magnitude effect, since it takes longer for them to distinguish the pair $\{6_2, 7_2\}$ than it does for them to distinguish the pair $\{7_2, 8_2\}$. This is because distinguishing the former pair requires comparing the last digit of '110' with that of '111', while, as we have seen, distinguishing the latter pair only requires comparing the first digit of '111' with that of '1000'.

Dehaene argues from this that the brain does not perform these tasks like an ordinary digital computer. Rather, he claims [1997, p. 220], it should be modeled using some sort of analog accumulator, since such machines are themselves subject to these effects:

> The peculiar way in which we compare numbers thus reveals the original principles used by the brain to represent parameters in the environment, such as number. Unlike the computer, it does not rely on digital code, but on a continuous quantitative internal representation. The brain is not a logical machine, but an analog device.

Quoting Gallistel [1990] approvingly, Dehaene continues:

> Instead of using number to represent magnitude, the rat [like the *Homo Sapiens!*] uses magnitude to represent number.

I will return to this claim in due course.

Another proposed reason to think that an innate number sense is still of use to numerate human adults, is that its integration with our culturally acquired abilities to represent numbers promises to solve a problem that has been posed, independently, by Marcus Giaquinto [2007] and Saul Kripke [1992]. The problem is to explain why decimal users have a good idea of how many members a given plurality has, or what number is the solution to an arithmetical problem, on being given a decimal numeral, even though they have little or no idea on being given the corresponding numeral from another system in which they have been trained. Giaquinto, in his [2007, p. 92], illustrates the problem as follows:

> You ask some students if any of them can work out the value of seven to the power of six; one of them quickly writes down ''1,000,000'' saying that this is the answer in base 7 notation. Understanding the place system of numerals, you will see that this smart-alec answer is correct. Even so, it will probably leave you feeling somewhat in the dark. Why? It correctly designates the number, and it does so in a language you understand. Given any other number in base 7 notation you would be able to tell which of the two is larger, and the algorithms you know for multi-digit addition and multiplication work just as well in base 7 notation. So what is missing?

To explain this phenomenon, Giaquinto suggests that

> a strong association of number size representations with decimal numerals and with your natural language number expressions has been established in your mind, while

> no such link has been established between number size representations and multi-digit numerals in other bases.

For example, if a student tells you that the answer is 117,649, you thereby have the answer, because your approximate sense of 117,649 is associated with the corresponding decimal and lexical numerals.[13]

Here one might wonder whether it is necessary for your number sense to be associated with lexical as well as decimal numerals in order to explain your preference for decimal. But, in any case, it is obvious that this association would not suffice. For, by hypothesis, you only have an exact sense of the size of pluralities of up to three or four members. So, your approximate sense of 117,649 could only help facilitate knowledge of an approximation of the answer to the question what is seven to the power of six. It could not explain why, when you are told that the answer is 117,649, you thereby know *exactly* what the answer is.

4. The Triple-Code Model

Of course there is no need to claim that all arithmetical tasks are performed by the accumulation and mental manipulation of quantities. For example, according to Dehaene and Cohen's triple-code model, numerate adults have two other kinds of mental representations at their disposal. Firstly, they have the mental correspondents of number words stored in '*a verbal word frame*, in which numbers are represented as syntactically organized sequences of words' [1995, p. 85]. Secondly they have the correspondents of positional numerals stored in 'a *visual arabic number form*, in which numbers are represented as strings of digits on an internal visio-spatial scratchpad' [p. 85]. Nevertheless, according to Dehaene and Cohen [1995, pp. 85-6], it is our

[13] This explanation is also endorsed in [Carey, 2009, p. 337-338].

number sense —with its accumulated analog representation of number— that gives meaning to these other kinds of representation:

> Under the assumptions of the triple-code model, neither the arabic number form nor the verbal word frame contain any semantic information. The meaning of numbers is represented only in the third pole of the model, the *analogical magnitude representation.*[14]

I will now say more about how the analog representation of number is supposed to be embodied and how it is supposed to sharpened by these other kinds of representation.

The accumulator is a metaphor that characterizes a neural network, which is in turn a metaphor that characterizes how neurons function. To put things in terms of the network metaphor, each object in a perceived plurality is allocated a quantity of neural activity, which is then normalized to an approximately constant quantity in case more activity is initially assigned to larger objects (or louder sounds, etc.). The normalizations allocated to each object are then summed by what Dehaene calls 'accumulation neurons' [1997, p. 251], and the resulting total is divided by the constant quantity, to yield an estimate of the size of the plurality as output, this being the final value of the metaphorical accumulator. So called 'detector neurons' [ibid] are disposed to fire when the estimates they receive from the accumulator are within fixed intervals. They reach a firing peak for the estimates to which they are 'tuned' [ibid], and show decreased firing activity on receipt of estimates that are larger or smaller than the ones to which they are tuned, in a way that is normally distributed around the peak.[15] According to Dehaene, the interval around the peak is the same for almost every detector neuron. The only exceptions are the

---

[14] Cf. [Mandelbaum, 2013, p. 369].

[15] See also [Dehaene and Changeux, 1993, pp. 394-395].

neurons tuned to one, two and three, which show much smaller intervals, modeling the fact that we can perceive the cardinal size of very small pluralities with great accuracy. For estimates of above three, the interval around the peak is fixed at about plus or minus 30%, so the range of estimates for which the various detector neurons fire increases with the estimate to which they are tuned. For example, a detector neuron tuned to five will fire less frequently on receipt of an estimate of approximately four but not at all for an estimate of approximately one; on the other hand a detector neuron tuned to an estimate of fifty will still fire on receipt of an estimate of approximately forty. So, as the mean of the numbers presented in the comparison task increases, so does the inaccuracy of the corresponding detector neurons as they begin to fire more frequently for estimates to which they are *not* tuned. Thus the inaccuracy of the model in performing comparison tasks conforms to the magnitude effect, as a result of the distribution around the mean increasing as a constant ratio of a growing mean. As for the model's being subject to the distance effect, according to Dehaene [1997, p. 251], this is also a result of the distribution of neuronal activity.

Finally, to account for the fact that we can represent numbers precisely with numerals and number words, Dehaene proposes that our analog representation of number is integrated with language in a way that gives content to the latter while making the former more precise: 'Symbols tune neurons much more sharply, thus allowing them to encode a precise quantity' [1997, p. 271]. Thus, he continues, a perceived plurality 'evokes broad and fuzzy activation in the parietal neurons, while symbols induce firing in a smaller but highly selective subgroup.' This promises to help solve the problem about our preference for decimal notation, which was raised in section 3. Here the idea would be that if a student tells you that seven to the power of six is 117,649, you thereby have a good idea of the answer, because your approximate sense of

117,649 is associated with the corresponding decimal numeral in your visual arabic number form, in a way that sharpens your number sense to encode a precise quantity. Thus the decimal user might be said to know more or less —if not quite exactly— what the answer is.

5. Why the Triple-Code Model is Not Sufficient to Explain Our Intuitive Grasp of Numbers

Frege warned us to be careful 'always to separate sharply the psychological from the logical, the subjective from the objective' [1884, p. x]. In the light of this, we need to separate two questions. Firstly, there is the question of what sorts of subjective mental representations must be posited by cognitive scientists in order to explain the relevant data. Secondly, there is the question of whether these representations represent numbers. This brings me to the first problem for the triple-code model, which is that the accumulator does not represent the following constitutive properties of numbers: (I) discreteness, (II) potential infinity, (III) general applicability and (IV) duality. I will now discuss each of (I) – (IV) in turn.

(I) The natural numbers are as a constitutive matter discrete. In contrast, the variable accumulated by an analog accumulator is continuous. For this reason, as Tyler Burge points out, while an analog "representation" can be *correlated* approximately with number, it cannot be accurate or inaccurate based upon whether or not it reflects the right discrete properties.[16] For example, it cannot accurately represent 117,649 as distinct from 117,650. But, Burge reasons, if it does not have accuracy conditions concerning discrete properties, then it cannot *represent* these properties at all. But then it cannot represent natural numbers, since these are as a constitutive matter discrete.

---

[16] [Burge, 2010], especially chapter 10.

One might try to meet this objection by appeal to the hypothesis that the accumulator accumulates a fixed unit of quantity, rather like an egg timer that is filled by pouring in cups of sand.[17] But this hypothesis remains subject to the previous objection, since the neural analog of one cup of sand will still be *approximately* one cup, and so for example will not be able to represent 1 as distinct from 1.00001. Further, the claim that the accumulator is integrated with numerals and number words is of no help either, since the resulting smaller intervals around the 'firing peak' are still fuzzy rather than discrete. Furthermore, and to return to the puzzle raised at the end of section 3, the triple-code model *still* cannot explain why, when you are told that the answer is 117,649, you thereby know *exactly* what the answer is, i.e. 117,649 as distinct from 117,649.00001.

(II) The assumption that there are potentially infinitely many sentences of English is a constraint on linguistic theorizing among cognitive scientists; likewise, the corresponding assumption about numbers is an equally reasonable constraint on cognitive accounts of our arithmetical capacities. But the accumulator embodies a perceptual, pre-linguistic capacity, and as such lacks the recursive or iterative capacity for potential infinity.[18] For example, it does not have the potential to repeat the step of accumulating a fixed unit of quantity indefinitely.

(III) Because the accumulator embodies a perceptual, pre-linguistic capacity, it can only detect the sizes of concrete pluralities. But as Frege pointed out, number is not simply a property of concrete pluralities, since almost anything that can be conceptualized in terms of a suitable kind-concept can also be numbered.[19]

---

[17] [Dehaene, 1997, p. 18; Gallistel and Gelman, 2000].

[18] [Carey, 2009].

[19] [Frege, 1884], especially §14.

(IV) The term ‘natural numbers’ can be used to denote two distinct sets of entities: the (finite) cardinal and (finite) ordinal numbers. That these coincide in the finite case is why, when we establish a one-to-one correspondence between ordinals and objects to be numbered, the last ordinal that corresponds with an object will be the same irrespective of the order in which the objects are counted, which in turn explains why the last ordinal is also the cardinality of the set of objects.[20] Conversely, given this cardinality, we can predict, with mathematical certainty, which ordinal will be the last one that corresponds with an object.[21] The problem is that Dehaene’s hypothesis, that the accumulator accumulates a fixed unit of quantity for each perceived object, does not distinguish these two aspects of natural numbers, and cannot explain the relationship between them.[22] How on earth, for example, can Dehaene explain that we know, with mathematical certainty, which number will be the last one that corresponds with an object, given that we know the cardinality of the set of objects?

To be clear, I do not deny that the variable accumulated by our accumulator represents a species of quantity than can be correlated approximately with number. But I do take these four objections to show, conclusively, that the accumulator does not represent natural numbers. Furthermore, it does not represent either rational or real numbers. This is because both sets of numbers share the following feature, under their usual ordering:

(V) Density: between any two numbers there is another number.

In contrast, while the states of the accumulator are continuous, they do not possess this structural feature.

---

[20] [Cantor, 1885].

[21] [Steiner, 2005].

[22] Thanks to Mark Steiner for this point.

My next objection to the triple-code model pertains to the fact that Dehaene appears to identify numbers with neural outputs of his model (call this 'Idealism'). For example, we are told:

> Number appears as one of the fundamental dimensions according to which our nervous system parses the external world [1997, p. 227].
>
> Numbers, like other mathematical objects, are mental constructions whose roots are to be found in the adaptation of the human brain to the regularities of the universe [p. 233].
>
> My proposal is that the brain evolved a number system to capture a significant regularity of the outside world, the fact that at our scale, the world is largely composed of solid physical objects that move and can be grouped according to the laws of arithmetic [2001b, p. 100].[23]

There are several problems here. Firstly, Idealism commits Dehaene to saying that the numbers are found only when and where suitably and contingently evolved brains are found, contrary to the fact that tense, location and modal distinctions do not apply to numbers.[24] Secondly, one might worry that Idealism renders Dehaene's account vacuous. This is because it requires identifying numbers with mental constructions that can only be characterized as the outputs of a neural network which are themselves numbers, the mathematical network metaphor being absolutely necessary to characterize how neurons function. In fact, I think that the situation is even worse than this, because Idealism, taken together with the fact that the mental constructions

[23] Dehaene is not alone in finding numbers in our brains. See also [Gelman and Gallistel, 2015, p. xviii].

[24] [Burgess, 2003].

in question lack the properties described in (I) – (V), commits Dehaene to what I will call "Revisionism," the view that natural numbers do not have the properties described in (I) – (IV), while neither rational nor real numbers have the property described in (V). In which case he is guilty of just the sort of absurd revisionist psychologism that Frege ridiculed.

Of course Dehaene could reject Idealism and accept that numbers are not mental constructions; rather, natural numbers are abstracta satisfying (I) – (IV), while rational and real numbers are abstracta satisfying (V).[25] In this way Revisionism would be avoided. However, the triple-code model is still subject to the objection that it is not sufficient to represent natural, rational or real numbers, since it is not sufficient to represent the properties described in (I) – (V). Further, this, together with the acceptance of numbers as abstracta, entails an absurdly strong form of Platonism, according to which are numbers are abstracta that no one ever succeeds in representing accurately.

My next objection is that either of Revisionism or the above form of Platonism is sufficient to commit the advocate of the triple-code model to a fallacy, which is that if the model is true, then humans cannot represent natural, rational or real numbers; this undermines the arguments for the triple-code model, which cannot be given unless humans can represent numbers. To see this, it will help to denote the collection of properties described in (I) – (V) as 'Φ.' First I concede, for the sake of argument, that if humans are able to represent natural, rational or real numbers that have the appropriate properties from Φ, then the triple-code model is true. Further, if the triple-code model is true, then, by either Revisionism or Platonism, humans are not able to represent natural, rational or real numbers that have the appropriate properties from Φ. From these premises it follows that if humans are able to represent natural, rational or

[25] Cf. [Mandelbaum, 2013, p. 382, fn. 11 and 12].

real numbers that have the appropriate properties from Φ, then humans are not able to represent natural, rational or real numbers that have the appropriate properties from Φ. From this it follows, by the principle that if p implies its own falsity, then p is false, that humans are not able to represent natural, rational or real numbers that have the appropriate properties from Φ.[26] The problem is that this pulls out the rug from under the triple-code model, since it's impossible to derive its explanations of the relevant data without representing natural, rational or real numbers that have the appropriate properties from Φ.[27] For example, in order to argue for the existence of detector neurons, Dehaene uses psychophysical bridging laws, to derive the magnitude and distance effects from distributions of simulated neuronal activity, distributions that model the neurons in question.[28] Clearly this relies on statistical argumentation, and in particular presupposes a distribution function that assigns numbers that have properties from Φ to events (i.e. the simulated firing of detector neurons). I, for one, cannot see how these arguments could be replaced with ones that only mention continuous quantities and do not mention natural, rational or real numbers.

Finally, there is the problem of vacuity. I have argued that numbers cannot be represented unless the system of representation has the structure of numbers. This shows that the structure of numbers must be assumed in explanations of how we grasp them. Further, having these structural

[26] Cf. [Russell, 1940, p. 15]. Thanks to Mark Steiner for suggesting that Russell's argument is a charitable reconstruction of Dehaene's position.

[27] Cf. [Steiner, 1978, pp. 19-20], in which he uses somewhat similar reasoning to show the circularity in the argument that numbers are empirical posits because they are inferred from their indispensability to science.

[28] [Dehaene, 1997, p. 250].

features seems sufficient to represent numbers without the help of the details of the triple-code model. Thus, the model seems doomed to be either insufficient, or to assume so much about what it purports to explain as to render it uninformative.

6. Burge on Numerical Concepts

Next I turn to Tyler Burge's attempt to explain our grasp of numbers in terms of our acquisition of numerical concepts. In Burge's view concepts are the constituents of thoughts, and are also shareable ways of thinking of subject matters. As for what numerical concepts are concepts of, Burge claims that Frege 'was correct in thinking of numbers as having a certain second-order status' [2007, p. 71]. He also follows Frege by claiming that numbers are grasped 'only through understanding arithmetical propositions' [2009, p. 315]. Here Burge has in mind applied arithmetical propositions like *that there are 2 houses of Congress.* The idea appears to be that we understand concepts such as those of 2 and 4 by understanding applied arithmetical propositions, and only then form unapplied arithmetical propositions such as *that 2 + 2 = 4*:[29]

> Understanding pure arithmetic requires understanding applications of it in counting. Understanding '3' involves understanding 'there are 3 F's', which in turn requires being able to count the F's—put them in one–one relation to the numbers up to 3 [2007, p. 72].

Burge continues that such understanding can be combined with three other capacities, to place one in an epistemic relation to numbers that is more immediate than conceptualization:

---

[29] A referee points out to me that a detailed development of this claim can already be found in [Tennant, 1987], in chapters 20 and 25.

> Being able to apply a canonical (numeral-like) concept for the number 3 in an immediate perceptual way, seems to me to constitute a 'not completely conceptual' relation to the number.

This requires some unpacking. Firstly, one must possess *canonical* concepts of numbers, these being the conceptual counterparts of numerals belonging to the decimal system, rather than descriptions of numbers in the conceptual counterpart of, for example, successor notation.[30] Secondly, canonical concepts corresponding to the digits must form the *base* of mental computation, so that they determine small numbers in a way that is computationally simple, rather than determining them as the results of recursive computation. Regarding the difference between computationally simple and complex concepts, Burge has this to say:

> Canonical concepts for larger numbers are built by simple recursive rules from the simplest ones…
>
> Understanding what larger numbers are derives from this immediate hold on the applicability of the smaller ones [p. 72].

Finally, one must be able to apply these simple concepts in a way that is guided by the ability to subitize (see section 2 of this paper). Thus for Burge it is only

> in individuals who have an understanding of a numerical system, [that] the primitive subitizing capacities join with conceptual abilities to support noninferential, noncomputational numerical assignments in thought to small groupings [2009, p. 313].

[30] Cf. [Tennant, ibid; Shwayder, 2008, p. 7; Peacocke, 2009; Kripke, 2011a] regarding canonical concepts.

7. Problems for Burge's Account

My first objection to Burge is that if we allow him the assumption that there are epistemically immediate thoughts about numbers, then his theory is too restrictive, since according to it we can have such thoughts about only the smallest few numbers. This is because we are only able to subitize the exact size of pluralities of up to three or four members, and have to figure out the exact number of even slightly larger pluralities by more discursive means. Burge is aware that this aspect of his theory makes it questionable, and is willing to grant that theories according to which multi-digit numerals facilitate epistemically immediate thoughts about numbers are also 'tenable' [2007, p. 74]. However, he has a clear preference for the very strict theory of epistemic immediacy described above, arguing that it is preferable to the less restrictive theory due to Kripke, according to which decimal numerals including multi-digit ones are suitably epistemically immediate. Kripke's theory is motivated by consciously accessible, intuitive evidence, in particular the intuition that decimal numerals —including multi-digit ones— are 'immediately revelatory' [2011a, p. 261]. By this Kripke means that if a decimal user counts or is told that the number of guests is 547, he thereby *knows how many* guests there are; likewise, if a decimal user calculates that the factorial of 5 is 120, he thereby *knows what number* the factorial of 5 is; no further inference or fact-finding is needed. In contrast, 'the factorial of 5' and '5 X 4 X 3 X 2 X 1' are not immediately revelatory. Further, as both Kripke and Giaquinto point out (see section 3), neither are the corresponding numerals from other systems like binary and base 7 in which one has been trained.

The issue here is whether we should characterize epistemic immediacy intuitively, from the point of view of someone engaged in arithmetical practice, as Kripke does, or in Burge's more restrictive theoretical terms. Burge's complaint, in his [2007], is that Kripke's intuitive notion ignores ‘evidence from psychology’ that the canonical concept of 547 is understood inferentially, by performing sub-personal, recursive computations on its psychologically basic, perceptually applicable elements, and that such evidence suggests we are immediately related to 4 but not to 547. But there are various reasons to resist these claims.

Firstly, Burge does not offer any specific evidence that the canonical concept of 547 is understood inferentially by performing sub-personal computations on its perceptually applicable components.[31] Secondly, there is the aforementioned worry that Burge's criterion is too restrictive to solve the puzzle raised by Kripke and Giaquinto. Is it a consequence of Burge's view that no one *really* knows how many F's there are or what number they are thinking about, by grasping a multi-digit numeral? If so, then the view has an absurd consequence. Of course Burge might respond that from the point of view of cognitive science this consequence is not absurd. However, I do not see why putative evidence about the time taken by sub-personal mental processes should trump the intuitive point of view of someone engaged in arithmetical practice, from which it is abundantly clear that we resolve computations with multi-digit base-10 numerals.

I now turn to problems with another aspect of Burge's criterion of immediacy: the perceptual applicability of understood numeral-like numerical concepts, which raises the question of what is required in order to understand canonical numerical concepts. Burge's proposal, in his [2007, p. 72], is that understanding such concepts requires understanding applied arithmetical propositions, which in turn requires the ability to count:

> Understanding ‘3’ involves understanding ‘there are 3 F's’, which in turn requires being able to count the F's —put them in one-one relation to the numbers up to 3.

---

[31] [Burge, 2007, pp. 73-74].

Obviously this gloss of counting presupposes a grasp of numbers, and, as earlier quotations show, it is Burge's view that that we cannot represent numbers prior to acquiring concepts of them. So if this account is to avoid circularity, it must assume that we can *already* represent numbers via propositions containing *non*-numeral-like concepts, and can use the latter in our counting experience, through which we come to understand propositions containing canonical numeral-like concepts. But in the present context this is a significant assumption about what is to be explained, one that needs spelling out.

One can remove reference to numbers from the requirements for counting, by stating them as follows. Firstly, the words in the count list must be recited in a stable order. Secondly, a one-to-one correspondence must be established between the words in the count list and the objects counted. Thirdly, one must be able to give the final word of the count in answer to the question 'how many F's?' However, as Burge himself realizes,[32] it is possible to meet all of these requirements without grasping the cardinal significance of counting or grasping which cardinals are denoted by the members of the count list. To see this, consider that there is a stage during development when children can recite a short list of numerals in a stable order, put them in one-to-one correspondence with the F's, and recite the last numeral in the count when asked 'how many F's?' And yet, when instructed to give the experimenter *m* F's —where *m* is the last numeral recited— they give the experimenter a random number of F's. This result is due to Karen Wynn, who summarizes in her [1992, p. 234]:

> In all cases, children could successfully count larger sets of items than they could give when asked… Thus children's ability to correctly give a certain number of items lags well behind their ability to successfully count that same number of items.

---

[32] [Burge, 2010, p. 491].

This suggests that meeting the above requirements on counting does not suffice for understanding the cardinal significance of numerical concepts; in Burge's terminology, it does not suffice for full competence understanding. So what does suffice? Burge does not say.

Finally, Burge's account (like Dehaene's) conflates the (finite) ordinal and cardinal numbers. This is because his claim, that counting is necessary for understanding 'there are 3 F's', entails that a process of reciting or generating numbers, to be corresponded with objects counted, is necessary for understanding 'there are 3 F's'. This is to conflate cardinals with ordinals; the latter are essentially connected to a process of reciting numbers, but the former are not.

To conclude this discussion, in purporting to explain epistemically immediate thoughts about numbers in terms of the conceptualized successor of subitizing, Burge's account is too restrictive —recognizing such thoughts for only the smallest numbers—, unsupported by evidence, and, most importantly, based on a significant assumption about what it purports to explain, one that cannot be spelled out in terms of meeting the above conditions on counting. Further, since Burge's theory of pure arithmetical understanding also appeals to the conceptualized successor of subitizing, it too faces all of these problems.

8. A Methodological Moral

I have argued that neither neuropsychologism nor Burge's cognitive account of numerical concepts can provide a complete explanation of our grasp of numbers. I would add that despite the points of divergence between these two accounts, they face analogous problems. Both presuppose too much about what they purport to explain to be informative,[33] while also trying to

[33] I think that a similar problem also afflicts the "Quinean bootstrapping" proposal in [Carey, 2009], but I cannot argue for this claim here.

characterize our grasp of numbers in a way that seems absurd in the light of what we already know from the point of view of mathematical practice. The moral that I draw from all of this is that, in addition to evidence from cognitive science, philosophical reflection from the point of view of one who participates in arithmetical practice is also needed,[34] including reflection on the representational properties of numerical notations. I will now illustrate how these two approaches can work in concert, by proposing a solution to the problem raised in section 3, due to Kripke and Giaquinto: of how to explain why decimal notation is immediately revelatory, in that decimal users know how many members a given plurality has, or what number is the solution to an arithmetical problem, on being given a decimal numeral, even though they do not on being given the corresponding numeral from another system in which they have been trained.

Kripke, in his [1992], claims that for the purposes of mathematics, one should, whenever possible, use a notation that is 'structurally revelatory' – one that has a structural affinity, akin to isomorphism, with the subject matter that it represents.[35] I would add that while having a structurally revelatory notation is sufficient for being able to represent a mathematical subject matter, being structurally revelatory is not sufficient for being immediately revelatory, as can be seen from the example of prefix notation. Consider the representations of the operations of addition, multiplication and exponentiation, as these are given, implicitly, by recursion equations. These equations are usually stated in infix notation, as in:

i. $x + 0 = x$

[34] Cf. [Nagel, 1997], especially chapters 2 - 4.

[35] See [Marshall, 2016] in which I describe Kripke's example of the standard structurally revelatory notation for the universe of hereditarily finite sets, as well as the example of stroke notation as a structurally revelatory notation for the so-called Frege-Russell numbers.

ii. $x + S(y) = S(x + y)$

But they can also be stated in prefix notation:

i.a. $+(x, 0) = x$

ii.a. $+(x\ S(y)) = S(+(x, y))$

I am surely not alone in finding the meaning of the latter equations much harder to discern than that of their infix equivalents.[36] This is also the case with the infix and prefix statements of the associative law for addition:

$$x + (y + z) = (x + y) + z$$

$$++xyz = +x+yz$$

So infix notation is much more immediately revelatory than prefix notation. Yet the notations are equally structurally revelatory of what they represent, since they have an equal structural affinity with their subject matter, having isomorphic parsing trees with the same roots, labels and orderings. With that said, I now turn to the question of how all this relates to the aforementioned puzzle raised by Kripke and Giaquinto.

There is a significant structural difference between our lexical numerical system and decimal notation, because each system is structured for a different modality.[37] On the one hand, the lexical system is structured for speech and hearing. On the other, decimal notation is structured so that decimal numerals can be read and visualized easily. The number of digits is small enough to be well within the bounds of what we can remember easily, while still being large enough for the notation to remain so concise that we can, up to a point, survey and

[36] Thanks to Haim Gaifman for this example.

[37] Our lexical system is multiplicative-additive (‘three-hundred and twenty-seven’), while our decimal notation is ciphered-positional (‘327’). See [Chrisomalis, 2010].

visualize multi-digit numerals. Then there is the fact that powers can be represented by the position of a single digit. It is because of these structural features that decimal notation helps us to overcome the limitations of our parsing ability.

Modifying Kripke's proposal that a notation should be structurally revelatory, I claim that it should also be *visually revelatory*. That is, a notation should reveal structural features of its subject matter visually, by helping one to see or visualize them. Further, I claim that there may be some tension between the requirement of having a visually revelatory notation and the requirement of having a structurally revelatory one, with the result that there is a trade-off between the two requirements. For example, decimal notation is visually revelatory, because we can visualize the decimal numerals in order, and this reveals, visually, the ordering of a progression of numbers. Thus decimal notation is also somewhat structurally revelatory. But as a result of being visually revelatory, and so structured to be read and visualized (as described in the previous paragraph), it is not as structurally revelatory as it might otherwise be. For example, it is not as structurally revelatory of a progression as stroke notation, which is not structured to be read and visualized.

As an aside, the case of infix notation is an example of a visually revelatory notation that improves upon a structurally revelatory one, without the need for a trade-off between the two requirements. For, I have argued, infix and prefix notation are *equally* structurally revelatory of the recursion equations and the associative law for addition, since the relevant infix and prefix statements have isomorphic parsing trees. But because we parse blackboard arithmetic and natural language in infix, not prefix, we find infix notation considerably easier to read, and can visualize repeated successions and additions in infix without conscious effort. We can also visualize these features by reading prefix notation, but only with more conscious effort.

I now turn to my explanation of why decimal notation is immediately revelatory. It is because it is visually as well as structurally revelatory, with the result that we can parse it with little conscious effort. This is why we know what is expressed by decimal notation, but do not know what is expressed by other notations that we understand. Here I am making two empirical claims. Firstly, we have a tacit grasp of the dictionary rule that gives the ordering of our decimal notation, which we can apply, in order to locate numerals in relation to one another, without conscious inference. Secondly, we are also able to visualize the decimal numerals in their dictionary order, with relatively little conscious effort.[38] Of course I concede that these abilities can be exercised for numerals from other systems that we are well practiced at reciting in order. The point is that doing so requires more conscious effort. For example, even after some training in binary notation, one has to discover where 101101 occurs in relation to many other sequences in the binary ordering by reciting or calculating. The position of 101101 is *not* known simply because we can apply understood rules and visualize the numerals without discursive effort.

My claim that subjects, who are trained in a visually revelatory notation, can exercise these abilities with little or no conscious effort, receives support from brain imaging studies of children. Subjects who are learning decimal notation show high levels of prefrontal activity, which is indicative of effort. This vanishes during development as the notation is mastered.[39]

---

[38] Cf. Dehaene and Cohen's claim that we have a visual arabic number form (see section 4).

[39] [Dehaene, 1997, pp. 269-70].

References

Almog, J. and Leonardi (eds.) [2009]: *The Philosophy of David Kaplan.* New York: Oxford University Press.

Burge, T. [2000]: ‘Frege on Apriority’, in [Burge, 2005], ch. 10, pp. 356-387.

-------- [2005]: *Truth, Thought, Reason.* New York: Oxford University Press.

-------- [2007]: *Foundations of Mind.* New York: Oxford University Press.

-------- [2009]: ‘Five Theses on De Re States and Attitudes’, in [Almog and Leonardi, 2005], ch. 15, pp. 246-316.

-------- [2010]: *Origins of Objectivity.* New York: Oxford University Press.

Burgess, J. P. [2003]: *Mathematics, Models and Modality: Selected Philosophical Essays*. New York: Cambridge University Press.

-------- [2003]: ‘Numbers and Ideas’, in [Burgess, 2003], ch 1, pp. 23-30.

Butterworth, B. [1999]: *What Counts: How Every Brain is Hardwired for Math.* New York: The Free Press.

Cambell, J. (ed.), [2005]: *The Handbook of Mathematical Cognition.* New York: Psychology Press.

Cantor, G. [1885]: ‘Review of Frege’s Die Grundlagen’, in [Ebert and Rossberg, 2009], pp. 341-8.

Carey, S. [2009]: *The Origin of Concepts.* New York: Oxford University Press.

Chrisomalis, S. [2010]: *Numerical Notation: A Comparative History*. New York: Cambridge University Press.

Dehaene, S. [1997]: *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford University Press.

-------- [2001a]: ‘Précis of *The Number Sense'*, *Mind and Language*, **16**, 16-36.

-------- [2001b]: ‘Author’s Response: Is Number Sense a Patchwork?’, *Mind and Language* **16**, 89-100.

Dehaene, S. and Changeux [1993]: ‘Development of Elementary Numerical Abilities: A Neuronal Model’, *Journal of Cognitive Neuroscience* **5**, 390-407.

Dehaene, S. and Cohen [1995]: ‘Towards an Anatomical and Functional Model of Number Processing’, *Mathematical Cognition* **1,** 83-120.

Ebert, P. and Rossberg [2009]: ‘Cantor on Frege’s Foundations of Arithmetic’, *History and Philosophy of Logic*, **30**, 341-8.

Franks, N.R. Dornhaus, Metherell, Nelson, Lanfear and Symes [2006]: ‘Not Everything that Counts Can be Counted: Ants Use Multiple Metrics for a Single Nest Trait’, *Proc. R. Soc. B* **273**, 165-169.

Frege, G. [1884/1953]: *The Foundations of Arithmetic.* New York: Harper and Brothers.

Gallistel, C. R. [1990]: *The Organization of Learning.* Cambridge: MIT Press.

-------- [2011]: ‘Review of *The Origin of Concepts*’, *Lingua* **121**, 308–312.

Gallistel, C. R. and Gelman [2000]: ‘Non-Verbal Numerical Cognition: From Reals to Integers’, *Trends in Cognitive Sciences* **4**, 59-65.

Geary, D.C., Berch and Koepke (eds.) [2016]: *Evolutionary origins and early development of number processing*. Elsevier/Academic

Gelman, R. and Gallistel [2015], in [Geary, Berch and Koepke 2016], Foreword, pp. xii-xix.

Giaquinto, M. [2007]: *Visual Thinking in Mathematics*. New York: Oxford University Press.

Hahn, M. and Stoness (eds.) [1999]: *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society.* Lawrence Erlbaum Associates.

Kripke, S. A. [1992]: ‘Logicism, Wittgenstein and *De Re* beliefs about Natural Numbers.’ In preparation.

-------- [2011a]: ‘Frege’s Theory of Sense and Reference’ in [Kripke, S. A. 2011b], 254-91.

-------- [2011b]: *Philosophical Troubles: Collected Papers Vol I*. New York: Oxford University Press.

Mandelbaum, E. [2013]: ‘Numerical Architecture’, *Topics in Cognitive Science* **5**, 367-386.

Marshall, O. R. [2016]: ‘Giaquinto on Acquaintance with Numbers’, *Journal of Philosophy,* to appear.

Nagel, T. [1997]: *The Last Word.* New York: Oxford University Press.

Peacocke, C. [2009]: ‘Frege’s Hierarchy: A Puzzle’, in [Almog and Leonardi, 2005], ch. 10, pp. 159-186.

Platt, J.R. and Johnson [1971]: ‘Localization of position within a homogenous behavior chain: Effects of error contingencies’, *Learning and Motivation* **2**, 386-414.

Rips, L. [2011]: *Lines of Thought.* New York: Oxford University Press.

Rocha, A., Massad and Pereira [2005]: *The Brain: Fuzzy Arithmetic to Quantum Computing.* Springer.

Russell, B. [1940]: *An Inquiry into Meaning and Truth.* George Allen and Unwin Ltd.

Shapiro, S. (ed.) [2005]: *The Oxford Handbook of the Philosophy of Mathematics and Logic.* New York: Oxford University Press.

Shwayder, D. [2008]: *Statement and Referent: An Inquiry into the Foundations of our Conceptual Order.* Stanford: CSLI Publications.

Steiner, M. [2005]: ‘Mathematics, Explanation and Scientific Knowledge’, *Noûs* **12**, 17-28.

-------- [2005]: 'Mathematics—Application and Applicability', in [Shapiro, 2005], ch. 20, pp. 625-650.

Tennant, N. [1987]: *Anti-Realism and Logic.* Clarendon Press.

Whalen, J. Gallistel, and Gelman [1999]: 'Nonverbal Counting in Humans: The Psychophysics of Number Representation', *Psychological Science* **10**, 130-137.