Naturalising Mathematics? A Wittgensteinian Perspective

Jan Stam 1,*, Martin Stokhof 2,3 and Michiel Van Lambalgen 2

1 Department of Neurology, Amsterdam UMC, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands
2 ILLC/Department of Philosophy, University of Amsterdam, 1012 GC Amsterdam, The Netherlands; m.j.b.stokhof@uva.nl (M.S.); m.vanlambalgen@uva.nl (M.V.L.)
3 Department of Philosophy, Tsinghua University, Beijing 100084, China
* Correspondence: j.stam@amsterdamumc.nl

Abstract: There is a noticeable gap between results of cognitive neuroscientific research into basic mathematical abilities and philosophical and empirical investigations of mathematics as a distinct intellectual activity. The paper explores the relevance of a Wittgensteinian framework for dealing with this discrepancy.

Keywords: neuroscience; mathematics; Wittgenstein; naturalism

1. Introduction

Research in cognitive neuroscience on mathematics is largely focused on basic mathematical abilities, such as counting, measuring, and spatial orientation. Higher mathematics has remained by and large outside its scope; it is studied mainly from historical, social, cultural, and philosophical perspectives. Thus, a gap appears between the results of the former and those of the latter.

This paper is concerned with the question of how this gap can be bridged. The leading idea is that a suitable conceptual framework can be culled from the work of Wittgenstein on the philosophy of mathematics and, more generally, that on epistemic practices. Wittgenstein’s analyses combine observations on natural abilities and (broadly) cultural dimensions in a unified framework and connect with a 4E approach to cognition [1] that transcends some of the limitations of neurocognitive research. By viewing the results from cognitive neuroscience from this perspective, we gain insight both into the content and scope of neuroscientific results and into the potential relevance of a Wittgensteinian naturalistic approach in the analysis of mathematics.

The paper is structured as follows. We start with an overview of the results of cognitive neuroscience on basic pre-symbolic numerical abilities (Section 2). Then we frame these results in terms of the naturalistic elements of a Wittgensteinian framework (Section 3). Next, we analyse the gap in terms of the role that training and education play in the Wittgensteinian framework and relate the results to insights from developmental psychology and other empirical disciplines (Section 4). In Section 5, we explore attempts to explain the connection between basic numerical abilities and the socio-cultural development of mathematics. Then we go back to some of Wittgenstein’s views on mathematics and their naturalistic implications and look at their relevance for this issue (Section 6). We end with conclusions and suggestions for further research.

2. Cognitive Neuroscience of Quantitative Competence

2.1. Pre-Verbal Cognition of Quantity

The early use and acquisition of numerals has been studied extensively with various methods and much is still controversial or simply unknown. We discuss the most stable findings and mention some controversies relevant for our argument.
An ability to estimate small quantities of objects, called the ‘Approximate Number System’ (ANS)$^1$, has been demonstrated in animals, in pre-verbal children, and in cultures without number language. Newborn babies react to relative differences between quantities of visual objects or sounds. The ratio of the difference that can be detected increases gradually with age, from 1:3 at birth to 3:4 in the fourth year up to the adult 7:8 ratio [3]. Anthropological studies of cultures without a symbolic number system confirmed that the ANS is a universal human ability [4]. Many vertebrate animal species and insects exhibit ANS abilities, and field studies show that animals estimate and compare quantities of conspecifics or food items [5]. Moreover, animals can be trained to react to sums and differences and even to the empty set (zero) [6]. A variety of experimental techniques (fMRI, Event Related EEG Potentials, single neuron recordings) show increased neural activity during numerosity estimation tasks in the human intra-parietal sulcus (IPS) and the prefrontal cortex, and in homologous cortical regions of monkeys. Number neurons in the latter respond selectively, but noisily, to the quantity of visually presented small sets [2]. It has been argued that the ANS is not dedicated to estimating discrete quantities but is a manifestation of a general system for estimating continuous magnitudes, such as size and duration. Walsh proposed a parieto-frontal system for guiding action by information about space, time, and discrete quantity, extracted by the IPS from sensory input [7]$^2$. A recent review of the experimental evidence concluded that the ANS is selectively tuned to discrete quantities [9], but others support the idea of a single system for discrete and continuous magnitudes [6].

Besides the approximate number system, there is a mechanism for accurate and rapid recognition of the exact quantity of small sets, called ‘subitizing’, described by Wittgenstein as ‘groups of objects that can be taken in at a glance’ (PI, 9)$^3$. Subitizing may be a manifestation of a more general-purpose object tracking system (OTS), a neural mechanism for tracking small sets of objects through space and time, constrained by the capacity of the visuo-spatial short-term memory of about 3–4 items [3]. Subitizing is demonstrated in animals and pre-verbal infants. It matures more rapidly than the ANS, and around their first birthday, children reach the adult limit of 3–4 items$^4$.

Neuro-cognitive aspects of geometry have been studied by similar methods as applied for numerical abilities. Individuals without geometric language or schooling, such as young children and adults from cultures without symbolic geometry or maps, can recognize differences in geometrical features among sets of otherwise similar figures, for instance between straight and skewed angles and between parallel and nonparallel lines [17]. The impressive navigation and construction skills of many animals suggest that they possess a kind of ‘natural geometry’. The discovery of neurons which become selectively active when rats or humans visit certain locations are evidence of a cerebral positioning system$^5$. These observations inspired the idea of universal inborn ‘core systems of geometry’ [18].

2.2. Limitations of Neuro-Cognitive Research

Neuroscience mainly studies individual subjects, isolated from their natural and social surroundings, and with a small range of possible movements. Brain imaging requires that subjects lie in the narrow core of the MRI magnet where they can hardly move. Single neuron recording in animals often requires restraining the subjects, and extensive training before they accept the experimental conditions and perform the standardized tasks needed to make reproducible measurements. These limitations may blind neuroscience to the contextual and social dimensions of behaviour and the cognitive aspects of action and interaction. favouring perception experiments in individual subjects can reinforce philosophically untenable intuitions, for instance that complex cognitive behaviour can be fully explained by processes in individual brains$^6$. It reproduces the neglect in the traditional philosophy of mind of situational and social embedding and practices (‘enactivism’), which are central considerations in Wittgenstein’s later thinking. Psychological experiments have their own limitations. Training animals to react to sums or differences of quantities, interesting as it is, does not show that they possess arithmetic skills. Looking time experiments
with babies, as evidence of their reactivity to differences in quantities, are hard to replicate and their interpretation is controversial [19].

Besides these experimental issues, neuroscience faces theoretical problems. One of them is the tendency to construct overarching theories of how the brain or the mind works based on limited data. Some of these theories contain dualist philosophical residues. A key example is the widespread use of the term ‘representation’ to describe processes in the brain. Uncritical use of the term confuses brain states with mental or symbolic content, obfuscating the huge explanatory gaps between the two [19]. Representational expressions may be innocent metaphors, but they often reveal an underlying intuition that to know something or to be able to act requires some sort of copy of (relevant parts of) the world in the brain or the mind. This mentalist view, nowadays mostly disguised as brain talk, was one of the targets of Wittgenstein’s criticisms.

Another problem is the habit to describe the results of experiments with animals or preverbal humans using terms borrowed from symbolic mathematical knowledge. The use of expressions such as ‘number sense’ [2] or ‘natural geometry’ [18] to characterize the nonverbal ability to react to discrete or continuous magnitudes suggests that animals or babies know about numbers or measures, or even Euclidian geometry. However, geometry is a mathematical technique applied to idealized pictures of reality, by far surpassing the basic abilities observed in those experiments [21].

How to connect the nonverbal systems of quantity perception causally with the human symbolic number system is an instance of the ‘symbol grounding problem’, a central problem in the philosophy of mind and language and one that is far from solved [22]. There is a huge gap between a nonverbal grasp of quantity and the ability to use number words and to count or measure. We will explore the development of symbolic abilities later (Sections 4 and 5), but first we discuss how Wittgenstein’s views may help to understand the nature of mathematics.

3. Wittgensteinian Naturalism and Mathematics

3.1. Wittgenstein and Naturalism

In recent years, there has been an increased interest in the role played by naturalistic considerations in Wittgenstein’s work. Although there are dissident [23] and sceptical [24] voices, it appears to be widely accepted that there are indeed such considerations, and a variety of analyses of them can be found in the literature. Rule following is a central topic [25], as are sensations [26], ethics [27,28], and mathematics [25,29].

The status of Wittgensteinian naturalism has been questioned. Dromm argues that it is ‘imaginary’, by which he means that it is first and foremost a methodological tool and not an appeal to facts about nature (human and/or physical) [23]. Baker [24] observes a tension with Wittgenstein’s emphatic rejection of the idea that philosophy formulates theses and provides explanations. However, other authors see a connection with Wittgenstein’s views on science and philosophy [30,31].

A complicating factor here is that not all authors understand naturalism in the same way. For example, Medina [32] makes a distinction between the kind of naturalism proposed by Quine and ‘anthropologism’, as conceptualised by Jacquette [33], and, quite rightly, dismisses the idea that Wittgenstein would endorse naturalism in the Quinean sense. The scientism inherent in the latter is alien to Wittgenstein’s views [34]. Liberal naturalism as defined by Macarthur [30] seems more akin to Wittgenstein’s conception.

Most congenial with the undertaking of this paper is the view that ties Wittgenstein’s naturalism to current 4E-theories of human cognition [36–38]. However, we will argue that the social dimension, especially the role of socialisation and education, is at least as important [39–41].

3.2. Wittgenstein and Practices

To provide a conceptual framework, we need to analyse where and how exactly naturalistic considerations enter into Wittgensteinian analyses. We focus on *On Certainty*...
(OC), noting that although it is sometimes considered to represent a distinct stage in the development of Wittgenstein’s thought [42], there is ample evidence that core ideas of OC go back to his work from the 1930s, which justifies also using other work from Wittgenstein [43,44].

The concept of a practice plays a central role in Wittgenstein’s later work. Practices can be of different kinds, and no general definition is forthcoming, but a couple of characteristic features can be identified from Wittgenstein’s descriptions: practices come with ‘a point’; they constitute meaning; they have both natural and social dimensions.

‘Having a point’ means that a practice allows for the question ‘Why do we engage in this?’ It should be taken in a broad sense (Wittgenstein includes singing rhymes, telling jokes (PI, 23)), but not everything goes. Practices constitute meaning in a broad sense, not just of what we say but also of what we do, and thereby they normatively constrain our doings and sayings. Practices themselves are determined by both social and cultural factors as well as natural ones. The former account for socio-cultural diversity and historical change, the latter for the stability and relative socio-cultural invariance of basic practices.

It is the third characteristic that provides a vector for naturalistic considerations. As Wittgenstein indicates in OC, practices are constituted by what he calls ‘certainties’. The concept of a certainty is complex, but some shared characteristics are forthcoming. First, certainties are not beliefs, they are not part of epistemic practices, and thus are not subject to doubt, justification, and so on. Second, although some certainties have linguistic counterparts, most of them are manifested in our ways of acting. Finally, certainties are not characterised by content but by function. Something’s being a certainty is not a matter of ‘what it is about’ but of how it functions: being exempt from doubt, assumed in the background, constitutive of a certain practice.

The framework of certainties that a community entertains constitutes its practices. As is evident from contemporaneous diversity and historical change, there is a plurality of such frameworks. However, this does not entail radical relativism: certainties, in their turn, are constrained by ‘very general facts of nature’ (PPF, xii, 365): the way the world is and the way we as human beings are, i.e., our physiology and basic psychology, constrain what makes sense for us to entertain as certainties.

The constraints that nature, broadly conceived, imposes can force some certainties while only ‘suggesting’ others. The latter relate to the point of a practice, and it is here that the possibility of change enters the picture. Nature may change, we may change, our needs may change, our knowledge of the world changes, and that may result in a change in our practices. The purpose they serve may become obsolete, or different ways of achieving the purpose may become available.

3.3. Mathematics as Practice

What does mathematics as practice come to? In the end, this calls for an extensive analysis of all that Wittgenstein has written about mathematics, but that is beyond the scope of this paper. Here we limit ourselves to reviewing the three general characteristics of practices outlined above, and then examine whether the concepts of certainty and natural constraints can fit into a conception of mathematics that can accommodate a natural dimension.

‘Having a point’ is a key consideration for Wittgenstein. What makes a practice one of mathematics is not the use of a particular set of expressions or symbols, or of operations on such. A fortiori it is not a matter of these expressions referring to a particular kind of entities. Actions—including, but not limited to, the use of expressions—are mathematical because they serve a particular kind of purpose. This is most obvious in the case of basic counting and measuring. These actions are what they are used for—concrete, mundane purposes that are shared across various practices.

For higher mathematics ‘having a point’, is much less obviously practical in the mundane sense. Here, the distinction between ‘pure’ and ‘applied’, which does not play any role at the level of basic counting and measuring, becomes relevant. One could surmise
that the point of applied mathematics is a derivative of the points of the practices in which a particular piece of mathematics (method, result) is applied. In pure mathematics, the point is internal to the practice itself\(^\text{19}\). A particular area in pure mathematics comes with its own, self-generated standards for what makes sense, what is important, what constitutes progress and success, and even beauty (RFM I, 166–167). In that respect, the practices of pure mathematics resemble certain artistic practices—an intuitive association that is not uncommon.

‘Constitute meaning’ is quite clear in the case of higher mathematics: what makes a particular procedure or statement meaningful is internal to the practice. For basic counting and measuring that may be less obvious, but this might very well be due to their being dispersed practices. After a certain stage of initial training, we hardly ever count or measure just for the sake of counting or measuring. Normally, counting and measuring occur as elements in a more encompassing practice\(^\text{20}\).

Regarding ‘natural and social dimensions’, viewing mathematics as a practice means assigning it an intrinsically social dimension: mathematics is learned, taught, practiced. This manifests itself in pluralism, and Wittgenstein takes that quite seriously. The case of the wood sellers who price their ware by the ground surface of a pile, and not by its volume (LFM, 202; RFM, I, 147–151), is a prime example of a practice involving ‘counting and measuring’ that is quite different from ours, and hence not readily understandable. Wittgenstein’s insistence that the rationality of their practice cannot be judged by the criteria that are internal to our practice clearly shows his commitment to pluralism\(^\text{57}\).

Finally, what about the natural dimension? The key here is the observation that in Wittgenstein’s analysis, a practice always assumes a shared set of capacities, to enable teaching and learning, and to validate and maintain a practice as something that a community of people is meaningfully engaged in. Without these natural abilities and their development, we could not enter the stage of learning more complex behaviour\(^\text{21}\). In the next sections, we will look at scientific findings that are consistent with this view.

4. Training and Education

4.1. Learning Number Names and Counting

The pre-symbolic numerical abilities, subitizing and the ANS, which we share with many animal species, are mostly innate, but mastering the correct use of symbolic (oral or written) numbers requires a long process of training and learning. Around their third birthday, most children correctly use the word ‘one’ and many can recite the numerals up to ten, but they cannot use them to count. About a year later, they usually can count small sets in the subitizing range, up to four. Counting to 10 is generally achieved around 5 to 6 years\(^\text{22}\). Correct use of the number words, just like mastering colour words, develops slowly compared to the rapid increase in the general vocabulary. These delays may be a consequence of the so-called ‘segregation problem’\(^\text{59}\)\(^\text{23}\). Children are confronted with countless (literally) different set sizes and varieties of hues. Properties such as number or colour are not encountered independently but immersed in a large manifold of other properties. This precludes simple association learning, such as by ostensive teaching (pointing and naming)\(^\text{24}\). Ramscar et al.\(^\text{59}\) argue that children learn abstract concepts by competitive discrimination, a mechanism of prediction and error to match the words they hear with the appropriate selection from environmental cues\(^\text{25}\).

An area of controversy is whether learning to count is dependent on or secondary to learning language. One theory, based upon Chomskyan linguistics, is that supposedly innate grammar rules or principles such as place-coding are also applied in number systems. The idea that complex rules are somehow genetically encoded in the neonate’s cerebral wetware is biologically implausible and is largely abandoned. Nonverbal quantity cognition and its early development are probably innate—as are, of course, the neural networks, which in humans continue to develop and enable the enormous learning capacity of children. This could be called ‘weak nativism’ as opposed to Chomskyan strong nativism. Rejecting strong nativism entails that more explanatory work must be attributed to devel-
opmental mechanisms and learning, such as anatomically developing neural networks and changing functional (synaptic) connectivity. Of course, language is indispensable in learning more abstract concepts.

4.2. Learning Experiments ‘In Silico’

Zorzi and Testolin examined the issue of nativism versus early learning of nonsymbolic number cognition in two computer simulations [62]. In a simulated evolution (A), simple ‘organisms’ living in a limited world had a small neural network that recognized food units and enabled movements to find and consume them. An intermediate network layer connecting input and output could mutate at random. Selection pressure for finding more food units resulted in offspring that was able to recognize the number of food units. In a second experiment (B), a single multilayer neural network was exposed to arrays of different amounts of dots. The network spontaneously learned to respond to the numerosity of the stimuli with minimal instruction. Its behaviour mimicked human and animal findings, such as a gradual increase in accuracy similar to that of the ANS. After training, number selective neurons emerged in intermediate network layers. Their response pattern to discrete quantities of dots closely resembled that of neurons in the IPS of monkeys.

Both experiments are supportive evidence for two explanatory theories of nonsymbolic number recognition: (A) evolution and nativism, and (B) spontaneous learning with minimal instruction. Both simulations resulted in systems with a number sense, but system B agreed better with behavioural and experimental data in animals. These observations support the view that both evolution and postnatal development can be conceived as learning processes, on vastly different time scales [63]. In humans, these processes result in limited innate abilities and extensive postnatal maturation and learning ([64], pp. 312–315), but the relative contributions of nature and nurture to numerical cognition is controversial [65].

4.3. Learning to Calculate

After nonverbal numerical cognition and informal pre-school learning, we enter the domain of formal education. In what follows, we review some observations which we think are relevant for the connection with Wittgenstein’s views. Having mastered counting to about 10, children enter primary school and start the long process to master elementary arithmetic and geometry. For those of us who forgot how much we learned, we summarize what children are supposed to master at the end of primary school, at 12 years of age in most countries (Table 1).

Table 1. Global educational objectives for mathematics at the end of primary school in the Netherlands. Extracted from [66].

<table>
<thead>
<tr>
<th>Arithmetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Names of numerals; digits; decimal position system and notation; concept of zero; number sequence and number line; ordinal relations; insight in orders of magnitude up to ca. one billion; estimating; fractions, decimals, and conversion ($\frac{3}{4} = 0.75$, etc.); rounding; calculating (+, −, ×, ÷) with small numbers; multiplication tables to $10 \times 10$ by heart; using an electronic calculator; order of arithmetic operations, grouping, brackets.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length, surface, volume (estimating, measuring, calculating); metric system; temperature; weight; time (units and intervals, clock, calendar); instrument readings; maps, scale, distance, route planning; simple 3D objects; 2D ↔ 3D projections; simple formulas (velocity, etc.); basic graphs and tables; symmetry; tiling.</td>
</tr>
</tbody>
</table>

Abilities such as reciting the multiplication tables are taught by training, in Wittgenstein’s strong sense of ‘Abrichtung’ [27]. For competencies such as basic arithmetic, the exact nature of the learning process is less evident. Adding two numbers a child never added before requires mastering the procedure of addition, but explaining beforehand the procedure or the rules to be followed will not work, just as we do not learn language by first having the grammar explained [28]. Children start with small and easy sums and a lot
of drill and practice, and not by first being taught general principles, such as the law of commutativity\textsuperscript{29}. This results in their ability to follow (simple) rules, which is still different from them having acquired knowledge of mathematical facts\textsuperscript{30}. When a child finally masters basic arithmetic, we say that it knows the principles. It has learnt the procedure, not before learning but while learning. As Wittgenstein observed, ‘We got to know the nature of calculating by learning to calculate’ (OC, 45). This is a special kind of knowledge, a practical know-how, rather than factual knowledge, because few children will be able to explain the principles, even though they master them\textsuperscript{31}. It is about knowing how to do something, which is an ability that is not derived from having acquired a prior, innate, or mental set of rules or principles. To know ‘how’ instead of to know ‘that’ is at the core of Wittgenstein’s rule-following considerations (PI, 138–242)\textsuperscript{32}.

5. Second Nature: The Socio-Cultural Acquisition of Symbolic Abilities

As we enter more complex domains of human behaviour, neuroscientific explanations fall short, and the biggest challenge to any kind of naturalism is to provide explanations at the socio-cultural level. The explanatory gap between biology and a cognitive achievement such as mathematics, which is transmitted, refined, and expanded over many generations, seems unbridgeable. We discuss two approaches which show how naturalistic explanations of cultural achievements could be developed but also demonstrate the huge lacunae in that kind of understanding of complex social behaviour and culture\textsuperscript{33}.

5.1. Connecting Culture to Brains

‘Cultural recycling theory’ proposes that symbolic cognition and behaviour reuses pre-existing cerebral networks, which developed through pre-cultural evolutionary mechanisms\textsuperscript{67}. The idea is that culturally invented and transmitted tools, such as the symbolic number system, make use of cerebral systems which had developed to serve biological functions. The central argument for this theory is the existence of cortical regions which are indispensable for certain symbolic abilities. This causal connection cannot be explained by evolution because the cultural development of approximately 6000 years is too brief for adaptation of the brain by natural selection. Therefore, some cortical structures, having first slowly developed by natural selection to enable certain biological functions, become hubs for rapid cultural developments. Such a change in function of a structure to enable new capacities, called ‘exaptation’, is a well-established biological phenomenon\textsuperscript{34}.

The strongest case for cultural reuse is a specialised region in the left occipito-temporal cortex, the Visual Word Form Area (VWFA). It is consistently activated by reading letters and words, across different cultures and writing systems. Injury to this area or its cortical input (disconnection) damages the capacity to read, while writing and speech are preserved, a condition called ‘pure alexia’. The VWFA is closely linked to cortical areas involved in high-resolution vision and object recognition. The hypothesis is that the VWFA, by virtue of its strategic position and its original function, is optimally prepared to serve as a cortical centre for reading\textsuperscript{67}. Another example is the consistent involvement of the intraparietal sulcus (IPS) in nonverbal quantity perception, across cultures and primate species. This cortical area might be reused in the development of symbolic number processing.

These theories are highly speculative and hard to confirm empirically. The cultural recycling theory leans heavily on the specificity of the VWFA for reading. It is, however, involved in many other capacities, such as auditory word repetition and colour naming and is, of course, an integral part of larger multifunctional networks. Price and Devlin\textsuperscript{68} therefore called the concept of a specialized VWFA a myth. Be that as it may, we think the cultural recycling theory presently has too many loose ends to be a plausible candidate for closing the biology–culture gap.

5.2. From Tool Use to Symbol Manipulation

Exaptation presupposes that some cortical networks are equipped to accommodate rapidly evolving cultural symbolic abilities. It can however not explain the emergence of
symbolic systems per se, the socio-genetic question. A promising, enactivist approach of this problem is based upon a historical reconstruction of the early development of number systems [69]. This approach, called ‘cognitive archaeology’, tries to reconstruct the emergence of symbolic cognition from simple practices and techniques. External tools for counting, such as fingers, pebbles and tallies are plausible candidates for primitive external number systems. Tally sticks have the advantage of providing a record, and very old, possibly prehistoric specimens have been found. Mesopotamian clay tokens of different shapes (later called ‘calculi’), dating from the fourth millennium BC, made it possible to record larger numbers by grouping. Tallies and tokens function as an external memory, spatially and temporally distant from the originally counted objects, substantial abstractions compared to finger counting. Archaeological findings suggest a gradual transition from clay tokens to images of the tokens on the outside of the clay containers used to keep them, and from there to simple inscriptions. The latter were intermediate stages towards the first written number systems in cuneiform [69]. This process demonstrates a plausible development of a symbolic number system from external counting tools. A comparable enactivist account has been proposed for the emergence of abstract geometrical concepts, such as parallel lines, square angles, circles, etc., from external tools used in measurement and construction [21].

One big question of course remains: How can external counting systems, first material and proximal to the objects to be counted, later abstract, distal, and symbolic, give rise to mental symbolic operations such as calculating in the head? Cognitive archaeology cannot answer this, but we think a radical enactivist account could be developed to provide naturalist explanations of mental phenomena: external tools as precursors of external symbols, which subsequently become mental ‘off-line’ symbolic tools. This process may occur on the timescale of cultural development but also on the timescale of individual development. It is tempting to speculate that the cultural (phylogenetic) development of societies is reproduced in every child on the ontogenetic timescale [37]. If the phylogeny–ontogeny analogy has any merit for the cultural domain, then the relevant ontogenesis is not prenatal, but postnatal—the long process of acculturation, of learning to read, write, count, and calculate. For oral language, the recapitulation analogy fails because the child grows up in an environment already saturated with language, slowly developed during the phylogenetic eons before it was born.

Despite these differences between the phylogenetic and ontogenetic processes, enactivism may help to explain the acquisition of symbolic capacities on both timescales. Most children start counting using their fingers [70] and slowly progress towards the culturally required level of arithmetic. Only after mastering counting with external tools do they become able to use numerical symbols and subsequently perform mental arithmetic. As Wittgenstein observed, ‘Only if you have learnt to calculate—on paper and out loud—can you be made to grasp, by means of this concept, what calculating in the head is’ (PPF, xi, 277). This is of course not a theory of how mental calculation develops from using external tools, but a meta-observation. We can only grasp what mental calculation is if we first have learnt to calculate, by manipulating the external symbols on paper and ‘out loud’. It is an imaginative statement (in Dromm’s [23] sense), which can, however, be transformed into a naturalist view about the enactivist origin of mental processes. Arguably, Wittgenstein makes this very move in the same paragraph when he observes the following: ‘You can only learn to calculate in your head by learning to calculate’ (PPF, xi, 304). In PI 385, he asks this: ‘Would it be imaginable for someone to learn to do sums in his head without ever doing written or oral ones?’ The philosophical error, one might say, the main target of Wittgenstein’s critique, is that the order is reversed. Instead of action first (’in the beginning was the deed’; OC, 402), the philosophical tradition had put the mental first.

6. Wittgenstein and Mathematics

That mathematics holds a special place in Wittgenstein’s oeuvre may not be obvious, as much of it deals with a whole range of other topics. However, Wittgenstein himself was
Philosophies 2022, 7, 85

quite insistent, stating about himself that “Wittgenstein’s chief contribution has been in the philosophy of mathematics” ([71], p. 466).

If one works through Wittgenstein’s remarks or reads the reports of his lectures, one is struck by the heterogeneity of the topics that he addressed, which range from observations about basic counting and addition to remarks on Cantor’s diagonalisation argument and Gödel’s incompleteness theorems. Puzzling as this may be for a reader trained in the philosophy of mathematics, for Wittgenstein, apparently this motley of topics all belonged together.

A persistent element in Wittgenstein’s thinking about mathematics is his radical antirealism, which can be traced from the Tractatus up to his latest remarks. When he famously compares the treatment of philosophical questions with the treatment of an illness, the exemplary disease is mathematical realism (PI, 254–255). His principal answer to mathematical realism is a particular kind of constructivism. Mathematicians do not discover facts waiting somewhere to be discovered; they invent them (RFM I, 168). The interpretation of his view of mathematical propositions—equations, theorems, etc.—is controversial, but one picture transpires: he compares them to imperatives, or normative rules of expression, with a remote empirical origin [29,73]. They are ‘invented to suit experience and then made independent of experience’ (LFM, 43). Thus, Wittgenstein characterizes arithmetical propositions as empirical propositions hardened or fossilized into rules or paradigms (RFM VI, 22–23; OC, 657). These rules are neither arbitrary nor necessary; our experience, i.e., nature and our access to it, provides the basis for a variety of practices. Thus, the unassailability of mathematics is not due to its subject matter but to its function; mathematical propositions are certainties (OC, 650–655) that constitute practices. Like the rules of chess, we do not make them up, but we have inherited them (LFM, 143), and from that perspective, mathematics can be conceived as an anthropological phenomenon (RFM VII, 33). The constructivism of Wittgenstein, with the remote origin of mathematical expressions being located in experience and practices, and the anthropological view on their function, is a useful point of departure for developing a naturalist conception of mathematics that can accommodate the empirical findings that we reviewed above.

Another naturalistic dimension originates with Wittgenstein’s frequent referrals to teaching and learning. Although he explicitly denies seeking explanations or ‘doing natural history’ (PI, 109; PPF, xii, 365), he repeatedly refers to how children learn words and arithmetic. Some of these remarks can be interpreted as imaginary examples that are used to change the readers’ way of looking. However, many of his pedagogical observations also serve, as we read them, to underscore Wittgenstein’s views of the nature of language, counting and arithmetic as tools that we must learn to use. When he sketches a simple language in a toy world (PI, 1–10), he observes that ‘a child uses such primitive forms of language when it learns to talk’ (PI, 5). It learns not through explanation but by training (‘Abrichtung’). Training in this simple world often is by ostensive teaching: pointing to objects, uttering their names and learning their use (PI, 6). Even the first numerals can be taught by ostension because of our ability to perceive small ‘groups of objects that can be taken in at a glance’, without counting. According to Wittgenstein, ‘Children do learn the use of the first five or six cardinal numerals in this way’ (PI, 9). This is not an imaginary example but an empirical observation by someone with actual teaching experience with children of that age.

These Wittgensteinian perspectives can, in our view, be developed into a proto theory of naturalized mathematics. Such a theory must globally explain how an individual learns numbers, counting and calculating as a necessary base for acquiring more advanced mathematical skills, but also how entire cultures developed mathematics as we know it. Framed in biological terms, it should sketch an ontogenetic and phylogenetic (including cultural) development which could explain the mathematical abilities of individuals and societies.
7. Conclusions and Further Research

The scientific results we summarized above may form the nucleus of the development of a naturalistic conception of mathematics. The philosophical groundwork for such explanation, we argued, has been laid by Wittgenstein’s insights, which demystified the traditional aura of mathematics as a system of eternal, unassailable truths. A Wittgensteinian naturalism, which has both natural and socio-cultural dimensions, may provide a suitable framework for analysing how mathematics works, both at the basic level of counting and measuring, as well as in the increasingly complex and abstract practices of higher mathematics.

Of course, it would be odd to suppose that all of mathematics can be naturalised in one of the senses discussed in this paper, but one would like to have some grasp of what makes a mathematical subject naturalisable (or not).

To this end we introduce an admittedly coarse distinction between mathematics that is primarily concerned with quantity, and mathematics dealing with structure. This is no hard and fast distinction, but the following will do for our purposes.

The first type of mathematics (which dominated until Gauss) looks at mathematical representations of extensive and intensive observable magnitudes (for example, speed). These representations are typically computable functions, reflecting their intended use in producing scientific, quantitative predictions. Computable functions can be conceptualized as formulas. Their relevance to our theme is that computability provides an intersubjective constraint on the meaningfulness of definitions and theorems; if one adopts the Church–Turing thesis (as most mathematicians do), then computability is unique and is a constitutive part of a practice.

This is very different for the part of mathematics that falls under the heading of ‘structure’. Here one proceeds by abstraction, isolating common features of systems of quantities. The natural numbers \( \mathbb{N} \) and the rational numbers \( \mathbb{Q} \) have very different properties with respect to the ‘less than’ order, and \( \mathbb{N} \) is a part of \( \mathbb{Q} \) but not conversely. However, viewed purely as sets and ‘forgetting’ all other structural properties, \( \mathbb{N} \) and \( \mathbb{Q} \) are equivalent in the sense that there is a bijection between them. One may take the existence of a bijection between sets \( A \) and \( B \) as defining the relation ‘\( A \) and \( B \) have the same cardinality’, taking one’s cue from the case where both \( A \) and \( B \) are finite. Doing so entails that the natural numbers \( \mathbb{N} \) and the rational numbers \( \mathbb{Q} \) have the same cardinality, over-riding the part–whole relationship. There is nothing inevitable about this definition, though; it is possible to construct a theory of cardinality of (denumerably infinite) sets in which \( \mathbb{Q} \) does have a larger cardinality than \( \mathbb{N} \).[48] We are now in the realm of creativity and freedom, aided by a good deal of powerplay: the freedom of fellow mathematicians to propose their own definitions is not always recognised.[44]

This is just one way of approaching the multi-faceted nature of modern mathematics, which is really a conglomerate of many different practices, rather than one, homogeneous, intellectual discipline. Despite this diversity, there are also commonalities. All mathematical practices have socio-cultural dimensions, which accounts for the pluralism of modern mathematics. However, there are also basic mathematical abilities, grounded in physical and human nature, that appear everywhere, as dispersed practices in the context of complex integrative practices.[45] As such, they may have different roles to play in different settings, but they do create a naturalistic dimension of even the most advanced mathematical practices.

**Author Contributions:** Conceptualization, J.S., M.S. and M.V.L.; investigation, J.S., M.S. and M.V.L.; original draft preparation, J.S. and M.S.; writing—review and editing, J.S., M.S. and M.V.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.
Notes

1. Also called ‘number sense’ [2]. We regard this expression as shorthand for a perceptual system enabling organisms to behaviourally react to (‘recognize’) different discrete quantities of objects.

2. Leibovitch et al. emphasized the difficulty to separate in experiments the numerosity of discrete quantities (dots, sounds) from the magnitude of the stimulus (circumference, surface, duration), which could make the question of the selectivity of the ANS for numerosity unanswerable. Their paper attracted more than 30 pages of critical peer comments [8]. Most experts stick to the concept of a specific pre-verbal number sense.

3. According to Wittgenstein, ‘One of the most dangerous of ideas for a philosopher is, oddly enough, that we think, with our heads or in our heads.’ (Z, p. 605).

4. Whether subitizing is continuous with the ANS is controversial. Its speed and accuracy versus the slower and inaccurate ANS argue for separate systems [15]. In fMRI experiments, subitizing and ANS tasks activate different regions in the parietal cortex [3].


6. According to Wittgenstein, ‘One of the most dangerous of ideas for a philosopher is, oddly enough, that we think, with our heads or in our heads.’ (Z, p. 605).

7. Bennet and Hacker ([20], p. 143) regard the concept of representation as ‘a weed in the neuroscientific garden, not a tool—and the sooner it is uprooted the better’. However, like most weeds, it proves to be quite resilient.

8. Dehaene et al. [17] for instance, reporting their experiments with Amazonian people who lack symbolic mathematical concepts, concluded that ‘core geometric knowledge, like basic arithmetic, is a universal constituent of the human mind’.

9. According to Macarthur ([30], p. 46), ‘a metaphysical quietist version of liberal naturalism, one that avoids supernatural theological commitments in philosophy as well as refusing to give science an unwarranted ontological significance’ This view not only rejects the metaphysics of ‘spooky stuff’ such as souls, eternal platonic forms, etc., but also denies scientism and physicalist reductionism. Nature in this view is not only what can be described by physics or biology but includes ‘human nature’, socialization, education, and our daily practices.


11. A referee asked us why we have not engaged with the vast corpus of Piagetian studies on the development of logico-mathematical knowledge. The main empirical claim of these studies is that cognitive maturation consists in the acquisition of structures (for example mathematical structures such as groups, logical structures such as propositional logic). The main theoretical claim is that the acquisition of these structures is driven by biological processes aiming at homeostasis between an organism and its environment. This is a multi-factorial dynamic process: the organism acts on the environment, monitors the effects of its actions, and adjusts its behaviour accordingly, and all this against the background of biological maturation. The interaction of these factors is bound to generate discrepancies between the organism’s epistemic state and what it discovers in the environment, and Piaget’s theory claims that the resolution of these discrepancies takes the form of a determined progression of identifiable developmental stages. This claim has been a focal point of criticism. Piaget’s theory of cognitive development is thus an example of a domain-general theory, whereas naturalism in the Wittgensteinian mode tends to be domain-specific. More importantly, Piaget downplays the role of language, whereas this is obviously a core issue for Wittgenstein. Since our intention is to explore the Wittgensteinian perspective and to try and link it to cognitive scientific results, we have left the comparison with Piaget’s naturalism (or more broadly, Kantian forms of naturalism) for another occasion.

12. Wittgenstein uses the term ‘language game’, but ‘practice’ is a more neutral term adopted in work on practice theory that is inspired by Wittgenstein (cf., e.g., [45]).

13. Similar considerations, though not phrased in terms of certainties, can be found elsewhere, e.g., in LFM, RFM, PI, PPF.

14. Hence the term ‘hinge proposition’, which goes back to OC 341, 655. For more on the specific approach in epistemology, called ‘hinge epistemology’, that OC has given rise to see for instance [46, 47].

15. In PI 206, Wittgenstein refers to this as ‘shared human behaviour’. The interplay between these natural constraints and various socio-cultural parameters then accounts for the heterogeneity of certainties and different levels of entrenchment. The terms ‘constrains’ and ‘constraint’ should not be over-interpreted: we do not mean to suggest that what we call ‘constraints’ serve to carve out a specific subdomain from a larger whole that is of the same nature. ‘Natural constraints’ indicates those features of nature (including human nature) that are in a sense ‘inescapable’ and that in virtue of that specify a space of possible frameworks of certainties that humans may entertain. To put it differently, there is a multiplicity of world views that humans may have, but all of them are views of the same world and are entertained by humans that share basic features, and because of these two factors, these world views will share certain characteristics.

16. Recent years have seen an increased interest among philosophers and social scientists in the practice of mathematics, i.e., in the social structure of mathematical communities, practices of teaching and training, the pragmatics of proof, the situatedness of the development of mathematical theories, and so on. Cf., e.g., [48–50]. A recent book length treatment is Wagner, 2017 [51]. Some of
the work in this area takes inspiration from Wittgenstein (for example [51–54]. The approach taken by these studies is largely complementary to the perspective of cognitive science that is our focus here, which is why an explicit comparison is beyond the scope of this paper and must be left for another occasion.

The example of the wallpaper decorators (LFM, 36), who produce wallpaper that is adorned with complex proofs of deep mathematical theorems but who lack any knowledge of mathematics, provides an illustration: what are proofs to us are decorations to them; they don’t engage in mathematics: their practice has a completely different point.

Counting and measuring is relevant in a wide variety of practices. That makes them more akin to what Schatzki ([45], chapter 4) calls ‘dispersed practices’, as opposed to ‘integrative practices’. One important distinction is that the latter have an intrinsic telos. Interestingly, Schatzki argues that dispersed practices adapt to the integrative practices in which they occur. This might be part of an explanation of the pluralism that Wittgenstein endorses.

A point that is also argued by Wittgenstein, cf., PI 26 ff.

A kind of learning by trial and error is also described by Wittgenstein: When someone learns a new language in a strange country, he ‘will sometimes learn the language from ostensive definitions they give him; and he will often have to guess the meaning of these definitions; and will guess sometimes right, sometimes wrong’ (PI, 32).

For example: Young children without numerical language can make simple ordinal (less, more) numerical judgements for small numbers, but their achievements improve after they have mastered verbal counting [61].

The German term ‘Abrichtung’ is normally used for the training of animals ([29], p. 161). Wittgenstein also uses it for teaching a relationship that comes long after (if at all) learning to count, to add, and so on. Thus, the status of a + b = b + a depends on how it functions irrespective of the kind of objects (abstraction); (d) the order of counting is irrelevant ([38], p. 31). This ‘knowing’ of procedural principles may well be a rational reconstruction which has nothing to do with what is going on in the mind or brain.

‘The question (...) is why young children, who have had some understanding of quantity since they were neonates, wait until they are five to six years old to fully understand number’ ([60], p. 186).

Calculating prodigies, ‘who get the right answer but cannot say how. Are we to say that they do not calculate?’ (PI, 236). The association of knowledge and understanding with an ability extends beyond practical know-how and understanding and includes propositional knowledge. Quite generally, Wittgenstein claims, the grammar of the word ‘know’ is related to that of ‘can’ and ‘be able to’: the mastery of a technique (PI, 150).

Wittgenstein describes the teaching of abstract concepts like ‘regular’ or ‘same’ to someone who only speaks French: ‘...if a person has not yet got the concepts, I shall teach him to use the words by means of examples and by means of practice.—And when I do this I do not communicate less to him than I know myself’ (PI, 208).

Cf., e.g., OC 113, where Wittgenstein makes the point that commutativity of addition is typically something that in teaching basic arithmetic is treated as a rule, not as a (mathematical) fact. Facts are things you can claim to know, in which case you need to be prepared to back them up by giving grounds. In this case, e.g., by providing a proof in some axiomatisation of arithmetic. But that comes long after (if at all) learning to count, to add, and so on. Thus, the status of a + b = b + a depends on how it functions in a particular setting. And if the setting is that of teaching, it simply functions as a rule that guides behaviour.

‘And hence also ‘obeying a rule’ is a practice. And to think one is obeying a rule is not to obey the rule. Hence it is not possible to obey a rule ‘privately’ (PI, 202). Following a rule is part of a practice, and as such it is taught and subject to social norms. This summarizes Wittgenstein’s anti-mentalism and his enactive and socially embedded view. Important to note here is that the rule-following considerations occur not only in PI, but also in RFM.

A disclaimer: we will not even try to summarize the huge literature in this field, which ranges from philosophy to psycholinguistics, social and evolutionary psychology, game theory, sociology, economics, etc. We select a couple of in our view promising approaches to naturalize conceptual levels of human mathematical cognition.

Already suggested by Charles Darwin, who proposed that the swim bladder of fishes developed into an organ of respiration in land animals.

Inspired by the work by Overmann and Malafouris. We use Zahidi’s summary.
In the case of functions which take real numbers as values, ‘computable’ means that the number can be approximated in a computable manner to any desired degree of accuracy.

In LFM he also rejects other positions which were mainstream in the 20th century philosophy of mathematics, such as Frege-Russellian logistic foundationalism (LFM 260-66), formalism (112, 142-3), intuitionism (237) and finitism (141). His views of the latter three ‘isms’ are more nuanced and complicated than we can summarise here [72].

The superficial similarity of mathematics to games might lead to formalist conclusions, but this would be misleading and even ‘very dangerous’ (LFM, p. 142–143).

As he declares for instance in PI 144.

Cantor, who took the existence of a bijection as definitional for equicardinality and then applied his diagonal argument to construct a hierarchy of infinities, wrote scathing criticisms of DuBois Reymond’s work which took part–whole relationships as the starting point of his theory of orders of infinity.

Biological phylogenesis and ontogenesis differ radically in their causal mechanism: Phylogenesis is driven by the mechanism of natural selection, ontogenesis (at least prenatally) by the genetically programmed growth and differentiation of the embryo. He constructs a hierarchy of infinities, wrote scathing criticisms of Du Bois’s work which took part–whole relationships as the starting point of his theory of orders of infinity.

In LFM he also rejects other positions which were mainstream in the 20th century philosophy of mathematics, such as Frege-Russellian logistic foundationalism (LFM 260-66), formalism (112, 142-3), intuitionism (237) and finitism (141). His views of the latter three ‘isms’ are more nuanced and complicated than we can summarise here [72].

The superficial similarity of mathematics to games might lead to formalist conclusions, but this would be misleading and even ‘very dangerous’ (LFM, p. 142–143).

As he declares for instance in PI 144.

Cantor, who took the existence of a bijection as definitional for equicardinality and then applied his diagonal argument to construct a hierarchy of infinities, wrote scathing criticisms of Du Bois’s work which took part–whole relationships as the starting point of his theory of orders of infinity.

References


