Open access • Journal Article • DOI:10.1007/S11229-010-9837-9

Mathematical symbols as epistemic actions — Source link

Helen De Cruz, Johan De Smedt

Institutions: Katholieke Universiteit Leuven, Ghent University

Published on: 01 Jan 2013 - Synthese (Springer Netherlands)

Topics: Language of mathematics, Mathematical psychology, Word problem (mathematics education), The Extended Mind and Numerical cognition

Mathematical symbols as epistemic actions

Helen De Cruz
Centre for Logic and Analytical Philosophy, Katholieke Universiteit Leuven
Kardinaal Mercierplein 2, 3000 Leuven, Belgium, Helen.DeCruz@hiw.kuleuven.be

Johan De Smedt
Department of Philosophy and Ethics, Ghent University
Blandijnberg 2, 9000 Gent, Belgium, johan.desmedt@ugent.be

Abstract

Recent experimental evidence from developmental psychology and cognitive neuroscience indicates that humans are equipped with unlearned elementary mathematical skills. However, formal mathematics has properties that cannot be reduced to these elementary cognitive capacities. The question then arises how human beings cognitively deal with more advanced mathematical ideas. This paper draws on the extended mind thesis to suggest that mathematical symbols enable us to delegate some mathematical operations to the external environment. In this view, mathematical symbols are not only used to express mathematical concepts—they are constitutive of the mathematical concepts themselves. Mathematical symbols are epistemic actions, because they enable us to represent concepts that are literally unthinkable with our bare brains. Using case-studies from the history of mathematics and from educational psychology, we argue for an intimate relationship between mathematical symbols and mathematical cognition.

This paper is the draft prior to peer review. The definitive version can be found in the *Online first* edition of *Synthese* at the following url: `http://www.springerlink.com/content/36670740x1387471/`.

1 Introduction

What are mathematical symbols for? A widely held view is that they serve as external representations of abstract mathematical objects, which exist independently from their symbolic notations. Many believe, as epitomized by Barabashev (1997), that modernizing ancient mathematical texts does not alter their content, since the underlying mathematical truths to which they refer exist outside of space-time, i.e., independent of human thought and culture. According to this view, mathematical cognition only takes place after we convert mathematical symbols into an appropriate inner code. It would not matter whether equations are solved by use of the western algebraic notation or by arranging Chinese counting rods in matrices, nor would it matter which numerical notation system is used to perform arithmetical operations. This paper considers an alternative approach, where mathematical symbols are intimately linked to the concepts they represent. We will develop the argument that symbols are not merely used to express mathematical concepts, but that they are constitutive of the concepts themselves. Mathematical symbols enable us to perform mathematical operations that we would not be able to do in the mind alone, they are epistemic actions. We begin this paper by an examination of how humans acquire knowledge of mathematical objects. Next, we consider the role symbolic notation systems play in human cognition, focusing on mathematical symbols and taking negative numbers and algebra as examples. We then examine how mathematical symbols function as epistemic actions due to their semantic opacity.

2 How we acquire mathematical knowledge

According to some philosophers of mathematics (e.g., Benacerraf, 1973) an acceptable epistemology of mathematics should provide a causal account of how humans are able to acquire mathematical knowledge. At first blush it appears hard to explain how we are able to acquire knowledge of mathematical objects, which are often conceptualized as abstract entities that do not seem to fit in the material causal order. Social scientists (e.g., Ernest, 1998) have presented the elegant idea that the locus of mathematical reality can be found within human culture. Mathematical entities arise in the brains of individuals within a specific culture, but once arisen, they are no longer reducible to it. This idea goes back to Popper's distinction between world-1 (physical objects), world-2 (cognitive processes) and world-3 (concepts, images and other products of human reasoning). As mathematical entities are situated in world-3 they can have properties that may be unexpected and unintended; they transcend the cognitive properties of any individual (Popper, 1994, 25–27).

Like most factual knowledge (e.g., that the Earth is round or that some diseases are caused by viruses), knowledge of mathematical truths does not derive from direct sensory observation but from cultural transmission. The mechanisms that underlie this transmission need unpacking. We cannot simply download ideas into our brains, but have to reconstruct them in our own minds. To understand the concept DEMOCRACY, for example, one needs to reconstruct all its constitutive elements, such as competitive elections, civilian control of the military, and government by elected representatives. In this reconstructive process, humans seldom start from scratch, but draw on preexisting knowledge: if a child learns that an echidna is an ANIMAL, it can draw on earlier acquired knowledge of animals to make inductive inferences about echidnas, such as that they are self-propelled, require food, sleep, and produce offspring looking like themselves. Inductive inferences constrain and guide the acquisition of concepts which is aptly demonstrated by religious concepts. Religious entities often have counterintuitive elements, which have to be transmitted explicitly (Boyer, 2001). For example, ghosts violate our basic expectations of how physical objects behave in their ability to walk through walls, and to appear and disappear at will. But the psychology of ghosts is conform to our basic ontological expectations: they have beliefs, desires and a distinct personality. These latter features are not explicitly transmitted when we acquire the concept GHOST, but are tacitly assumed. The fact that all cultural transmission requires reconstruction implies that not all concepts are equally easy to transmit. Some cultural concepts will be relatively easy to reconstruct, because they provide a close match to our intuitive expectations (examples include most folk theories, such as folk biology and folk psychology); others will be harder to transmit because they elicit relatively few inductive inferences (they are *non-intuitive*; most scientific concepts fall into this category). Moreover, as minds are not high-fidelity copying machines, cultural transmitted concepts will tend to be distorted to fit pre-existing biases and ideas (Sperber, 1996).

To some extent, the acquisition of mathematical ideas can be characterized by this reconstructive process of culturally acquired information. For example, understanding the twin prime conjecture (i.e., there are infinitely many twins of primes such that if p is a prime, so is $p+2$) requires more than being able to derive or refute it from explicit definitions and axioms. Although it remains as yet unproven and may perhaps never be proven, we find the conjecture intelligible: by having found pairs of primes that are larger than any we have found before, we can wonder if there are more to find, and if eventually they will run out (Goodman, 1981). Interestingly, the initial belief that a proof may be correct, such as Andrew Wiles' proof of Fermat's last theorem, usually does not depend on thorough scrutiny, but on concurrence with high-level ideas long before the details are checked (Thurston, 2006). Mathematics education is organized as a stepwise progression from more elementary to complex notions, as can be witnessed in students who learn to solve equations with negative terms: they build on their already acquired ability to solve equations involving positive terms (Vlassis, 2004). By explaining the individual learning of mathematical concepts in this way, from more elementary to complex, we are faced with the problem of origin of the most elementary mathematical knowledge, which cannot be derived from earlier knowledge.

3 Elementary numerical knowledge

Empirical evidence from disparate disciplines suggests that humans are equipped with specialized cognitive capacities that facilitate their understanding of elementary mathematics. For brevity's sake we will focus on the example of number. Given that numbers seem far removed from elementary sense data, children from most cultures learn to count and reckon in a remarkably short span of time. After all, two bicycles and two cows do not share obvious perceptual properties, but both share the property of 'twoness'. Developmental and comparative psychologists have argued that our ability to represent natural numbers has precursors in perceptual capacities of preverbal infants and nonhuman animals. A growing body of experimental literature indicates that infants can estimate cardinality, predict the outcome of simple arithmetic operations, and compare the magnitudes of different collections of items (e.g., Feigenson, Dehaene, & Spelke, 2004) . The study of numerical cognition in animals also provides a wealth of evidence of successes in estimating cardinalities, comparing numbers of different magnitudes, and predicting the outcomes of elementary arithmetic operations. Such capacities have been attested in all vertebrate species examined for it, and recently even in invertebrates such as bees (Dacke & Srinivasan, 2008). Added to this is neuroscientific evidence, which shows that some areas of the human brain are consistently involved in arithmetical tasks (e.g., Cantlon, Brannon, Carter, & Pelphrey, 2006), strengthening the case for an evolved numerical competence.

It is quite likely that children draw on this early-developed knowledge when they learn to use natural numbers. For instance, the idea that numbers are abstract, amodal entities predates explicit instruction. Seven-month-olds can compare numerosities across the visual and auditory modalities: they can match the number of voices they hear to the number of speaking people they see (Jordan & Brannon, 2006). Five-month-olds successfully detect correspondences between the number of objects placed in their hands (tactile presentation) and the number of objects viewed on a screen (Féron, Gentaz, & Streri, 2006).

But as Rips, Bloomfield, and Asmuth (2008) have demonstrated, there is still a gap between natural numbers and the intuitive representations of number in prelinguistic infants and animals. Intuitive numerosities are approximate, becoming less precise as magnitude increases: six-month-olds can discriminate between 2 and 3 (two small numerosities) and between 8 and 16 (two larger numerosities that lie far apart), but not between 8 and 12 (Xu & Spelke, 2000). In a classic study in which rats were trained to press a lever n times (with n ranging from 4 to 36) to obtain a food reward, Meck and Church (1983) found that the animals became less and less accurate as numerosities increased. Intuitive number representations are subject to size effects (less accuracy with increasing size) and distance effects (confusion between numerosities that are close together). Several studies have shown that young children (Siegler & Booth, 2004) and nonhuman animals (Nieder & Miller, 2003) represent numbers on a logarithmic, rather than a linear mental number line. By contrast, natural numbers fit best onto a linear number line. In brief, a logarithmic mental number line is one where estimations of numerosities conform to the natural logarithms (Ln) of these numbers. This typically leads to an overestimation of the distance between small numbers, such as 1 and 2, where the psychological distance is typically judged to be much larger than between larger numbers like 11 and 12. Young children make characteristic errors when plotting numbers on a scale. Siegler and Booth (2004), for example, gave five- to seven-year-olds a number line with 0 at the left side and 100 at the right. The younger children typically placed small numbers too far to the right, e.g., 10 was placed exactly in the middle of the line. It remains as yet unclear how children can make the transition from these approximate logarithmic magnitudes to a linear natural number concept.

Learners run into additional difficulties when they acquire the concept ZERO, as they need to correctly acquire a host of properties, such as that zero is the neutral element in addition ($a + 0 = a$), or that for any a elevated to zero potency $a^0 = 1$, or

$0! = 1! = 1$. As a consequence, it might be difficult for a learner to correctly infer all of zero's properties. Several lines of experimental evidence indeed indicate that zero is more difficult to represent than the other natural numbers. For example, when three-year-olds are asked whether 4 or 0 is larger, they are as likely to answer that four or that zero is the largest, but they can effortlessly decide that 4 is larger than 2. During the early preschool years, children do not have the idea that 0 is a natural number, but rather treat it as a synonym for nothing (Wellman & Miller, 1986). The one chimpanzee tested for it likewise experienced more difficulties in representing the concept ZERO compared to other natural numbers. In a long-term study, running over 10 years, where the chimpanzee Ai was trained to use natural numbers, she kept on confusing 0 with 1 and 2; by contrast, she was highly accurate for other natural numbers 1 to 9. The fact that she never reached complete accuracy in the former task suggests that she did not possess a true concept ZERO, but that to her the symbol 0 meant something like 'very few' (Biro & Matsuzawa, 2001). If cognitive reconstruction were the only force guiding the cultural evolution of mathematical concepts, the concept ZERO would be quickly swamped by inference biases for positive quantities, and we would end up confusing zero with the smallest positive integers, not unlike the chimpanzee Ai. But this is not what happens. Although it is more difficult to acquire zero compared to other natural numbers, by the end of their preschool years, children have typically acquired an understanding of zero and correctly identify it as the smallest natural number. More impressively, Wellman and Miller (1986) found that six-year-olds are actually better at solving abstract mathematical operations with zero such as $a + 0 = a$ compared to other natural numbers. Interestingly, this psychological difference in conceptions of zero versus other natural numbers is echoed in the ontological status of the natural number zero. Only in the late 16th century did authors like Stevin begin to accept zero, rather than one, as 'le vrai et naturel commencement', the first natural number (Naets, in press). Interestingly, Stevin's conception of number was based on numerical operations rather than on general notions about numerosity. If natural numbers, including zero, cannot be directly constructed from elementary mathematical skills, it is even harder to explain how we acquire non-intuitive concepts (i.e., those that elicit very few tacit assumptions to a neophyte) like functions, vector spaces or infinitesimals.

4 Symbols and mathematical cognition

4.1 The role of external media in mathematical cognition

Some scholars (e.g., Clark, 2006; De Cruz, 2008) have explored the possibility that external media play an important part in mathematical cognition. Humans regularly supplement their internal cognitive resources with external support: hard discs and books serve as non-biological memory devices, and slide rulers and calculators make difficult tasks more tractable. This kind of action, where part of the cognitive load is delegated to the environment is termed an epistemic action (Kirsh & Maglio, 1994). Epistemic actions differ from pragmatic actions in that their primary goal is to obtain information about the world, whereas the latter are performed to bring about changes in the world.

The claim that external media enhance cognitive processing in this way is hardly controversial. Because part of the cognitive task is offloaded into the environment, performance often improves—it is obviously more difficult to count coins piled into a heap when one is not allowed to sort them into stacks (Kirsh, 1996). To be of philosophical interest, the extended mind thesis should be stronger than this and yet should not make claims that over-stretch the concept of cognition. For example, Clark and Chalmers (1998, p. 13) claim that a notebook, used by an Alzheimer patient as external memory device, is on a par with the use of internal memory to retrieve facts: 'the essential causal dynamics of the two cases mirror each other precisely'. Adams and Aizawa (2001) have raised concerns about this interpretation of the extended mind thesis, as it leads to cognition oozing into the world to the extent that the term loses its meaning. It is not because we use pen and paper to solve an equation that these objects are actually involved in mathematical cognition. Compare this to lopping shears: although this instrument enables humans to accomplish something that they would not be able to do with their bare hands (cutting thick branches), this does not imply that the muscular processes within our hands and arms actually extend into the shears. Similarly, although microscopes and hadron colliders are involved in our epistemic actions, this does not imply that one should attribute cognitive agency to these objects.

A way to interpret the extended mind without contributing cognition and agency to artifacts is to argue that not all concepts are mental representations. One must then suppose that not all concepts can be entertained by human minds, due to intrinsic limitations of human cognition. One can easily bring such limitations to mind: humans have only three types of color receptor cells in their eyes, mantis shrimps have as many as twelve, whereas most birds and insects have four types, which allow them to observe ultraviolet light. From the perspective of the representational theory of mind, mental representations of colors falling outside of what humans can observe will never be instantiated in any human mind (Margolis & Laurence, 2007, p. 568). Still, although human vision does not have receptor cells for ultraviolet light, we are able to capture it with special instruments and adapted cameras, and these can give us impressions of what some species of bird would look like to their conspecifics, or what some flowers would look like to pollinating insects. Hence, due to these cameras, humans can entertain the concept ULTRAVIOLET LIGHT without actually having a mental representation of it. Likewise, if mathematical symbols can represent objects that are not representable with our internal cognitive resources alone, this would imply that not all mathematical concepts are mental representations. This view corresponds closely to Millikan's (1998, p. 59) notion of concepts as abilities: a concept is the ability to reidentify entities with fair reliability under a wide variety of conditions. Concepts are abilities that are particular to cognitive agents; they enable the agent to make meaningful inferences. Having a mathematical concept can be thought of as the ability to make meaningful inferences about a particular mathematical object, to understand the relationship between this mathematical object and others, and to successfully use it to solve mathematical problems.

Having the concept π, for instance, means that one can make meaningful inferences about it, that one understands the relationship between π and other mathematical objects (e.g., the circle, the Cauchy distribution) and that one can use π to solve various mathematical problems (e.g., in geometry and statistics).

The symbol i illustrates how mathematical symbols can represent ideas that are not intuitively accessible. This symbol was introduced to denote an operation that is cognitively impossible, namely taking the square root of -1 (i.e., $\sqrt{-1}$). Taking the square root of any negative number is cognitively intractable: the result cannot be negative, since multiplying two numbers with the same sign is always positive. Nor can it be positive, since multiplying two positives cannot yield a negative. Nevertheless, early mathematicians like Cardano and Bombelli (16th century) allowed for calculations involving square roots of negative numbers in order to solve particular equations, especially cubic equations. In one problem, Cardano attempted to find a solution to the problem of dividing 10 into two parts, the product of which is 40. His solution was ingenious: first, divide 10 into two equal parts, we have five on each part. $5 \times 5 = 25$, which is still 15 short of 40. Dividing this remainder of -15 equally between the two parts, we have $\sqrt{-15}$. Thus the solution becomes $(5+\sqrt{-15})(5-\sqrt{-15})$. He remarked that this solution 'adeo est subtile, ut sit inutile' (it is as refined as it is useless) and considered the operation to be a 'mental torture'. Nevertheless, the formal correctness of the operations led Cardano to accept square roots of negative numbers (Ekert, 2008). However, real progress with even roots of negative numbers was only booked when Euler (18th century) introduced the symbol i. Once introduced, mathematicians no longer needed to worry about square roots of negative numbers, because the symbolism effectively masks this cognitively intractable operation, for example $\sqrt{-15}$, which bothered Cardano, can be elegantly rewritten in Euler's notation as $i\sqrt{15}$, hence the equation would be written as $(5-i\sqrt{15})(5+i\sqrt{15}) = 40$. Once Euler introduced the symbol i, where $i^2 = -1$, mathematicians had a convenient shorthand to incorporate even roots of negative numbers. This allowed such numbers to be incorporated in number theory, allowing for the representation of complex numbers of the form $x + yi$, with x as the real component and yi as the imaginary one, which vastly extended both number theory and algebra. It allowed, for example, for a proof of the fundamental theorem of algebra, for which real numbers do not suffice; it also allowed for Euler's identity $e^{i\pi} + 1 = 0$. Complex numbers were only generally accepted in the mathematical community once Argand (19th century) offered a geometric interpretation of the complex numbers, using a modified Cartesian plane, with the real part of a complex number represented by a displacement along the x-axis, and the imaginary part by a displacement along the y-axis. Again, the use of an external representation facilitated the acceptance of a new mathematical entity.

Denoting a cognitively intractable operation with a symbol makes it more manipulable, which effectively enables mathematicians to overcome human cognitive limitations. It is as if by sweeping nonintuitive mathematical operations under the carpet mathematicians need no longer fret about them. Whitehead (1911) already pointed out that mathematical notation systems free up cognitive resources because they allow us to offload ideas into the environment that are difficult to represent mentally.

> By relieving the brain of all unnecessary work, a good notation sets it free to concentrate on more advanced problems, and in effect increases the mental power of the race. In mathematics, granted that we are giving any serious attention to mathematical ideas, the symbolism is invariably an immense simplification [...] [B]y the aid of symbolism, we can make transitions in reasoning almost mechanically by the eye, which otherwise would call into play the higher faculties of the brain. It is a profoundly erroneous truism [...] that we should cultivate the habit of thinking what we are doing. The precise opposite is the case. Civilization advances by extending the number of important operations which we can perform without thinking about them (Whitehead, 1911, 58–61).

4.2 Costs and benefits of symbol use

Symbols are cognitively costly. Experimental evidence (Spelke & Tsivkin, 2001) indicates that bilinguals experience difficulties transferring exact multiplication and addition facts learned in one language into their other language. Typically, speakers of two or more languages resort to just one language—the one in which they learned basic arithmetical procedures—when counting or doing arithmetic. A historical examination of numerical notation systems (Chrisomalis, 2004) indicates that the adoption of the hindu-arabic numerals has almost invariably resulted in the disappearance of the indigenous systems, suggesting it may be too cognitively demanding to keep using both notation systems. Also suggestive of the claim that using symbols places heavy demands on cognition is the repeated observation that students, even those familiar with symbolic notation systems, consistently prefer to use verbose methods or to imagine particular situations to solve questions, like 'how to obtain the number of girls in a class when the number of boys is known, and you know that boys outnumber girls by four' (Harper, 1987). Children also experience notable difficulties when shifting from one symbolic representation to the other. When asked to lay out with blocks simple problems posed in Arabic notation (e.g., $1 + 7$), 5- to 7-year-olds typically fail to transcribe such problems, in this case, the solution is to place the blocks in two groups, one of one block and one of seven blocks. Instead, they often try to lay the blocks in such a configuration as to copy the written problems (e.g., laying '$1 + 7$' out in blocks). Only with the guide of explicit instruction can children successfully use objects in a symbolic way (Uttal, Scudder, & DeLoache, 1997).

To complicate matters further, there is no prescribed methodology to weigh the relative cognitive costs and benefits of particular notation systems. Zhang and Norman's (1995) analysis of internal and external operations of numerical notation systems suggests that multiplication in purely additive number systems, such as the roman numerals, demands more internal cognitive resources compared to positional systems. In contrast, Schlimm and Neth (2008) have developed a model which shows that—at least for artificial intelligent agents—addition and multiplication with roman numerals are perfectly feasible, although they place heavier demands on working memory, especially in the case of multiplication.

As we are immersed in a world of visual markings, such as arabic digits, letters, and pictures, it is difficult to assess how these markings influence the way we think. We here take DeLoache's (2004) very broad working definition of symbols as objects that someone intends to represent something other than itself (for clarity's sake, we disregard here the possibility of self-referring symbols, although these are in principle possible). Following this definition, a painting of a house is a symbol because it represents a house and is meant to represent a house, yet it is a painting, made with brush and paint on canvas or paper. This definition does not draw fine-grained distinctions between indexes or icons, but it is useful from a cognitive point of view. How do we come to understand that the symbol 2 stands for the object natural number two? Understanding symbols requires us to decouple meaning from materiality. This decoupling of the material nature of a symbol and its referent emerges early in development, but it is not self-evident. Prior to 18 months of age, young children have difficulties discriminating representations from real objects: controlled experimental studies have shown that infants attempt to suck pictures of bottles or put on photographs of shoes. By the second year of life, however, children can interpret pictures correctly and pay more attention to their meaning than to their shape (DeLoache, 2004). In one experiment (Preissler & Bloom, 2007), two-year-olds were shown a picture of an unfamiliar looking artifact which was called a 'wug'. When asked to give the experimenter a wug, the children gave the experimenter an object that resembled the depicted object. However, when asked 'look at the picture, can you give me another one?' the children gave the experimenter another picture with a dissimilar looking object. Such studies indicate that by the age of two, children can flexibly switch between the material nature of a symbol and its referent.

Once symbolic systems are acquired, they have a profound influence on human cognition. Neuropsychological studies indicate that the human brain, both in structure and function, is susceptible to both the types of symbols we learn and the way in which they are acquired. Reading Chinese logographs, for example, recruits brain areas additional to those typically recruited in alphabetic writing systems (Tan, Feng, Fox, & Gao, 2001). This may be caused by the fact that logographs are visually more complex, and that they represent syllables rather than individual phonemes. Whereas English native speakers use mainly the perisylvian language-related brain areas when performing arithmetical tasks, Chinese-speaking subjects rely more on motor-related brain areas (Tang et al., 2006). These differences may be related to a disparity in mathematics education: whereas western children solve arithmetic problems by retrieving addition and multiplication facts from verbal memory, Chinese children make extensive use of the abacus—the motor patterns of the imagined finger movements manipulating the abacus beads show up in the neuroimaging studies. More profound differences can be found when comparing cerebral organization between literate and illiterate people: a study of Portuguese women (Petersson, Silva, Castro-Caldas, Ingvar, & Reis, 2007) found that illiterates have more white matter and are more left-lateralized compared to literates from the same socio-economic background. As the function of white matter is to connect different functional areas of the nervous system, this indicates that literate people have less connections in their brain. Perhaps this puzzling finding can be explained by the fact that literate people can afford to forget, as they can allocate their memories to scribbled notes, computers and other artifacts.

Initially, learning to use symbols is cognitively costly, but once acquired, they provide considerable computational advantages. It is therefore not surprising that symbolic notation systems are widespread across cultures. Indeed, artificial languages are nearly as universal as natural languages, and have emerged independently in many civilizations, including Mesopotamia, India and Maya (Staal, 2006). Interestingly, cultures with strong oral traditions such as Homeric Greece, Vedic India or Old Norse culture often drew on fixed formulaic expressions (strings of words) that could be flexibly combined, such as the phrases ‘swift-footed Achilles’ in Homeric prose or ‘steed of the sea’ in Norse eddas. These formulæ are repeated so often that they lose much of their meaning. As they can be combined at will, they free up the narrator’s cognitive resources (Netz, 1999). In scientific texts, phrases like ‘arguably’, ‘but see’, ‘taken together, this evidence suggests’ might serve a similar purpose. Mathematics takes a prominent position in the use of artificial language, being invented numerous times independently and predating writing by 10,000 years or more. Indeed, visual notation systems can be traced back to at least 30,000 BC in the form of notched bone and antler objects which show a purposeful differentiation between the notches. The groupings of notches suggest a mathematical function, and many of these objects are interpreted as calculators or tallies (Tratman, 1976).

5 Specific properties of mathematical symbols

To understand how mathematical symbols enhance cognition, in particular, their ability to hold and stably represent non-intuitive ideas, it is useful to take their properties under scrutiny. Most mathematical symbols can be conceived of in two ways: they refer to objects *and* to procedures (Sfard, 1991). Structurally, all mathematical symbols refer to objects, such as the symbol π that represents the irrational number 3.14159265... Operationally, they denote procedures, in Euclidean geometry the value of π is the ratio of a circle’s circumference to its diameter.

Although mathematical symbols do not refer to material objects, mathematicians often treat them as if they were real things, as is evident in standard phrases like ‘there exists a function x such that ... ’ Incidentally, the view that mathematical symbols refer to objects does not imply commitment to a Platonist view of mathematical objects as entities outside of time and space. As Krieger (1991) noted, their ontology is similar to that of other cultural institutions, such as norms, laws or recipes. To give but one example, the central limit theorem can be conceptualized as a tool that allows statisticians to make tests of significance, to predict confidence intervals and to extrapolate findings on a small sample to an entire population. This cultural artifact sanctions not only the practice of statistics, but also social laws that rely on its implication that a collection of individuals exhibits simpler and less fluctuating behavior than the individuals themselves.

Mathematical symbols also denote computational procedures; they are representations of specific mathematical operations. In the history of mathematics, many symbols started out as nothing more than abbreviations of names of mathematical operations. However, because they provide a shorthand for operations that are difficult to conceptualize, they effectively divert our attention away from them. This often turned out to be advantageous for mathematical creativity. As Sfard (1991) and Muntersbjorn (2003) have demonstrated, notational innovations that were first invented to denote new mathematical operations recurrently gave rise to new mathematical concepts. In her analysis of the development of the calculus, Muntersbjorn (2003) describes how methods to study the properties of conic sections gave rise to the description of very small parallelograms. When summed, their surface area could be directly compared to the figures in which they were inscribed. Building on the work of Wallis and others, Leibniz (17th century) developed the idea that tiny parallelograms approximating a curvilinear area may be manipulated as if they possess numerical value: for any series of abscissas $x_1, x_2, x_3, \ldots, x_n$, we can name the difference between successive individuals dx. For any series of tiny parallelograms, with lengths given by ordinates $y_1, y_2, y_3, \ldots, y_n$ and widths by dx we can name their sum $\int ydx$. The symbols d and $\int$ were thus not arbitrary labels for immutable abstract objects, but rather shorthand descriptions of procedures to be performed on other mathematical objects. The ontological status of infinitesimally small quantities (differentials, dx) was attacked by Berkeley (18th century) and others, who questioned them for their lack of mathematical rigour. With the introduction of the limit concept in mathematical functions (19th century), these discussions petered out: d/dx should be regarded as an operator on a function, while dx has no independent meaning.

5.1 The case of negative numbers

Studies in educational psychology indicate that this evolution from procedure to mathematical object also occurs in individual students as they attempt to grasp novel mathematical ideas. In her study of adolescents' understanding of negative numbers, Vlassis (2004) noticed how 14-year-olds made surprisingly many errors on equations with negative terms. Interviews with these pupils revealed that they primarily relied on procedural rules that they had learned by heart, such as 'if both terms are negative, the sum is negative'. Even high-level students experienced difficulties explaining these procedures, showing that their performance is due to a studious application of rules, rather than an intrinsic understanding of negative numerosities. The best predictor of success in individual students was their ability to use the minus sign correctly, not their conceptual understanding (Vlassis, 2008). Difficulties with negative integers remain into adulthood: adults are very fast in judging which is the largest of two positive integers, such as 4 and 9, but take longer when one or both digits have a minus sign, such as in deciding whether -4 or -9 is the largest (Fischer, 2003).

Similarly, in the history of mathematics, we can observe that negative numbers were used long before they were accepted as legitimate mathematical objects. Arabian mathematicians rejected negatives altogether. Indeed, the term algebra, derived from the seminal text by al-Khwārizmī, *Al-jabr watl muqaal-jabr* (830), means restoration, as in the sense of adding equal terms to both sides of an equation to remove negative quantities (Stedall, 2001). Although Chinese mathematicians did use negative integers to solve equations, they were reluctant to accept a negative number as a result of an equation, resulting in many problematic reformulations of what would otherwise have been straightforward solutions (Joseph, 2000). Prior to the 19th century, the majority of European mathematicians, including Vieta and Pascal, discarded negative numbers as unintelligible, even though the minus sign appeared in print as early as 1489, in a treatise on arithmetic by Widmann. In the 17th century, Masères even devoted an entire treatise on how to avoid negative numbers, especially as solutions. This rejection of negative numbers continued well into the 19th century: De Morgan (1830, 103–104) wrote that '$3-8$ is an impossibility; it requires you to take from 3 more than there is in 3, which is absurd'. Those authors prior to the late 19th century who did accept negative numbers, such as Leibniz, did so because of their usefulness in numerical operations. As we saw earlier, Stevin accepted zero for a similar reason.

The non-intuitive nature of negative integers might be explained by their lack of ecological salience. As the intuitive number representation has probably evolved to enable animals to keep track of quantities in the environment, such as in the case of foraging, there seems to be no compelling reason why natural selection should have equipped their brains to deal with negative quantities (De Cruz, 2006). Indeed, early mathematicians who accepted negative numbers as proper mathematical objects did so in spite of their non-intuitive properties. Leibniz, for example, was convinced that Arnauld's argument against negatives was valid: this 17th century French mathematician argued that, given two (positive) integers $a > b$, it is the case that $\frac{a}{b} > \frac{b}{a}$. As $1 > -1$, it should follow that $\frac{1}{-1} > \frac{-1}{1}$. Since this is not the case, negatives were deemed to be impossible. Leibniz nevertheless argued that one could calculate with them because their form was correct, just as one can calculate with imaginary numbers (Buzaglo, 2002, 10). Without the prior existence of the mathematical operation of subtraction, symbolized by the minus sign, negative numbers would probably not have come into existence. Negative numbers are a good example of how notational innovations can foster new mathematical concepts.

5.2 Algebra

The cases of calculus, i and negative numbers show that symbols may be indispensable in the growth of mathematical knowledge, as they provide a material support for non-intuitive ideas, thus making them manipulable. But it remains unclear whether symbols play a role in the acquisition of more intuitive mathematical ideas, such as algebra.

Historically, the symbolic notation of algebraic problems preceded conceptual advances in algebra. Up until the 16th century, algebraic problems were formulated as concrete problems, such as calculating interest, land surveying or the division of inheritances. In cases where nonnumerical symbols were used to represent quantities, they were restricted to the representation of unknowns, such as in the case of Diophantine (named after a third-century Hellenic mathematician) and Indian mathematics. Mathematicians like Vieta en Descartes replaced this verbose style by symbolic notation. Importantly, they introduced symbols not only to represent unknowns but any given quantity, i.e., variables. During the 17th century, these innovations allowed for the development of a formal algebraic calculus, which specified the ways in which equations should be manipulated. Only in the 19th century, formalists like De Morgan and Gregory decided that algebra should be more than a kind of arithmetic, and started to treat algebraic formulæ as things in themselves, developing the field of abstract algebra. This evolution is reflected in the use of symbols in students' solutions to mathematical problems. Whereas younger students (12–14 years) rely on a purely verbose method, and work by example, older students (13–16 years) typically employ a Diophantine method, restricting their use of symbols to represent unknowns, whereas the oldest students (16–18 years) generalize the problems in terms of both variables and unknowns (Harper, 1987). Although these parallels between historical developments and individual learning may strike one as mysterious, they probably reflect inherent properties of mathematical concepts and their relationship to human cognition, in particular the use of symbols as a way to extend human cognition. As in the case of negative numbers, students' success in algebraic problem solving is best predicted by the extent to which they manipulate symbols. To gain a better insight into how adolescents learn to solve first-degree equations, Qin et al. (2004) measured their patterns of brain activation. These indicated that the students relied on a set of distinct brain circuits, including the intraparietal sulci, which have been implicated in numerical cognition (e.g., Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003) and the posterior parietal cortex, which is normally activated in visuospatial tasks, including spatial working memory and attention orienting. As the pupils got better at solving the equations, the activation in this posterior parietal area decreased, indicating that they relied less on spatial working memory and attention. This is in line with our hypothesis that fluency with symbols frees up cognitive resources: by using external mathematical symbols we do not need to represent mathematical operations internally.

The influence of symbolic notation in the development of algebra is well illustrated by the different historical trajectories of three systems for solving equations. Chinese algebra was based on the manipulation of counting rods, which were arranged in groups of five. By laying out these rods, mathematicians of the Han dynasty (200 BC–200 AD) could elegantly visualize simultaneous linear equations: they simply arranged the rods in rows and columns, where each row corresponded to the coefficient of an unknown and each column to an equation. This in turn promoted the invention of matrix methods to solve simultaneous linear equations and higher-order equations. (This type of solution was not reached by European mathematicians before the 18th century, probably because arabic numerals are less suited for this external cognitive operation.) However, these rods were less useful to express general abstract rules other than actual calculations which preserved the concreteness of Chinese mathematics. Consequently, Chinese algebra textbooks never attempted to give an abstract formulation of a general rule, but presented examples that served as paradigms to solve similar problems (Chemla, 2003). In contrast, once European mathematicians had invented symbols to represent variables and unknowns from the 16th century onwards, they could provide general solutions to problems. To give but one example, it is quite straightforward to prove in European algebra that the sum of two odd integers is an even integer. This stands in stark contrast with Arabic algebra, which was a sophisticated verbose method for solving equations. Each power of an unknown has its own term, e.g., *shay* (thing) means x, *māl* (sum of money) means x^2, and so on. In contrast to western algebra, there are no operations with polynomials. Thus, the expression 'four things' is not equivalent to $4x$, i.e., things multiplied by four. Instead, it merely indicates how many things are present, as in 'four bottles'. Even seemingly mundane operations like addition and subtraction are not represented consistently as at least eight synonyms denote subtraction. Remarkably, in his analysis of Arabic algebra, Oaks (2007) found only one instance where *shay* actually meant an arbitrary quantity. The verbose character of Arabian mathematics, which was intimately tied to mathematical practice, did not allow for the development of formal algebraic proofs or the formulation of general solutions. Rather, memory and recitation by heart were central, as in the performances of prodigies in mental calculation that haunted our television shows until recently. As a result, Arabian algebra did not experience much conceptual progress between the 9th and 16th centuries. For example, Omar Khayyám had to resort to a geometric solution in order to solve cubic equations (Joseph, 2000). This all too brief discussion of algebra in imperial China, early modern Europe and the medieval Islamic world indicates that the way algebraic ideas were expressed symbolically was more than a mere format to represent internally held mathematical thoughts—rather, as we can glean from their influence on the divergent history of mathematics in those cultures, these notation systems did exert a profound influence on the subsequent development of mathematical ideas.

6 The opacity of mathematical symbols

Many mathematical symbols denote operations that are difficult to perform; they provide material anchors for thoughts that are difficult to understand or represent. Without external symbols, such thoughts would not survive long in the competition for attention and cognitive resources that characterizes the cultural transmission of ideas. Concepts that are hard to learn and hard to represent are typically distorted by inferential processes, such as intuitive number representation. As we have seen in the case of zero, i, and negative numbers, human intuitive expectations of how mathematical objects should behave, can sometimes be a hindrance to the progress of mathematical theory. But once these mathematical objects are represented externally, and not just stored as mental representations, their cultural transmission is less vulnerable to distortion. The symbols protect mathematical ideas from being distorted or confused by inference biases, and provide a point of focus to manipulate these objects with higher fidelity.

Since mathematical symbols can express operations that cannot be performed by the naked mind (i.e., the mind without the help of external representations), they can convey a range of ideas that are semantically opaque. A transparent symbol is one to which we have semantic access, we intuitively grasp its meaning. Hindu-arabic numerals which denote positive integers $1, 2, 3 \ldots$ are examples of transparent symbols. Studies of brain activation in adults and five-year-olds show that a number comparison task with these numerals activates the same brain areas as one that involves arrays of dots; even the speed of computation is identical (Temple & Posner, 1998).

Apparently the brain immediately translates a positive integer into a mental representation of its quantity. By contrast, symbols that represent non-intuitive concepts remain partially semantically inaccessible to us, we do not reconstruct them, but use them as they stand. Indeed, when using such mathematical symbols in everyday discourse, we do not elaborate on them—many users would be hard put when pressed. But it is precisely due to this semantic inaccessibility, that mathematical symbols remain underdetermined, allowing further creative processes as symbols can be used in novel contexts for a variety of epistemic purposes. Their generality enables us to use identical functions and equations in disparate contexts, like the Price equation that is used in disciplines as diverse as economics, biology, and anthropology. This is not only because they provide a formal, precise way to describe processes and observations, but also because their semantic opacity allows us to reason more correctly, without letting intuitive biases get in the way. Indeed, several experimental studies have shown that children's performance on non-intuitive tasks improves dramatically when they learn how to solve the problems in a mathematical way. For example, 9 to 11-year-olds can solve fractions and balance problems better when questions are presented in a discrete, mathematical rather than in an uncalibrated format. In these studies, the children were presented with a non-intuitive setup, a full glass close to the balance point, and a half-empty glass further away from it, and had to predict the behavior of the scale. In such cases, the subjects were exclusively guided by weight bias (i.e., the balance should tip towards the fullest glass). However, when presented in a discrete format (with different pegs placed in discrete positions along the lever of the scale), more children correctly pitted mass and distance against each other, and realized that in some cases an object of lower mass can tip the scale in its favor (Schwartz, Martin, & Pfaffman, 2005).

In conclusion, mathematical symbols are more than external representations of mathematical concepts. They enable us to express mathematical operations that are unthinkable with our bare brains. Using mathematical symbols can be seen as epistemic actions, not unlike the use of other external tools in scientific practice, such as microscopes, particle accelerators and slide rulers. As we have aimed to demonstrate with the examples of negative numbers and algebra, denoting mathematical operations by symbols enables us to treat such operations as if they were real entities. Conceptual progress critically depends on the ability to use mathematical symbols as expressions of operations, a process that can be observed in the history of mathematics and in the minds of students.

References

Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, *14*, 43–64.

Barabashev, A. G. (1997). In support of significant modernization of original mathematical texts (in defense of presentism). *Philosophia Mathematica*, *5*, 21–41.

Benacerraf, P. (1973). Mathematical truth. *Journal of Philosophy*, *70*, 661–680.

Biro, D., & Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (*Pan troglodytes*): Cardinals, ordinals, and the introduction of zero. *Animal Cognition*, *4*, 193–199.

Boyer, P. (2001). *Religion explained. The evolutionary origins of religious thought.* New York: Basic Books.

Buzaglo, M. (2002). *The logic of concept expansion.* Cambridge: Cambridge University Press.

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology*, *4*, e125.

Chemla, K. (2003). Generality above abstraction: The general expressed in terms of the paradigmatic in mathematics in ancient China. *Science in Context*, *16*, 413–458.

Chrisomalis, S. (2004). A cognitive typology for numerical notation. *Cambridge Archaeological Journal*, *14*, 37–52.

Clark, A. (2006). Material symbols. *Philosophical Psychology*, *19*, 291–307.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*, 7–19.

Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Animal Cognition*, *11*, 683–689.

De Cruz, H. (2006). Why are some numerical concepts more successful than others? An evolutionary perspective on the history of number concepts. *Evolution and Human Behavior*, *27*, 306–323.

De Cruz, H. (2008). An extended mind perspective on natural number representation. *Philosophical Psychology*, *21*, 475–490.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in cognitive sciences*, *8*, 66–70.

De Morgan, A. (1830). *On the study and difficulties of mathematics.* London: Paul Kegan.

Eger, E., Sterzer, P., Russ, M. O., Giraud, A.-L., & Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, *37*, 1–20.

Ekert, A. (2008). Complex and unpredictable Cardano. *International Journal of Theoretical Physics*, *47*, 2101–2119.

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics.* Albany: State University of New York Press.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Féron, J., Gentaz, E., & Streri, A. (2006). Evidence of amodal representation of small numbers across visuo-tactile modalities in 5-month-old infants. *Cognitive Development*, *21*, 81–92.

Fischer, M. H. (2003). Cognitive representation of negative numbers. *Psychological Science*, *14*, 278–282.

Goodman, N. D. (1981). The experiential foundations of mathematical knowledge. *History and Philosophy of Logic*, *2*, 55–65.
Harper, E. (1987). Ghosts of Diophantus. *Educational Studies in Mathematics*, *18*, 75–90.
Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences of the USA*, *103*, 3486–3489.
Joseph, G. G. (2000). *The crest of the peacock: Non-European roots of mathematics* (2nd ed.). Princeton: Princeton University Press.
Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior*, *4*, 415–452.
Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.
Krieger, M. H. (1991). Theorems as meaningful cultural artifacts: Making the world additive. *Synthese*, *144*, 135–154.
Margolis, E., & Laurence, S. (2007). The ontology of concepts—Abstract objects or mental representations? *Noûs*, *41*, 561–593.
Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 320–334.
Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs, and real kinds: More mama, more milk, and more mouse. *Behavioral and Brain Sciences*, *21*, 55–65.
Muntersbjorn, M. M. (2003). Representational innovation and mathematical ontology. *Synthese*, *134*, 159–180.
Naets, J. (in press). How to define a number? A general epistemological account of Simon Stevin's art of defining. *Topoi*, *xx*, xx–xx.
Netz, R. (1999). Linguistic formulae as cognitive tools. *Pragmatics and Cognition*, *7*, 147–176.
Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, *37*, 149–157.
Oaks, J. A. (2007). Medieval Arabic algebra as an artificial language. *Journal of Indian Philosophy*, *35*, 543–575.
Petersson, K. M., Silva, C., Castro-Caldas, A., Ingvar, M., & Reis, A. (2007). Literacy: A cultural influence on functional left-right differences in the inferior parietal cortex. *European Journal of Neuroscience*, *26*, 791–799.
Popper, K. (1994). *In search of a better world. Lectures and essays from thirty years*. London: Routledge.
Preissler, M., & Bloom, P. (2007). Two-year-olds appreciate the dual nature of pictures. *Psychological Science*, *18*, 1–2.
Qin, Y., Carter, C. S., Silk, E. M., Stenger, V. A., Fissell, K., Goode, A., & Anderson, J. R. (2004). The change of the brain activation patterns

as children learn algebra equation solving. *Proceedings of the National Academy of Sciences of the USA*, *101*, 5686–5691.

Rips, L., Bloomfield, A., & Asmuth, J. (2008). From numerical concepts to concepts of number. *Behavioral and Brain Sciences*, *31*, 623–642.

Schlimm, D., & Neth, H. (2008). Modeling ancient and modern arithmetic practices: Addition and multiplication with Arabic and Roman numerals. In V. Sloutsky, B. Love, & K. McRae (Eds.), *Proceedings of the 30th annual meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Schwartz, D. L., Martin, T., & Pfaffman, J. (2005). How mathematics propels the development of physical knowledge. *Journal of Cognition and Development*, *6*, 65–88.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*, 1–36.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428–444.

Spelke, E. S., & Tsivkin, A. (2001). Language and number: A bilingual training study. *Cognition*, *78*, 45–88.

Sperber, D. (1996). *Explaining culture. A naturalistic approach.* Oxford: Blackwell.

Staal, F. (2006). Artificial languages across sciences and civilizations. *Journal of Indian Philosophy*, *34*, 89–141.

Stedall, J. A. (2001). Of our own nation: John Wallis's account of mathematical learning in medieval England. *Historia Mathematica*, *2*, 73–122.

Tan, L. H., Feng, C. M., Fox, P. T., & Gao, J. H. (2001). An fMRI study with written Chinese. *NeuroReport*, *12*, 83–88.

Tang, Y., Zhang, W., Chen, K., Feng, S., Ji, Y., Shen, J., Reiman, E., & Liu, Y. (2006). Arithmetic processing in the brain shaped by cultures. *Proceedings of the National Academy of Sciences of the USA*, *103*, 10775–10780.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences of the USA*, *95*, 7836–7841.

Thurston, W. (2006). On proof and progress in mathematics. In R. Hersh (Ed.), *18 unconventional essays on the nature of mathematics* (pp. 37–55). New York: Springer.

Tratman, E. K. (1976). A late Upper Palaeolithic calculator (?), Gough's cave, Cheddar, Somerset. *Proceedings of the University of Bristol Spelæological Society*, *14*, 123–129.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology*, *18*, 37–54.

Vlassis, J. (2004). Making sense of the minus sign or becoming flexible in 'negativity'. *Learning and Instruction*, *14*, 469–484.

Vlassis, J. (2008). The role of mathematical symbols in the development of number conceptualization: The case of the minus sign. *Philosophical Psychology*, *21*, 555–570.

Wellman, H. M., & Miller, K. F. (1986). Thinking about nothing: Development of concepts of zero. *British Journal of Developmental Psychology*, *4*, 31–42.

Whitehead, A. N. (1911). *An introduction to mathematics.* London: Williams & Northgate.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

Zhang, J., & Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition*, *57*, 271–295.