Mathematics in Mind

Marcel Danesi *Editor*

Interdisciplinary Perspectives on Math Cognition

Springer

Mathematics in Mind

More information about this series at http://www.springer.com/series/15543

Marcel Danesi
Editor

Interdisciplinary Perspectives on Math Cognition

Springer

Editor
Marcel Danesi
University of Toronto
Toronto, ON, Canada

Cognitive Science Network
Fields Institute for Research in Mathematical Sciences
Toronto, ON, Canada

ISSN 2522-5405 ISSN 2522-5413 (electronic)
Mathematics in Mind
ISBN 978-3-030-22536-0 ISBN 978-3-030-22537-7 (eBook)
https://doi.org/10.1007/978-3-030-22537-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The study of mathematical cognition has become an ever-broadening interdisciplinary field of inquiry that aims to understand the neural basis of mathematics and, more specifically, how mathematical concepts emerge. Starting with the work of Brian Butterworth, Stanislas Dehaene, Keith Devlin, Lakoff, and Núñez, among others, the field started burgeoning in the early 2000s, having provided today a huge database of research findings, theories related to math learning, and insights into how mathematics intersects with other neural faculties, such as language and drawing. The field has not just produced significant findings about how math is processed in the brain but also reopened long-standing philosophical debates about the nature of mathematics.

The number of journals, book series, and monographs that is now devoted to the study of math cognition is enormous. The purpose of this volume aims not to add merely to the accumulation of studies but to show that math cognition is best approached from various disciplinary angles. The goal is to broaden the general understanding of mathematical cognition through the different theoretical threads that can be woven into an overall understanding. The groundwork for establishing an interdisciplinary approach was laid, in recent times, by George Lakoff and Rafael Núñez in their book *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*, in which they discussed a coherent, albeit controversial, view of how mathematicians come to use and invent their proofs and theorems through the use of analogies and metaphor. The idea that proofs and mathematical concepts stem from such "rhetorical thinking" certainly resonates with semioticians, linguists, literary critics, and other humanist and social scientists. In this framework, mathematics can be seen to be an offshoot of the same neural-conceptual system that leads to the processing of language and other human skills and faculties. Whether or not this tenet can be proven empirically, the point is that it is plausible and highly interesting and, thus, needs to be explored seriously if we are ever to come to an understanding of what mathematics is and why it leads to knowledge of the world.

The overall perspective that this volume aims to adopt can be called *hermeneutic*. Like the philosophy of mathematics, the hermeneutic approach must consider the ontological source of mathematics; like psychology, it should try to understand the nature of mathematics as a product of brain and culture interacting in specific ways; like semiotics, it must connect mathematics to signs and symbols and should look at the relationship of mathematics to other human faculties and how it connects to the outside world.

This anthology is hermeneutical in the interdisciplinary way that it explores how the math mind manifests itself in various ways and what implications for math learning and teaching it has. The theme woven throughout is that math cognition is interconnected with other processes, such as spatial reasoning and metaphor, which lead us to contemplate deeper structures hidden or implicit in mathematical creations. This book is part of a series of projects undertaken at the Fields Institute for Research in Mathematical Sciences under the aegis of its *Cognitive Science Network: Empirical Study of Mathematics and How It Is Learned*. The present series is published under this aegis as well. As a Co-director of the network, I would like to express my sincere gratitude to Dr. Edward Bierstone, the Director of the Fields, for allowing me to explore mathematics from an interdisciplinary perspective with the collaboration of colleagues from mathematics to neuroscience.

Overall, mathematicians, cognitive scientists, educators of mathematics, philosophers of mathematics, semioticians, psychologists, linguists, anthropologists, and all the other kinds of scholars who are interested in the nature, origin, and development of mathematical cognition will hopefully find something of interest in this volume. The implicit claim in all the studies is that in order to penetrate the phenomenon of mathematics, it is necessary to utilize methods and theoretical frameworks derived from a variety of disciplines.

Toronto, ON, Canada Marcel Danesi

Contents

Chapter 1
From Biological Brain to Mathematical Mind: The Long-Term Evolution of Mathematical Thinking

David Tall

Introduction

In this chapter we consider how research into the operation of the brain can give practical advice to teachers and learners to assist them in their long-term development of mathematical thinking. At one level, there is extensive research in neurophysiology that gives some insights into the structure and operation of the brain; for example, magnetic resonance imagery (MRI) gives a three-dimensional picture of brain structure and fMRI (functional MRI) reveals changes in neural activity by measuring blood flow to reveal which parts of the brain are more active over a period of time. But this blood flow can only be measured to a resolution of 1 or 2 s and does not reveal the full subtlety of the underlying electrochemical activity involved in human thinking which operates over much shorter periods.

Here we use available information about the brain to consider aspects of mathematical thinking that can be observed by teachers and learners. For example, by understanding how the brain interprets written text and hears spoken words, it becomes possible not only to reveal *why* individuals have difficulty in making sense of expressions in arithmetic and algebra but also *how* sense-making can be improved at every level from the full range of young children to the varied needs of adults. One possibility involves noticing aspects that are intuitively grasped by more successful thinkers that give them advantage and introducing these insights explicitly to improve mathematical sense-making for the broader population.

Another aspect relates to the difference between the way that the eye reads text and follows moving objects. This offers fundamental insight into human perceptions of constants and variables that are foundational in the calculus. At a higher level of abstraction there is the manner in which a written proof may be scanned by someone attempting to make sense of it. These diverse insights are used to build a coherent

D. Tall (✉)
University of Warwick, Coventry, UK
e-mail: David.Tall@warwick.ac.uk

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_1

picture of how the biological brain can develop into a mathematical mind capable of contemplating and sharing increasingly subtle mathematical theories at all levels from newborn child to adult.

Building such a theory must take into account that different communities of practice may interpret situations in ways that may be in conflict with one another so that the conclusions in one community may not be appropriate for another. This applies to many different communities engaged in mathematical activity, including mathematicians in different specialisms, teachers in various educational contexts, philosophers, psychologists, neurophysiologists, curriculum designers, politicians and so on.

In addition to considering competing theories, the proposed framework seeks a higher level multi-contextual overview that takes account of the natural ways that mathematical thinking develops over the longer term, both in the individual and also corporately in different communities. The evidence presented here suggests that accepted approaches to teaching and learning mathematics by established communities of practice may be counterproductive in supporting the long-term development of mathematical thinking.

Differing Conceptions of Mathematics and a Multi-Contextual Overview

The term "community of practice" was initially introduced by Lave and Wenger (1991) as "a group of people who share a craft or a profession". In any community of practice, individuals may have differing personal viewpoints but, overall, they agree (or believe that they agree) to certain shared principles. Communities of practice incorporate a wide range of participants, including "experts" who are well versed in the practices and "novices" who are being introduced to the practices of the community in various contexts.

In the long-term learning of mathematics, the contexts encountered will change substantially. For the purposes of this chapter the term "mathematical context" will refer to a specific mathematical topic being experienced by a particular individual or group of individuals in a specific community of practice. The topic may relate to a single example or to a longer term sequence of activities.

As mathematics becomes more sophisticated, some ideas that worked well in a previous context may continue to be *supportive* in a new context, while others may become *problematic* (Tall 2013). For instance, simple number facts such as "2 + 2 makes 4" encountered with whole numbers continue to be supportive when dealing with more general numbers, such as fractions or signed numbers. Other experiences, such as the fact that the product of two whole numbers gives a bigger result but the product of two fractions can be smaller, may be problematic for many learners.

Our main strategy is to seek fundamental principles that remain supportive through many contexts over the long term, so that they can be used as a stable basis for learning while identifying successive problematic aspects that arise as the context changes to help learners become aware of them and address changes appropriate to support long-term learning.

Problematic changes in context often occur as mathematical thinking evolves, both in history and in the individual. This can be seen in the language relating to new kinds of number—positive and *negative*, rational and *irrational*, and real and *imaginary*—which involve significant boundaries in the evolution of ideas that need to be addressed.

Crossing a boundary may be termed a "transgression" from the Latin for "going across", which carries with it a sense of moving to previously unacceptable territory (Kozielecki 1987; Pieronkiewicz 2014). It is used not only in a religious context, but also in a geographic context such as when water flows across a flood plain. It is also appropriate in a historical or personal transition across a boundary in mathematics.

The changes in context may be interpreted in different ways by different communities. If a given community A has a particular belief that is problematic for community B, and an individual or a subgroup S in community A changes to adopt the beliefs of community B, then this change will be seen by community A as a *transgression* while community B will see it as an *enlightenment* (Tall 2019).

In historical development, such transitions from transgressions to enlightenments occurred with the introduction of negative or complex numbers, or the use of infinitesimals in seventeenth-century calculus, which was criticised and later rejected by the introduction of epsilon-delta analysis at the beginning of the twentieth century, and then reintroduced, subject to great dispute, in non-standard analysis in the 1960s.

Similar conflicts occur in individual learning as mathematics shifts to new contexts, say from whole number arithmetic to fractions, to signed numbers, to finite and infinite decimal expansions, and to real and complex numbers and from various contexts in arithmetic to algebra.

It is not simply a matter of shifting from one level of insight to a higher level. Often it is important to be aware that apparently conflicting possibilities can coexist in different contexts at the same time. For instance, in whole number arithmetic there is a theory of unique factorisation into prime numbers which can be extended to fractions and signed numbers by allowing the powers to be positive or negative, and to factorise polynomials in algebra. But, for highly technical reasons, it fails for certain more general algebraic numbers that mix whole numbers and square roots (Stewart and Tall 2015).

Why should the average reader care? The answer is that average readers are unlikely to encounter this particular problem in algebraic number theory, but they *will* encounter many examples where new experiences conflict with previous experiences that makes them feel *uncomfortable*, *unwilling* and even believing that they are *incapable* of thinking about mathematics. Matters are made worse when learners are subject to the beliefs of communities of experts that are at variance with their own current level of development.

The response to these conflicts is to identify their possible sources not only in the thinking of the student or the teacher, but also in the mathematics itself as it develops in sophistication. This offers new ways of addressing the problem of making long-term sense of mathematical thinking.

We begin by considering:

- How the biological brain operates as it encounters increasingly sophisticated mathematical constructs in successive contexts over the long term.

Then we consider:

- How the brain makes sense of space and number.
- How the eyes and brain interpret written text.
- How the brain interprets spoken and aural expressions.

This information will be used to formulate a framework for the long-term meaningful interpretation of expressions in arithmetic and algebra.

More generally, we will briefly consider:

- How the eye follows a moving object, giving meaning to constants and variables.
- How the eye reads through a written proof to make it meaningful.

This will be shown to be part of an overall framework for the long-term evolution of mathematical thinking in the individual (and in corporate society) that takes account of cognitive and affective growth through increasingly sophisticated mathematical contexts.

The Biological Brain

The biological brain is far too complicated to describe in detail in a chapter such as this. It has evolved over many years where more successful variants in individuals are passed on to later generations without any overall grand design. The individual grows from a single fertilised cell and develops by successive cell subdivision guided by the genetic structure from the parents to construct an essentially symmetric brain in two halves with complex links between them.

Evolution works in unexpected ways. For example, the left side of the human brain receives signals and sends output to the right side of the body and the right side of the brain deals with the left side of the body. The two halves cooperate together: almost all right-handed individuals and most left-handers deal with sequential operations such as language, speech and calculation in the left brain while the right brain focuses on global aspects such as interpreting visual information and estimating size.

Neuroscience studies the brain in a variety of ways. These include the use of electrodes on the scalp to detect electrical activity in the cortex (the "grey cells" on the surface where sophisticated thinking takes place). Magnetic resonance image scanners (MRI) take cross-sectional scans of the brain to give a three-dimensional picture of brain structure including the internal connections. Functional MRI scanners (fMRI) trace blood flow over a period of 2 s or so, as the blood carries more oxygen to areas where the brain is more active. Both give valuable insight into brain structure and a broad view of its operation while being too coarse to trace the detail of human thinking occurring in milliseconds.

Initially fMRI studies of mathematical activity focused mainly on simple arithmetic tasks. More recent studies (e.g. Amalric and Dehaene 2016) suggest surprising possibilities in the relationship between language and mathematical thinking. They say that Chomsky (2006) declared that "the origin of the mathematical capacity [lies in] an abstraction from linguistic operations", while Einstein insisted: "Words and language, whether written or spoken, do not seem to play any part in my thought processes" (quoted in Hadamard 1945: 142–1433). Of course, different individuals may think in different ways and Einstein certainly used imaginative thought experiments in developing his theories of relativity.

However, when Almaric and Dehaene studied mathematicians working in very different research areas (abstract algebra, analysis, geometry, topology), they found that all of them activated areas of the brain related to spatial sense and number which are present in young children before they develop language and are also found in many other non-human species.

Apart from linguistic memory for arithmetic facts, these areas rarely link to areas processing language (Dehaene et al. 1999; Shum et al. 2013; Monti et al. 2012). In addition, brain imaging studies of nested arithmetic expressions reveal little or no links with language areas (Maruyama et al. 2012; Nakai and Sakai 2014).

While language may be used as scaffolding to link different aspects of mathematical thinking, deeper levels of mathematical thought link with spatial imagery and mathematical operations. In this chapter we seek to link the natural use of language to fundamental human ways of thinking flexibly about spatial imagery and number.

Brain activity, as a whole, deals not only with cognitive issues. The limbic system (Limbic System n.d.)[1] in the centre of the brain handles a complex array of tasks including laying down and retrieving long-term memories; it also reacts immediately to threats in a primitive "fight or flight" mechanism. This suffuses the brain with biochemicals (neurotransmitters) that enhance or suppress connections that can affect mathematical thinking in emotional ways. These may be positive in terms of determination and resilience or negative in terms of anxiety or avoidance.

To make sense of how the human brain builds mathematical connections, it is therefore important to complement what is known about cognitive development with affective reactions to mathematical ideas.

[1] Wikipedia: Limbic System: https://en.wikipedia.org/wiki/Limbic_system

How the Brain Makes Sense of Spatial Information and Number

In the early years a young child develops the capabilities to recognise a given object seen from different viewpoints and in different orientations as being consistently the same. In mathematics, over a period of years, the child builds what Piaget (1952) called “conservation of number”. This means that a given collection of objects has a consistent number attached to it and that if the objects in the collection are rearranged spatially or if they are counted in a different way, then the number of objects remains the same.

Mathematicians formulated the properties of number and arithmetic using rules such as the “commutative”, “associative” and “distributive” laws for addition and multiplication. This was taken as a foundation of the “New Math” of the 1960s, but failed to take account of the reality of the development of mathematical thinking in the learner. On the other hand, mathematics educators studied the difficulties encountered by learners and formulated more child-centred approaches including the elaboration of different methods of counting and whole number arithmetic, such as count all, count on, known facts and derived facts. International comparisons such as TIMSS (2015) and PISA (2015) brought politicians into the act as they sought to improve international competitiveness. Multiple communities of practice sought to influence the curriculum in very different ways that could be in conflict.

In this chapter we will not enter into a comparison between the practices of different communities. Instead we focus on the increasing sophistication of mathematical structures and operations and how they develop from fundamental human ideas of time, spatial sense and number.

The concept of number does not start with rules of arithmetic. Instead it builds from a sense that when a collection is reorganised in space or counted in different ways, then some things remain the same. The most important of these, which is not immediately obvious to the child, is that the number of objects in a collection remains the same, no matter how it is rearranged or how it is counted. Figure 1.1 (taken from Tall 2019) shows how a collection of six objects has the same number of objects no matter how it is arranged or counted. The number 6 is chosen because it is the smallest number that allows not only different methods of counting and addition, but also two different methods of multiplication.

Young children will have many life experiences that contribute to the development of mathematical thinking, including shared singing and dancing with rhythmic

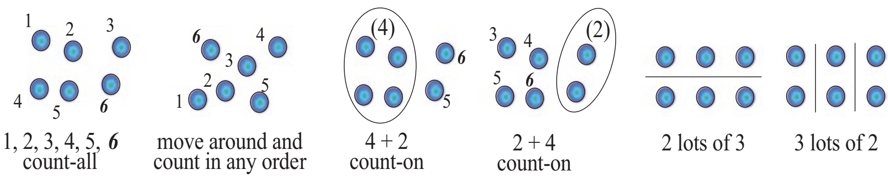


Fig. 1.1 Conservation of the number 6

representations of the number sequence: "One, two, three, four five, once I caught a fish alive; six, seven, eight, nine, ten, then I let him go again".

As children mature, they will have many experiences, playing games, practising arithmetic techniques and exploring patterns. Opportunities arise to focus on an increasing awareness of the conservation of number. For example, in counting a collection of objects in different ways, the total number remains the same; in adding together two or more collections, the total number of items is the same, no matter how it is calculated:

$$3+4+6+15 \text{ gives the same result as } 4+15+6+3.$$

This is a fundamentally important principle over the long term. It applies not only to whole numbers, but also to fractions, signed numbers, decimal notation, infinite decimal expansions, real numbers and even complex numbers. For instance:

$$7+\tfrac{3}{4}+1.414+(-5) \text{ gives the same result as } 1.414+7+(-5)+\tfrac{3}{4}.$$

This leads to a major underlying principle that is supportive throughout the number systems encountered in school mathematics:

The General Principle of Addition for Numbers: A finite sequence of additions of numbers is independent of the order of calculation

For individuals who attain more sophisticated levels of mathematical thinking, this can lead to a further generalisation in algebra and calculus:

The General Principle of Addition: The sum of a finite collection of constant or variable quantities is independent of the order of calculation

There are corresponding principles for multiplication, such as

The General Principle of Multiplication: The product of a finite collection of constant or variable quantities is independent of the order of calculation.

The multiplication principle works for most situations in school mathematics, though it fails in more sophisticated contexts such as matrix multiplication. Both principles can be extended to other operations, such as subtraction, division and powers and this will be addressed later.

To support meaningful learning of mathematics over the long term, the aim is for teachers and learners to become aware of properties that remain supportive through several changes in context to give a stable foundation for new learning. The plan is to use the sense of security in such general principles to encourage learners to address situations where the context changes and previously supportive ideas become problematic, to seek meaningful reasons why changes in meaning need to be incorporated into long-term thinking. This can be assisted by reflecting on how we humans make sense of our perceptions and actions.

How the Eyes Read Text and Symbolic Expressions

When we read text on a page, we do not scan the lines smoothly. Instead the retina in the eye has a small area called the macula which registers much higher detail and takes in successive parts of the text in a succession of jumps (called “saccades”) that the brain puts together to build up the meaning of the text. Figure 1.2 (taken from Tall 2019) shows printed text on the left and a representation of how human vision focuses on a small part of the text on the right. Read the clear text on the left several times to sense how your eye jumps along the lines to make build the meaning of the text. *Do this now ….*

It transpires that when we speak words, we do so as a sequence in time, and when we read text, we do so in an ordered sequence in a direction dependent on the language concerned—usually left to right in Western languages (Tall 2019). This sets an implicit mode of thinking that can become problematic when interpreting expressions in arithmetic and algebra.

How the Brain Interprets Spoken and Aural Symbolic Expressions

When a mathematical expression such as 1 + 2 × 3 is spoken or heard, it occurs *in time*, as “one plus two times three”. The traditional sequence of spoken and written language suggests that the operations should follow the sequence in time: first carry out the operation “1 + 2” which is “3”, and then “3 × 3” which gives **9**.

Children are given a different convention in mathematics that contradicts this natural sequence with the rule “multiplication takes precedence over addition”. This requires first performing the second operation “3 × 3” to get “6” and then calculating “1 + 6” to get the “*correct*” answer, **7**.

If children learn to follow the rule without reason, as expressions become more sophisticated and the rules more complicated, then, over the longer term, arithmetic, and subsequently algebra, becomes increasingly difficult. Simply learning by rote fails to take into account the subtle ways in which language is expressed. It is not just a matter of *what* we say, but *how* we say it.

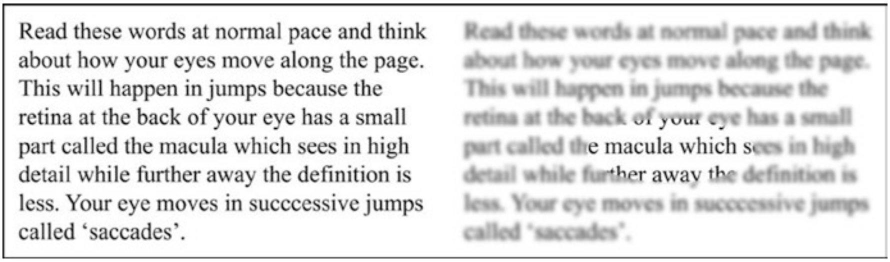


Fig. 1.2 Reading text

In written text we use punctuation to distinguish subtle differences in meaning. In spoken text, we can use tone of voice and articulation to express meaning. If we say mathematical expressions, by leaving gaps in different places, we can emphasise which operations are linked together. For example, saying

$$\text{``one plus}\,[\text{gap}]\,\text{two times three''}$$

emphasises that the words in the phrase "two times three" are to be taken together, suggesting that the result is "one" plus "two times three", which is "one" plus "six", which is "seven".

In Tall (2019) I played with the idea of writing an ellipsis (…) to denote a gap, so that

$$1 + 2 \ldots \times 3 \text{ means } 3 \times 3, \text{which is } 9,$$

while

$$1 + \ldots 2 \times 3 \text{ means } 1 + 6, \text{which is } 7.$$

At this point it is helpful to speak the two expressions "1 + 2 … × 3" and "1 + … 2 × 3" out loud to yourself and, if possible, to another person, to see how these two ways of speaking give two clearly different meanings, not only to the person speaking, but also in communication with others.

Do this now before proceeding. It is essential that you experience this for yourself.

This reveals that the meaning of a sequence of operations in arithmetic, and also later in algebra, depends on the way it is spoken. It can be formulated as follows:

> **The Articulation Principle**: The meaning of a sequence of operations can be expressed by the manner in which the sequence is *articulated* (Tall 2019).

It is essential to realise that this principle does not act like a definition in mathematics that can be used to make a formal deduction or to prove a theorem. It makes us aware that we need to think very carefully how we interpret and communicate mathematical expressions.

The principle operates with other expressions. For instance, it clarifies the possible meanings of an expression such as 2 × 3 + 4 which could be interpreted as

$$2 \times 3 \ldots + 4 \text{ which gives } 6 + 4, \text{which is } 10$$

or

$$2 \ldots \times 3 + 4 \text{ which gives } 2 \times 7, \text{which is } 14.$$

This makes it imperative to introduce suitable conventions to clarify the precise meaning, such as introducing brackets around operations that should be performed first. Thus 2 × 3… + 4 can be written as (2 × 3) + 4, which is 10 and

$$2 \ldots \times 3 + 4 \text{ can be written as } 2 \times (3 + 4), \text{which is } 14.$$

It is then possible to introduce further conventions to reduce the length of expressions. For instance, the convention "multiplication takes precedence over addition" allows us to remove brackets around a product, in the knowledge that the convention requires multiplication to be calculated before addition to rewrite

$$(2 \times 3) + 4 \text{ as } 2 \times 3 + 4$$

while retaining the notation for 2 × (3 + 4).

The principle of articulation is widely applicable throughout mathematics. For instance, my 11-year-old grandson surprised me one day when he asked me

$$\text{“What is the square root of 9 times 9?”}$$

Knowing that he was familiar with negative numbers, I replied that the answer could be "+9 or −9". "No", he replied, "it's 27". Then he explained that he meant

$$\text{“the square root of 9 ... times 9”}$$

which gives 27 (Tall et al. 2017).

Subsequently, we found that the principle works not only for simple arithmetic expressions but also throughout the whole range of mathematical expressions used to specify mathematical operations as mental objects.

Flexible Use of Symbolism Dually Representing Process or Concept

The idea of an operation becoming a mental object of thought has permeated research on mathematical thinking for many years. Piaget referred to this transition at "reflective abstraction" and many other authors have formulated similar ideas using different terminology. (See Tall et al. 2000 for a general discussion.) Broadly speaking, there are two essentially different mental constructs—a *process* (or *operation*) which occurs in time, either as a procedure with a specific sequence of actions or as a more general input-output process, and a *concept* (or mental *object*) that can be conceived as a mental entity that can be manipulated in the mind. In what follows, when referring to mathematical expressions, the terms "process" and "operation" will be used interchangeably as will the terms "concept" and "object". Often the situation is seen as having two different states: one as process (or operation), and the other as concept (or object) with distinct acts of passing from one to the other.

Gray and Tall (1994) realised that an expression such as 2 + 8 can be conceived either as a process to be carried out, such as "add 2 and 8", or as a concept, the "sum of 2 and 8", which is 10. They responded to this dual and ambiguous use of the symbol by naming it a "procept". This offers a major theoretical advance because it refers to the possible use of the symbol flexibly, either as a process that could be carried out in a variety of ways or as a single mental entity that can be manipulated as a mathematical object, whichever is more useful in a given context.

Sometimes it is important to distinguish between the two meanings. As operations, two different operations can give the same object, so we often speak of them as "equivalent operations". For instance, when we speak of fractions, we say that $^3\!/_6$ and $^2\!/_4$ are "equivalent fractions" because they are certainly different as operations, but they are the same rational number, represented on the number line by a single point.

This flexible duality of expressions as process or concept occurs throughout arithmetic, algebra, calculus and more sophisticated use of symbols. Often the curriculum is designed to start with examples of specific procedures to convert one expression to another. For instance, algebraic expressions such as $(x + 1)(x - 1)$ can be multiplied out to get x^2-1, and this can be factorised to return to $(x + 1)(x - 1)$. Initially these two expressions may be seen as "equivalent" but they are also different ways of representing the same underlying mathematical object which has a single graph.

This is not the only way in which sophisticated ideas evolve. It is also possible to begin with an intuitive sense of a concept and then seek ways of constructing and calculating it. Applied mathematicians do it all the time. They start with a situation that they seek to model and use mathematics to construct and test the model to see if it gives a good prediction.

Making Sense of Mathematical Expressions Dually Representing Operation or Object

Given the way in which rules of precedence violate the directional way of reading text, Tall (2019) proposed a simple notation to use the distinction between process and concept to give a natural meaning to the rules of precedence. Starting with a single operation such as 2 + 8, simply put boxes around the objects:

$$\boxed{2} + \boxed{8} \text{ is the operation of adding the objects } \boxed{2} \text{ and } \boxed{8}.$$

If the whole expression is conceived as an object, put the box around the whole expression:

$$\boxed{2+8} \text{ is the object which is the result of adding 2 and 8.}$$

This relates directly to the different ways we articulate an expression to indicate which operations should be performed first and then the result should be considered as an entity to be operated upon. For instance

$$2\times 3\ldots+4 \text{ can be interpreted as } \boxed{2\times 3}+4 \text{ which is } 6+4, \text{ giving } 10,$$

while

$$2\ldots\times 3+4 \text{ can be interpreted as } 2\times\boxed{3+4} \text{ which is } 2\times 7, \text{ giving } 14.$$

The general principle of addition tells us that if there are several additions in a box, then the order does not matter, so

$$2\times\boxed{3+4+5} \text{ is the same as } 2\times\boxed{4+5+3}.$$

There is a corresponding principle for a box containing several multiplications.

There are a few conventions that require individual treatment. For example, if letters are involved in an algebraic expression, such as $2 \times a \times b$, then the convention is to omit the multiplication signs, writing it as $2ab$. There is no problem here:

$$\text{The operation can be written as } \boxed{2}\ \boxed{a}\boxed{b} \text{ and the object as } \boxed{2\,a\,b}.$$

In dealing with the contraction 2½ for 2 + ½, boxing the expression as an operation requires an explicit addition sign $\boxed{2}+\boxed{{}^{1}\!/_{2}}$ as the symbol $\boxed{2}\boxed{{}^{1}\!/_{2}}$ could be confused with the product.

The exponent notation for x^2 can be written as $\boxed{x}^{\boxed{2}}$ as an operation and as $\boxed{x^2}$ as an object.

What is important for the human brain is to reduce the complication by not using unnecessary notation. What matters is the principle of seeing operational symbols flexibly as process or concept and to interpret the operations in an expression according to their precedence. For instance, in the quadratic expression

$$2x^2+7x+6$$

it is not necessary to put boxes around the numbers. Visualising the terms x^2, $2x^2$ and $7x$ as single objects, the expression can be seen as

$$2\boxed{x^2}+7x+6$$

or, in the usual notation, as

$$2x^2 + 7x + 6$$

where now the reader can flexibly see x^2 as an object and $2x^2$ and $7x$ as objects which are also the product of objects. The expression is now the sum of three terms and the general principle of addition allows them to be written in any order. Individual terms can be manipulated to see $5x$ as $3x + 2x$ and 6 as 3×2 and the expression $2x^2 + 7x + 6$ can be factorised as $(2x + 3)(x + 2)$.

To be able to manipulate expressions in this way requires considerable flexibility on the part of the individual. In a traditional algebraic curriculum, the reading of more complicated expressions is often guided by mnemonics such as PEMDAS in the USA or BIDMAS in the UK to specify successive levels of precedence.

PEMDAS, remembered as “Please Excuse My Dear Aunt Sally”, sets the order of precedence as “Parentheses, Exponents, Multiplication, Division, Addition, Subtraction”; BIDMAS gives “Brackets, Indexes, Division, Multiplication, Addition, Subtraction”. The situation is more complicated because the order is actually P > E > M = D > A = S or B > I > D = M > A = S where > denotes a higher level of precedence and = denotes an equal level. The rule states that higher precedence operations are performed first and equal precedence are performed left to right.

The use of this mnemonic proves to be highly problematic. Brain research reported earlier shows that merely learning the mnemonics by rote may link to language areas in the brain but not to the areas involved in fundamental human sense of space, time and number. Our new view of understanding meaning through articulation and flexible interpretation of symbol as process or concept now offers a new way of linking visual symbolism to fundamental human ideas of spatial layout and number.

There is also a further limitation of the mnemonics PEMDAS and BIDMAS because they only apply to binary operations $a + b$, $a - b$, $a \times b$, $a \div b$ and a ^ b (written as a^b) and not to unary operations such as the additive inverse, $-a$ and square root $\sqrt{a}$, nor to more sophisticated operations such as matrix multiplication, limits, differentiation, integration and other more advanced symbolisms that require new rules of operation.

The principle of articulation generalises naturally to give meaning to more advanced concepts. A typical instance is the square of a negative quantity $-x^2$ which can be articulated as

$$\text{“minus } x \,[\text{gap}]\, \text{squared”, or as “minus } [\text{gap}]\, x \text{ squared”.}$$

These give the two different meanings:

$$(-x)^2 \text{ and } -(x^2).$$

The same idea also clarifies the meaning of x^2 when a negative number is substituted for x. College students may find difficulty in substituting "x equals minus 2" in "x squared". Is it "minus *two squared*" as −4 or "*minus two* squared" as +4? (McGowen and Tall 2010). The articulation principle clarifies this distinction.

What becomes apparent in this long journey through sense-making in arithmetic and algebra is that it is possible to make sense of the conventions adopted in traditional algebraic notation by building from the principle of articulation, the general principles of addition and multiplication and the duality of expressions as process and concept. This approach links naturally to what has been discovered about the workings of the brain where mathematical thinking at all levels benefits from making mental links between concepts in space, time and number.

What is even more remarkable is that this analysis generalises to more sophisticated expressions written spatially using templates as laid out in modern digital software.

We have already seen a spatial layout when a power is written raised as a superscript. Possibilities proliferate with symbolism for limits, summation, integrals, matrix layouts and so on. These can be written by hand or built up using software templates such as MathType or specified symbolically using languages such as TeX. Figure 1.3 shows the layout for the general solution of a quadratic equation.

Reading an expression involves scanning the spatial layout to make sense of it. By starting with the whole expression as an object, it is possible to see it as flexibly as a process with sub-expressions as objects in boxes, and then to dig hierarchically down into the objects re-imagined as processes to give flexible meaning for the whole expression (Fig. 1.4).

This successive focus on the whole as an object and then as a process, to see the constituent parts of the process as objects that can then be further broken down, is essentially how successful thinkers can intuitively see the hierarchical structure of the expression. Now it can be explained explicitly to encourage a broader range of the population to make sense of expressions.

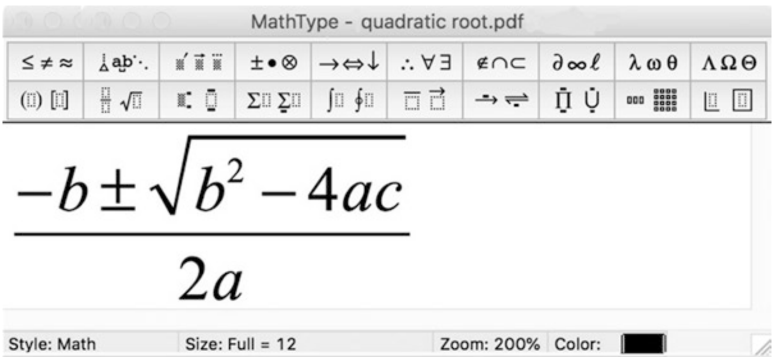


Fig. 1.3 Spatial layout of an expression

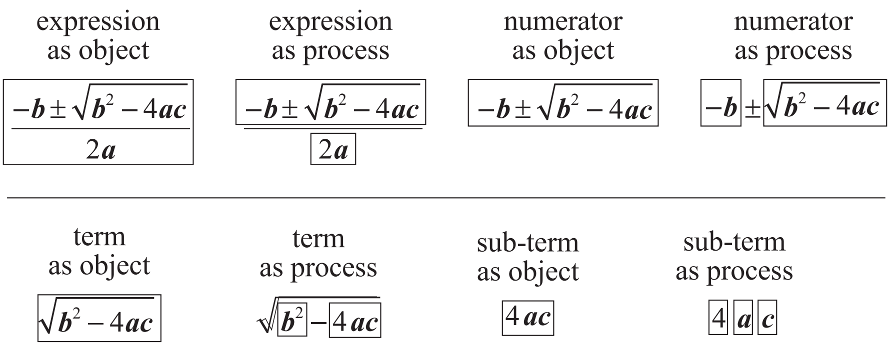


Fig. 1.4 Sub-expressions as operation or object

Equations

Once teachers or learners have the insight afforded by the meaning of the principle of articulation and the flexibility of expressions as process or object, this can give meaningful new ways of interpreting equations. An equation consists of two expressions with an equal sign between them. The new insight allows an equation to be seen in different ways, depending on whether either or both of the two expressions are process or concept.

For a young child, an equation in arithmetic such as

$$2+3=5$$

is usually read from left to right as an *operation* in which 2 + 3 is seen to give the result 5. This is in the form “process = number”. An algebraic equation in the same form, such as

$$2x+3=9$$

can be seen as a process to produce the output object 9:

$$2x+3=\boxed{9}.$$

Seen as a succession of steps, the process can be written as

$$\boxed{x}\xrightarrow{\times 2}\boxed{2x}\xrightarrow{+4}\boxed{10}$$

The process can then be undone by reversing the steps:

$$\boxed{3}\xleftarrow{\div 2}\boxed{6}\xleftarrow{-4}\boxed{10}$$

which immediately tells us that the original input x must be 3.

However, an equation with an expression on both sides, for example,

$$3x-2=2x+1$$

cannot be "undone" in the same way. This might be solved by guessing a value for x which works, or by seeing both sides as the same object, written as

$$\boxed{3x-2}=\boxed{2x+1}.$$

We can then operate on the equation by "doing the same operation to both sides" which retains the equality of the new sides. Once the original equation can be imagined as having an object on either side, we can do this in standard notation, by adding 2 to both sides to get

$$3x-2+2=2x+1+2.$$

Using the general principle of addition, this simplifies to

$$3x=2x+3$$

and, taking $2x$ from both sides gives

$$x=3.$$

A student who sees an expression only as a process and not as an object is more likely to be able to solve an equation of the form "expression = number" by "undoing" than solve an equation with expressions on both sides. This is studied extensively in the literature and was named "the didactic cut" (Filloy and Rojano 1989).

A teacher who has given meaning to expressions using the principle of articulation and has grasped the flexibility of expression as process or concept has a new way of giving meaning to equations. "Doing the same thing to both sides of an equation" in the form "object = object" ends up either with both sides always being the same (an "identity") or with the equation only being satisfied by certain values of the unknown (an "equation"). The first case occurs with an equation such as

$$2\left(x+3\right)=2x+6$$

or

$$\left(x+y\right)\left(x-y\right)=x^2-y^2.$$

In our new way of thinking, this is a single object (a procept) with different processes to calculate it.

A teacher who belongs to a community of practice that makes sense of expressions in this way may offer enlightenment where others may only see many complications arising from the didactic cut. But whether this transition to a new way of thinking is possible for teachers depends on their current beliefs and whether the transition, as seen from their current practice, is a transgression or an enlightenment.

The ability to see the equals sign used in a flexible way has further benefits as the mathematics evolves in sophistication. An equation in the form "variable = expression" may take the form of a *definition* of the variable (as a mental object) given by a process. For instance, the equation $y = x^2$ defines the (dependent) variable y in terms of a process operating on the (independent) variable x.

This applies in more sophisticated theory, such as infinite sums in the calculus where

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots$$

which defines the mathematical expression $\sin x$ computed as a potentially infinite process. The left-hand side is now a mental object which can be calculated as accurately as required by adding enough terms.

Over the long term, making sense of expressions, initially through the principle of articulation and then through the flexibility of expressions as process or concept, offers a consistent approach that encourages the learner to build on supportive ideas with confidence and deal with problematic aspects as they arise. Whether this approach is successful or not will depend on how current communities of practice see it as a transgression from their accepted practices or an enlightenment to move forward into the future.

Building a New Framework for Long Term of Mathematics

While various approaches to the curriculum have led to "Math Wars" arguing between approaches to mathematics learning, we can now shift to a higher multi-contextual level where learning "the basics of arithmetic" can be related flexibly to the meaning of expressions.

As children experience mathematical ideas in practical contexts, they will naturally pick up aspects related to each context. Making sense of different contexts to draw out common ideas is more complicated than having available simple principles that work in multiple contexts.

This is part of a much broader framework for making long-term sense in mathematics as a whole. In *How Humans Learn to Think Mathematically* (Tall 2013) I formulated a framework for long-term mathematical thinking beginning from the child's perceptions and operations with the physical world and with others in society. One strand of development senses the properties of objects, initially physical, and then constructed mentally, which I termed *conceptual embodiment*. Another strand focuses on the properties of operations that I termed *operational symbolism*. Both of these develop in sophistication from *practical mathematics* based on the *coherence* of properties that occur in practice to *theoretical mathematics* where properties are defined and relationships are deduced one from another in what may be termed *consequence*.

At the turn of the twentieth century, a further strand developed based on *properties* defined using set theory or logic which I termed *axiomatic formalism*. For many mathematicians, formal mathematical proof starts with Euclidean geometry. But there is a huge difference between mathematics based on properties of pictures or on known calculations and mathematics based on formal definition and proof. Prior to the end of the nineteenth century, the study of mathematics and science based on naturally occurring phenomena was described as "natural philosophy". I therefore distinguish "theoretical mathematics" based on "natural phenomena" from "axiomatic formal mathematics" based on set theory and logic (Fig. 1.5).

Figure 1.5 is a much simplified view of the theoretical framework developed in Tall (2013), based on the new information available from neuroscience. I termed the three main strands as "worlds of mathematics" because each world represents a fundamentally different way of thinking that evolves both in history and in the individual. Conceptual embodiment exists in many species and in human ancestors several hundred thousand years ago. Operational symbolism evolved in Homo sapiens in the last fifty thousand years, proliferating in various communities in Egypt, Babylon, India and China around five thousand years ago, becoming increasingly theoretical in Greek mathematics with the first flowering of mathematical proof two and a half thousand years ago. Axiomatic formal mathematics has been around for little more than a century. Now new possibilities are emerging in our digital age enabling Homo sapiens to use new digital tools to enhance the possibilities of enactive interface, dynamic visualisation, symbolic computation and emergence of new forms of artificial intelligence.

In this ongoing evolution, the biological brain evolves slowly. There is no reason to suppose that the biological brain of the ancient Greeks is substantially different from our own. In contrast, the technical evolution of digital tools

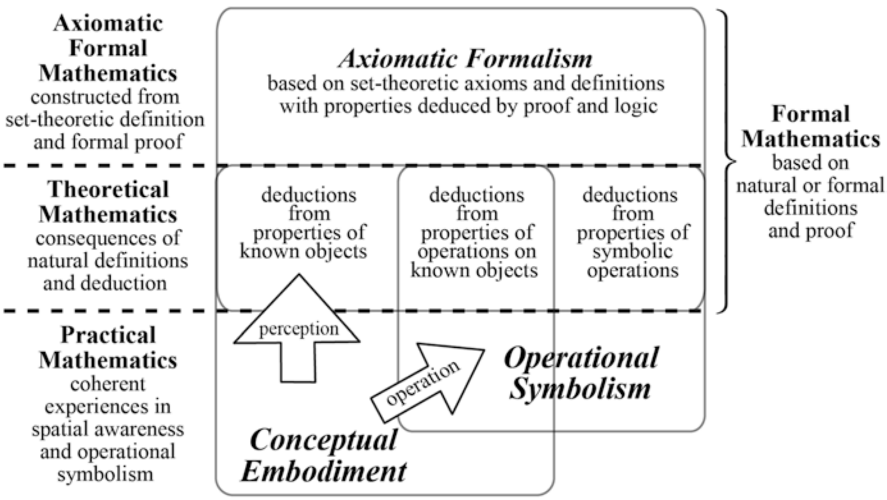


Fig. 1.5 The long-term development of mathematical thinking

available to support the mathematical mind that have occurred within a generation is immense. Although we now know that the biological brain is more complex than a simple duality between left and right brain, it still continues to support conceptual embodiment and operational symbolism with the forebrain taking an increasing role in integrating mathematical thinking in new forms of axiomatic formalism.

It is interesting to note that the diagram in Fig. 1.5 nowhere explicitly mentions the role of language. Instinctively, when I originally thought about the framework, I saw mathematical thinking to be related to the complementary roles of visual imagination, sequential symbolic operation and later logical deduction, with verbal language being used to describe connections between different parts of the framework.

The resulting two-dimensional picture gives only a partial idea of the broader complexity of the workings of the human brain. For example, it focuses on cognitive aspects that occur in the surface areas of the cortex and says little about the activity of the limbic system in the centre of the brain that not only performs many cognitive tasks relating to short- and long-term memory but also responds emotionally to supportive and problematic aspects of mathematical thinking.

Individuals do not operate in all areas of the framework. For everyday mathematics, as used by the vast majority of the population, all that is required is practical mathematics focusing on the coherence of spatial perception of the properties of objects and symbolic operation.

Those involved in mathematical applications including STEM subjects (science, technology, engineering, mathematics) usually only require practical and theoretical mathematics.

Only a small percentage of the population studying pure mathematics and logic use axiomatic formal mathematics.

Extending the Framework

Although the picture places axiomatic formal mathematics at the top of the figure, this is by no means the end of the story. Among the properties proved in axiomatic systems, certain theorems called "structure theorems" prove properties that reveal new forms of conceptual embodiment and operational symbolism. Sometimes the structure is unique in the sense that any two structures satisfying the definition have the same properties (said to be "isomorphic"). We can now see that the two structures may not only be "essentially the same", but they may also be conceived as a single entity that can be represented in different ways.

Two examples of such unique structures are "the natural numbers" and "the real numbers" which can be represented visually as points on a number line and symbolically using decimal notation. Other axiomatic systems may have many different examples, such as the concepts of "group" or "vector space". These have structure theorems that allow them to be classified and represented as mental objects or as operational structures. For example, a "finite dimensional vector space" can be proved to have a coordinate system, visualised as two- or three-dimensional space or imagined mentally in higher dimensions, where the coordinates allow the vectors to be manipulated symbolically (Tall 2013).

How the Brain Makes Sense of More Sophisticated Mathematics

Knowledge of the workings of the brain can now explain in simple terms how the biological brain can make sense of sophisticated mathematical ideas that are considered to be problematic for some and enlightening to others. Again, this explains evolution of ideas corporately in history as well as in the growing individual.

How the Eye Follows a Moving Object, Giving Meaning to Constants and Variables

When the eye follows a moving object, it starts using the same initial action as reading text, with a single jump to focus on the object, but then it follows the object smoothly as it moves. You can sense this by holding a finger in front of your eye at a comfortable distance away and move it sideways, keeping your gaze on the finger as it moves. The finger stays in focus while the background is blurred. In this way the eye is set up to follow moving objects smoothly. It is therefore natural to imagine a point on a line which moves. It is also natural to distinguish between a fixed point on a line (a constant) or a moving point on a line (a variable) (Fig. 1.6).

This has profound implications for the historical and individual imagination for constant quantities and variable quantities, including variables that can become arbitrarily small. In history this gave rise to ideas of indivisible quantities that are small but no longer further divisible and infinitesimal quantities, either as potential never-ending processes or as actual mental objects. This interpretation of infinitesimals as variable quantities offers a new way of considering the Greek arguments about potential and actual infinity. It sheds new light on the development of infinitesimal ideas in the calculus, in particular, how Leibniz may have imagined different orders of infinitesimality (Tall 2013, Chap. 13) or how Cauchy imagined infinitesimals as sequences that tend to zero (Katz and Tall 2012; Tall and Katz 2014).

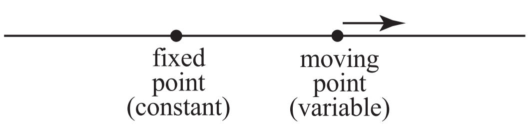


Fig. 1.6 Constant and variable points on a line

How Dynamic Movement Can Represent Infinitesimals as Process and Concept

An infinitesimal may be visualised as a variable point on a line. For example, consider a rational function $f(x) = p(x)/q(x)$ where p and q are polynomials with q non-zero. Draw the graph of $y = f(x)$ and the vertical line $x = k$. Figure 1.7 shows the vertical line intersecting the graphs of $y = c$, $y = x$, $y = x^2$ at heights c, k, k^2. For constant $c > 0$, as k decreases to zero, the points height k and k^2 fall below the point height c and the variable points k and k^2 are eventually less than any positive real number c. In this sense they are *infinitesimal*. Moreover, k^2 is smaller than k. If we think of k as being of order 1, then k^2 is of order 2 and, in general, as n increases, k^n is an even smaller infinitesimal of order n. Using such a visual representation, we can imagine infinitesimals of any order.

Of course, this argument may be rejected as a transgression, as it was by many contemporary critics of the early calculus. But for others, it offers enlightenment.

Using axiomatic formal arguments, we can go even further. Consider any ordered field K that contains the real numbers as an ordered subfield. (Remember that the field of real numbers is unique as the one and only *complete* ordered field, in the sense that any non-empty subset of real numbers has a least upper bound.) Any element x in K can be compared with any real number c, so we know that either $x > c$ or $x = c$ or $x < c$. We can then separate the elements of K into three distinct categories:

(1) Those x in K such that $x > c$ for all real c
(2) Those x in K such that $x < c$ for all real c
(3) Those x in K which lie between two real numbers, $a < b < c$.

We call those in (1) *positive infinite* elements, those in (2) *negative infinite* and those in (3) *finite*. If x is finite, then the set of real numbers less than x is non-empty (it contains a) and bounded above (by b); therefore, by the completeness axiom for the real numbers, it has a least upper bound c. It is then a straightforward deduction

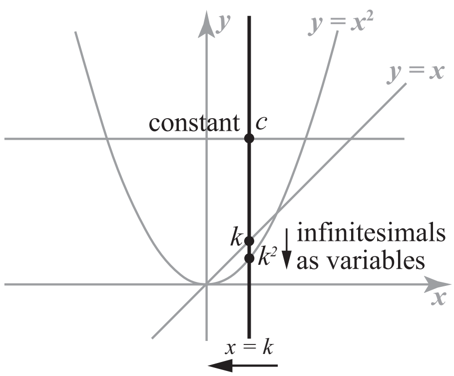


Fig. 1.7 Infinitesimals as variables tending to zero

to prove a structure theorem that the element $e = x - c$ is either a (positive or negative) infinitesimal or it is zero. This tells us that every finite element x is (uniquely) of the form $x = c + e$ where c is a unique real number and e is an infinitesimal or zero. We call this number c the *standard part* of x.

This transition to the new axiomatic formal context transforms a transgression into an enlightenment. It opens the flood gates. The structure theorem enables us to visualise infinitesimals (and their multiplicative inverses which are infinite) using simple algebraic maps (which I term *optical microscopes*) mapping x onto the standard part of $(x - c)/e$. This maps c to 0 and $c + e$ to 1, spreading out infinitesimal detail near c so that we can see it as a real picture! It also reveals that, seen through an optical microscope, the image of an infinitesimal part of a differentiable function is a real straight line. This links the intuitive idea that a differentiable function is "locally straight" under high magnification to a perfect visualisation in the axiomatic formal world (see Tall 2013, Chap. 11 or Stewart and Tall 2014, Chap. 15, for details).

These ideas easily extend to see infinitesimal detail for complex functions (Stewart and Tall 2018, Chap. 15). They are part of a much bigger framework of multidimensional analysis including visualisations of analysis in higher dimensions and sensible meanings for partial derivatives (Tall 2013, Chaps. 11 and 13).

How the Eye Reads Through a Written Proof to Make It Meaningful

In addition to the way the brain can interpret pictures, it can also scan a written proof, not just line by line, but also by looking back at significant steps and getting an overall grasp of the proof structure. Using eye-tracking techniques, Inglis and Alcock (2012) confirmed that undergraduates devoted more of their attention to parts of proofs involving algebraic manipulation and less to logical statements than expert mathematicians. Hodds et al. (2014) developed a technique of "self-explanation" in which students were encouraged to read a proof line by line, to identify the main ideas, get into the habit of explaining to themselves why the definitions are phrased as they are and how each line of a proof follows from previous lines. They were counselled not to simply paraphrase the lines of the proof by saying the same thing in different words, but to focus on making connections to grasp the main argument and explain how the given assumptions and definitions in previous lines led to the current line and contribute to the following lines. Students who had worked through these materials before reading a proof scored 30% higher than a control group on a subsequent occasion.

Notice that, in this case, the explanations were expressed linguistically, but the focus once more is on the relationships between ideas. A focus on making personal links is more likely to give a more coherent personal knowledge structure in the longer term.

The Role of the Limbic System in Enhancing and Inhibiting Mathematical Thinking

Up to this point, the presentation has focused on:

- *Cognitive* aspects of how individuals think about mathematical structures.

To gain a broader understanding of the long-term development of mathematical thinking, it is also essential to consider:

- *Affective* aspects that enhance and suppress the making of mental connections.

Mathematics evokes a wide range of emotions in different individuals. Some experience great pleasure in solving a difficult problem, even relishing the challenge. Others suffer a sense of tension and anxiety that interferes with their ability to answer a mathematical question or manipulate numbers. The anxiety can range from a mild sense of insecurity to a full-blown fear and loathing of mathematics.

These emotions arise in the limbic system in the centre of the brain. This is a collection of structures that support a variety of functions, including cognitive links between short-term and long-term memory, but also gives rise to primitive emotional responses of pleasure or pain. In particular, it responds to challenges or to danger with an immediate “fight or flight” reaction that suffuses the whole brain with neurotransmitters that excite or inhibit mental connections.

Confident students who rise to the challenge are placed on alert, ready to tackle the situation. Those who find the mathematics difficult or even impossible are likely to have their mental connections suppressed, causing them to freeze mentally and even be unable to respond. It is not just that students suffering from mathematics anxiety are unwilling to think mathematically. When their mental connections are depressed, they may not be able to think about mathematics at all.

Research identifies many diverse factors related to mathematics anxiety, including negative experiences of mathematics, fear of being asked questions in front of others, social deprivation, poor self-image, poor memory and so on. Here we are only concerned with one aspect: the long-term relationship between the individual and mathematics. A biological brain which has rich flexible connections and an awareness of the need to deal with problematic aspects of new contexts is much more likely to succeed than one which has limited rote-learned knowledge. A brain suffused with neurotransmitters that enhance mental connections is better placed to construct new meanings than a brain with mental connections that are suppressed.

Strategies for Enhancing Long-Term Mathematical Thinking

Moving to the Future in Different Communities of Practice

Recent international comparisons in TIMSS (2015) and PISA (2015) reveal widespread differences in long-term mathematical competence. PISA shows East Asian countries scoring highly in the first seven places out of 65 participants, with the Netherlands (10th) among those above average, the UK (26th) being average and the USA (36th) slightly below average, with a long tail including Brazil (58th). I selected these countries because they include some of the areas where I have had direct research experience.

As a consultant in a project involving 20 economic communities around the Pacific Rim, it was my privilege to participate in a multicultural overview of different communities developing Japanese Lesson Study and editing (but not writing) the English version of the first three volumes of the Japanese Junior High School mathematics (Isoda and Tall 2018). These books are written by mathematicians and teachers to encourage students to think for themselves, informed by research in mathematics education. The lesson sequence is organised to give the students experiences that will be useful for solving problems encountered later in the sequence. The sequence is then modified over successive implementations to build a stable version intended for general use.

The development of Lesson Study is broadly consistent with the framework formulated here with some differences. For example, mathematics education research distinguishes between "three twos" and "two threes" and the curriculum initially retains this difference as processes rather than seeking their unity as an object. Perhaps the next iteration of the curriculum will address this aspect.

The building of the long-term curriculum reveals a problematic transition from practical to theoretical mathematics. In the Netherlands, "realistic mathematics" introduces children to make sense of practical situations as active participants solving meaningful problems in imaginative ways (Van den Heuvel-Panhuizen and Drijvers 2014). This approach has spread internationally with widely acclaimed success. Yet it proved to lead to a situation where students in the Netherlands going to university were less well prepared.

Advocates of realistic mathematics investigated this phenomenon in three PhD studies involving "subtraction under 100", "fractions" and "algebra", to show that:

> Dutch students "proficiency fell short of what might be expected of reform in mathematics education aiming at conceptual understanding. In all three cases, the disappointing results appeared to be caused by […] the textbooks' focus on individual tasks […] with a lack of attention for more advanced conceptual mathematical goals, constitut[ing] a general barrier for mathematics education reform" (Gravemeijer et al. 2016: 25).

The authors came to the conclusion that it is not a weakness in the theory of realistic mathematics, but in the implementation of theory: that "realising" mathematical ideas needs extending to grasp the underlying theoretical ideas in more advanced mathematics.

An attempt to use Lesson Study in the Netherlands to address the problem for teenagers studying calculus proved initially to be problematic as the teachers followed their experience of Dutch culture including “following the textbook closely, the strict school guidelines and the pressure for high exam results” (Verhoef and Tall 2011). Only in the second year of the study did teachers begin to grasp the students’ personal ways of thinking to make sense of the relationship between dynamic visualisation and symbolism using *Geogebra* (Verhoef et al. 2014).

In the USA there is a vast quantity of research literature studying the complication of ideas in arithmetic, fractions and algebra. In general, this literature focuses more on the complications of mathematics and its implementation in the classroom. But where is the extended research to consider how to make sense of the simple idea of the principle of articulation and its resulting flexibility of symbolism as process or concept?

In the UK with a maximum political cycle of 5 years, politicians need results that vindicate their policies within such a period. Given the perceived lack of competitiveness in international comparisons, they sought to find how the more successful countries operate, seeking insights from Singapore, Shanghai, Finland and elsewhere, finding that different social and cultural attitudes made it problematic to transfer the expertise. In Brazil, which scores low in PISA studies, research revealed teachers teaching students rote-learnt rules to pass tests which work in simple cases but fail in general: for example, solving a quadratic equation using the formula, when many students could not manipulate a quadratic into the form $ax^2 + bx + c = 0$ to use the formula (Tall et al. 2014).

In both high-scoring and low-scoring communities on the PISA scale outside East Asia, the desire to “teach to the test” may offer some short-term success, but over the long term, rote learning of a range of disconnected methods may act as a barrier to the development of more sophisticated long-term mathematical thinking.

Reflections

This chapter has offered evidence relating to how the human brain makes sense of increasingly sophisticated mathematical ideas by referring to neurophysiological research and simple ideas that can be observed by teachers and learners in the classroom.

It acknowledges changes in meaning over the longer term as the learner encounters more sophisticated contexts. To offer positive support to address these changes, it focuses on fundamental aspects that remain supportive over several changes of context as a secure foundation to help learners make sense of problematic changes in meaning.

Focusing on how we articulate mathematical expressions can offer profound insight into the long-term learning of arithmetic and algebra. Other observations into the workings of the human brain offer insights into how we think about mathematical ideas at all levels from newborn children to the wide variety of adult thinking.

This is part of a broader theory of long-term mathematical development including both historical and individual growth that takes account of cognitive, affective and social aspects.

However, participants in different aspects of the enterprise will have their own views on how they should proceed. Different communities of practice may have radically different approaches that conflict with each other and one community may see a change in meaning as an enlightenment, while another community may see as a transgression. This has led to widespread differences involving "math wars" between different approaches and it is highly unlikely that a single approach will provide a universal solution.

The contribution of this chapter is to reflect on simple yet profound ideas that may enlighten different communities in ways that offer each community appropriate insight.

Difficulties encountered by young children in arithmetic may grow into mathematics phobia in adults. Mathematics educators often focus on creativity, encouraging young learners to see a specific pattern in many imaginative different ways. The framework recommended here uses the principle of articulation to clarify the meanings of expressions and, by interpreting expressions flexibly as process or object, it goes on to show how equivalent, but different, processes can be conceived as a single object. This makes explicit a long-term implicit development in the curriculum, where equivalent fractions are later seen as the same rational number marked as a single point on the number line, and equivalent algebraic expressions are later seen as a single entity with the same graph. A parallel focus on specific examples and underlying structure offers the possibility of a closer relationship between arithmetic and algebra.

At a more sophisticated level, by realising how the human eye sees variable quantities, the framework offers a new understanding of the use of infinitesimals in the calculus, linking together the different approaches in pure and applied mathematics.

At an even more advanced level of thought, the notion of "structure theorem" links set theoretic mathematics and logic back to visual intuition and meaningful symbol manipulation.

The possibilities are immense, especially at a time in history where new technology enables the fundamental operation of Homo sapiens to function in new ways that not long ago would have been inconceivable. Digital technology offers enactive control of dynamic imagery to support visual intuition, and symbolic manipulation to support operational symbolism, together with the ongoing evolution of artificial intelligence that currently falls short of the full capacity of the human brain. It is an exciting time to see how the biological brain uses new facilities to operate evermore powerfully as a mathematical mind.

References

Amalric, M. and Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians, *Proceedings of the National Academy of Science of the US*A. 113 (18) 4909–4917. https://doi.org/10.1073/pnas.1603205113

Chomsky, N. (2006). *Language and mind*. Cambridge: Cambridge University Press.

Dehaene, S., Spelke E., Pinel, P., Stanescu, R., and Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science* 284 (5416): 970–974.

Filloy, E., and Rojano, T. (1989). Solving equations, the transition from arithmetic to algebra, *For the Learning of Mathematics* 9 (2): 19–25.

Gravemeijer, K., Bruin-Muurling G., Kraemer J. M., van Stiphout, I. (2016). Shortcomings of mathematics education reform in The Netherlands: A Paradigm Case? *Mathematical Thinking and Learning* 18 (1): 25–44, DOI: https://doi.org/10.1080/10986065.2016.1107821.

Gray, E. M. and Tall, D. O. (1994). Duality, ambiguity and flexibility: A proceptual view of simple arithmetic. *Journal for Research in Mathematics Education* 26 (2): 115–141.

Hadamard, J. (1945). *The psychology of invention in the mathematical field*. Princeton: Princeton University Press (Dover edition, New York 1954).

Hodds, M., Alcock, L., and Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education* 45: 62–101.

Inglis, M., and Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*, 358–390.

Isoda, M. and Tall, D. O. (eds.) (2018). *Mathematics: Junior high school 1, 2, 3*, English Edition. Tokyo: Gakkotosyo.

Katz, M. and Tall, D. O. (2012). The tension between intuitive infinitesimals and formal analysis. In: Bharath Sriraman, (ed.), *Crossroads in the history of mathematics and mathematics education, (The Montana Mathematics Enthusiast Monographs in Mathematics Education 12),* pp. 71–90.

Kozielecki, J. (1987). *Koncepcja transgresyjna człowieka* [Transgressive concept of a man]. Warszawa: PWN.

Lave, J. and Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press. ISBN 0-521-42374-0.

Maruyama, M., Pallier, C., Jobert, A., Sigman, M., Dehaene, S. (2012). The cortical representation of simple mathematical expressions. *Neuroimage* 61(4):1444–1460.

McGowen, M. and Tall, D. O. (2010). Metaphor or met-before? The effects of previous experience on the practice and theory of learning mathematics. *Journal of Mathematical Behavior* 29: 169–179.

Monti, M. M., Parsons, L. M., Osherson D. N. (2012) Thought beyond language: Neural dissociation of algebra and natural language. *Psychological Science* 23(8): 914–922.

Nakai, T., Sakai, K. L. (2014). Neural mechanisms underlying the computation of hierarchical tree structures in mathematics. *PLoS One* 9(11): e111439.

Piaget, J. (1952). *The child's conception of number*. New York: Norton.

Pieronkiewicz, B. (2014). Report of mathematical transgressions conference, *Annals of the Polish Mathematical Society, 5th Series: Didactica Mathematicae* 36: 163–168.

PISA (2015). *Programme for International Student Assessment*. Retrieved from https://en.wikipedia.org/wiki/Programme_for_International_Student_Assessment, November 30, 2018.

Shum, J., Hermes, D., Foster, B. L., Dastjerdi, M., Rangarajan,V., Winawer, J., Miller, K., Parvizi, J. (2013). A brain area for visual numerals. *Journal of Neuroscience* 1: 6709–6715.

Stewart, I. N. & Tall, D. O. (2014). *Foundations of Mathematics*, Second Edition. Oxford: OUP. ISBN: 9780198531654

Stewart, I. N. & Tall, D. O. (2015). *Algebraic Number Theory and Fermat's Last Theorem*, Fourth Edition. Boca Raton: CRC. ISBN 9781498738392

Stewart, I. N. and Tall, D. O. (2018). *Complex analysis*, Second edition. Cambridge: Cambridge University Press. ISBN: 978-1108436793.

Tall, D. O, Thomas, M. O. J., Davis, G. E., Gray, E. M., and Simpson, A. P. (2000). What is the object of the encapsulation of a process? *Journal of Mathematical Behavior* 18 (2): 1–19.

Tall, D. O. (2013). *How humans learn to think mathematically*. New York: Cambridge University Press. ISBN: 9781139565202.

Tall, D. O. (2019). Complementing supportive and problematic aspects of mathematics to resolve transgressions in long-term sense making. *The Fourth Interdisciplinary Scientific Conference on Mathematical Transgressions*, Krakow (to appear).

Tall, D. O. and Katz, M. (2014). A cognitive analysis of Cauchy's conceptions of function, continuity, limit, and infinitesimal, with implications for teaching the calculus. *Educational Studies in Mathematics* 86 (1): 97–124.

Tall, D. O., Tall, N. D. and Tall, S. J. (2017). Problem posing in the long-term conceptual development of a gifted child. In: Martin Stein (ed.) *A life's time for mathematics education and problem solving*. *On the occasion of Andràs Ambrus' 75th Birthday*, pp. 445–457. WTM-Verlag: Münster.

Tall, D., de Lima R. and Healy, L. (2014). Evolving a three-world framework for solving algebraic equations in the light of what a student has met before. *Journal of Mathematical Behavior* 34: 1–13.

Van den Heuvel-Panhuizen M. and Drijvers P. (2014). Realistic mathematics education. In: S. Lerman (ed.), *Encyclopedia of Mathematics Education*, https://doi.org/10.1007/978-94-007-4978-8, Springer: Dordrecht. https://www.icrme.net/uploads/1/0/9/8/109819470/rme_encyclopaediamathed.pdf.

TIMSS (2015). *TIMSS 2015 results*. Retrieved from http://timss2015.org, November 30, 2018.

Verhoef, N. C., and Tall, D. O. (2011). Teacher's professional development through lesson study: effects on mathematical knowledge for teaching. *Proceedings of the 35e Conference of the International Group for the Psychology of Mathematics Education*, Vol. 4 (pp. 297–304). Ankara, Turkey.

Verhoef, N. C., Tall, D. O., Coenders, F. and Smaalen, D. (2014). The complexities of a lesson study in a Dutch situation: Mathematics Teacher Learning *International Journal of Science and Mathematics Education* 12: 859–881. doi: https://doi.org/10.1080/19415257.2014.88628.

Limbic System (n.d.) Wikipedia: Limbic System. Retrieved from https://en.wikipedia.org/wiki/Limbic_system. November 30, 2018.

Chapter 2
Compression and Decompression in Mathematics[1]

Mark Turner

The mathematician Hermann Weyl (1952) explained that a body in the world incorporates its environment, or rather, its ancestors' environment of evolutionary adaptation. An unmoving minuscule organism that floats in the ocean at a depth where gravity and water pressure balance each other out is nearly spherical, because for such an organism all directions are functionally the same, and so selection produced a suitable body. Its experience has spherical symmetry and so does its body. A plant fixed to the ground—like a tree—is asymmetric top to bottom because gravity creates an environment where all directions are not the same. The tree's environment is characterized by a constant difference: the gravity vector points down; a tree's form must deal with that. On the other hand, trees have mostly equivalent environments in any direction perpendicular to the vertical gravity vector—"*mostly* equivalent" because there are variations in the relative path of the sun, the flow of water, a strong onshore wind, and so on. Accordingly, trees, ignoring these local differences, for the most part have bodies that are the same in all directions perpendicular to the vertical axis. An animal on the ground that moves has different experience in the direction it is headed than it has from the direction whence it came, and so has a body that is different front to back. We run into things we are moving toward, not things from which we are moving away. We experience gravity and we move. Accordingly, our bodies are, on the outside, anatomically, pretty much different up-down and front-back, but not so much left-right. What can happen from the left can happen from the right. What we can do to the right we can pretty much do to the left. We can mostly reverse our experience to the left versus right just by doing an about-face. We are set up for this: it would

This chapter is based on chapters in Turner, M. (2014). *The Origin of Ideas: Blending, Creativity, and the Human Spark*. New York: Oxford University Press.

M. Turner (✉)
Case Western Reserve University, Cleveland, OH, USA
e-mail: mark.turner@case.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_2

be extremely inefficient if we had to learn everything twice, once to the left and once to the right. Instead, our brains are built to be able to map left to right and right to left and pick out the midline, the axis, about which the pattern is symmetric. This kind of symmetry is called "bilateral symmetry," or "heraldic symmetry" (Turner 1991).

We have several immediate, human-scale ideas that come from our bodily knowledge of bilateral symmetry. Most important, we know that when we are symmetrically positioned about the plane that separates the left side of our body from the right side of our body (called the "sagittal" plane), whether we are sitting or standing or even moving, then we are stable, balanced, able to deal with the world, and ready with power to engage what is in front of us. We know that the same is true of other people, because of course we can make an easy mental blend of our body and their bodies to have an idea of how their bodies work. We also know that we most easily attend to something by facing it, standing full-frontal to it, locating it in our sagittal plane directly ahead. And of course we know, by blending, that other people have this experience and behavior. Just by putting our two hands against each other and pressing with equal force, we understand the equilibrium and stability that come from equal opposing forces. And we understand that, along this lateral axis, there is a linear order from one side to the other that advances to our sagittal plane and then reverses to reach the other side. Like Leonardo's Vitruvian Man, we can mark off from left to right—fingernails, dactyls, palm, wrist, forearm, elbow, upper arm, shoulder, pectoral, neck and head, pectoral, shoulder, upper arm, elbow, forearm, wrist, palm, dactyls, and fingernails. It is like a bodily palindrome. A palindrome is a piece of writing that reads the same backward and forward, like "able was I ere I saw Elba," or "Madam, I'm Adam." Whether you start on the left or the right of the human body, one side is the mirror image of the other.

This bodily knowledge of symmetry is used to make new, compressed ideas that go far beyond human scale. An institutional power, like a king, is frequently represented by bilateral symmetry, as we see in the Lion Gate at Mycenae, where two beasts (scholars debate whether these beasts are actually lions) in heraldic symmetry lean rampant against a central pillar. We understand immediately a "stability" and "power" that go a vast distance beyond our human-scale idea of symmetry, because we have blended a vast mental web of political and institutional power with that human-scale, bodily idea of symmetry.

Vast conceptions are often given an artistic or poetic or symbolic presentation that starts at the periphery, works toward the center, and then works back out toward the periphery. When Odysseus meets his mother Anticlea on a visit to the underworld (Book XI of the *Odyssey*, lines170–203), he asks her six questions, which she answers in reverse order, something like this:

A. What killed you? (171)
B. A long sickness? (172)
C. Or Artemis with her arrows? (172–173)
D. How is my father? (174)

E. How is my son? (174)
F. Are my possessions safe? (175–176)
G. Has my wife been faithful? (177–79)
G’. Your wife has been faithful. (181–83)
F’. Your possessions are safe. (184)
E’. Your son is thriving. (184–87)
D’. Your father is alive but in poor condition. (187–96)
C’. Artemis did not kill me with her arrows. (198–99)
B’. Nor did a sickness kill me. (200–201)
A’. But my longing for you killed me. (202–3)

What falls at the center of this orderly symmetry is the great theme of the *Odyssey.* Odysseus will fight his way back to Ithaka, discover that Penelope has been faithful, and, against the greatest odds, regain his position with her, their son, and his people.

This may seem at first to be the stuff of old epic, but a few minutes’ reflection will bring to mind many examples from current political rhetoric, advertising, and popular entertainment that use such heraldic bodily balance.

Much of the virtue of this blending is that, instead of having to think about many things simultaneously, we can think of one thing—a compressed blend—and use that one thing to help us key into this or that part of the decompressed mental web that is so vast. It is easier to use a *compressed* blend than a diffuse network. If we can place a compressed blend inside the decompressed network, the compressed blend can help us manage the decompressed network. Blending other ideas and concepts with bilateral symmetry creates a highly compressed, experientially grounded blend, likely to be highly intelligible, memorable, and tractable. It gives us one compressed, tractable thing instead of two or more, where some of those inputs are diffuse and mentally intractable.

Here is a demonstration of the power of thinking about one compressed thing instead of two. Consider the sequence 1, 3, 5, 7, 9 …. What is the next number? It’s easy: 11, followed by 13, 15, 17 …. Now consider a second sequence: 2, 4, 6, 8, 10 … What is the next number? It’s easy: 12, 14, 16 …. Hold those two integer sequences in mind. Now alternate between them, starting with 1. That is, take the first element from the first sequence, then the first element from the second sequence, then the second element from the first sequence, then the second element from the second sequence, and so on, like this: 1, 2, 3, 4, 5, 6, 7 …. How does the sequence continue? It’s easy: 8, 9, 10, 11.

But now, using the same two sequences, combine them again, in just exactly the same way, but instead start with the second sequence rather than the first: 2, 1, 4, 3 …. How does the sequence continue? Well, uh, 6, 5, … uh, uh, um, 8, 7 …. If you run this demonstration in a lecture hall full of extremely clever people, they are guaranteed to start laughing quickly at how they stumble over the sequence.

Why? Both tasks have the same input mental spaces, that is, the two different integer sequences. Both tasks have the same rule for constructing the third sequence—or rather, the same rule except for where you start. To that extent, the tasks place identical demands on the mind. The obvious difference is that for the first task, there is a unified, compressed blend of the two sequences, a blend in which the sequence proceeds by taking one number, and adding one, and doing that again for the next number, and so on. That blend of the two sequences can be held in the mind all at once, and we can think about this one thing instead of having to alternate back and forth between two. But in the second task, it is much harder to get a single, unified, compressed blend that can be used to juggle and access the input mental spaces.

If we want to put some stuff into a room and it does not fit, there are in general two different ways to succeed: First, get a bigger room; second, change the stuff so that it will fit. These are very different, if complementary, strategies. Changing the stuff can include folding it, packing it, stacking it, filtering the stuff so as to throw away what we do not need to keep, and so on. Most interesting, changing the stuff can include adding things to it, like, say, stackable storage bins. If we want to stack a lot of fine wine in a small space, it might be best to build good racks for it. This may seem nonsensical, because adding racks increases the amount of stuff we must fit into the room. But that is often the right strategy. The specific details of the packing can vary. Here is an analogy: In Robert Crichton's *The Secret of Santa Vittoria,* the Italian villagers have hidden very many bottles of local wine underground from the German army at the end of the Second World War. They stacked it tightly. To mislead the Germans, they also stacked a lot of wine above ground, in plain view, but stacked it using a method that requires a great deal of space per bottle. So you can store a lot more or a lot less, depending on how you arrange it. The same is true of numeric sequences. One way to arrange the combination of the two input sequences 1, 3, 5 … and 2, 4, 6, … is to create the blend 1, 2, 3, 4, 5 …. This is a compressed blend that can be held as a unified mental space in the mind: Instead of working in the mental web of inputs, you can work in the blend. The other way to arrange the combination of these two sequences is to create the blend 2, 1, 4, 3 …, but that attempt to blend is poor. It produces something that does not fit so well in the mind.

Of course, we have heard and memorized the sequence 1, 2, 3, 4 … many times, and never heard or memorized the sequence 2, 1, 4, 3 …. Accordingly, somebody might intelligently object: Does the difference in our ability to manage the two sequences stem from the fact that in the first we are reciting from long-term memory, but in the second we lack such assistance? Does this exercise merely demonstrate the obvious: that we know what we have memorized but do not know what we have not?

We can run a different demonstration to answer the question. We can show the same effect without calling on long-term memory, by working with sequences that we have never heard or memorized. Consider a sequence defined by this rule: Take every other *even* number, beginning at 256. So, 256, 260, 264 …. What is the next number? It is easy to generate it, because there is a unitary rule: Just add four to the last number. This rule makes the sequence seem like one thing. Now hold that sequence in mind. At the same time, consider a very similar sequence with the identical rule: Take every other even number beginning at 254. So, 254, 258, 262 …. What is the next number in this sequence? It is easy to generate it, because there is a unitary rule, and it is the same unitary rule: Just add four to the last number. This rule makes this second sequence seem like one thing. Now hold that second sequence in mind along with the first. What is the sequence that consists of numbers taken sequentially in alternation from the two sequences, beginning with 256? So, 256, 254, 260, 258, 264, 262 …. What is the next number? Keep going. Everyone finds it difficult not to stumble almost immediately. It is crucial to realize that a computer, doing the actual math and not needing the compressed blend, would have not the slightest difficulty.

Why is that? We have no difficulty holding *each* of the sequences in mind because for each we can make a simple blend: The blend has only two numbers, and the second is four more than the first. Wherever you want to be in the sequence, project it to the first number in the blend, and project the next number in the infinite sequence to the second number in the blend, which is four more. So just add 4, and you are done. That little blend gives you the entire sequence, or rather you can expand from the human-scale blend to any part of the infinite sequence you like. The mental web has an infinity of numbers, but the blend has only two elements, and it repeats. So there is no difficulty in keeping either one of these sequences in mind, because we can use the compressed blend. If we could hold *both* of the input sequences in mind and go back and forth between them, choosing at each turn the next number for the new sequence, we could answer the question and continue indefinitely, switching back and forth in working memory. But we stumble almost immediately.

But now, here is a similar question that is much easier to answer. This time, start the new sequence at 254 and switch back and forth. Then the resulting sequence is 254, 256, 258, 260 …. What is the next number? Of course, the answer is 262, and then 264, and then 266, and so on forever. Everyone finds it very easy to continue this sequence indefinitely, even though we haven't forgotten that the inputs are two separate sequences, the first one being 256, 260, 264 … and the second one being 254, 258, 262 …. Why is it so difficult to run the sequence 256, 254, 260, 258, 264, 262 … but so easy to run the sequence 254, 256, 258, 260 …? The answer is not that we have already heard one but not the other. It is not that in one case we are reciting from long-term memory but in the other case we are not. The answer is not that they are put together from different inputs. Again, a computer would not have the slightest difficulty running either of these sequences, and indeed a mathematical ranking of the two sequences would assign them equal computational complexity. How does our mind work so that running the two sequences feels so different?

Everyone knows the answer immediately. In both cases, we have the same two input sequences, and in both cases we have the same sequence rule: Take numbers sequentially in alternation between the two input sequences. Or at least, we have the same sequence rule except for where we start. The computer would easily generate the needed parts of the two sequences and just alternate between them. But that is not how we do it.

For a person, there is a big difference between the two tasks. Starting with one of the two inputs makes it very difficult to keep going, and starting with the other makes it very easy, because in the second case there is a single, compressed blend, namely a single, unitary integer sequence defined by a rule: Start at 254 and keep adding 2 to the last number. We can use the tight blend to do our thinking, and expand it to the input sequences when we want to. (Of course, there is a rule for the sequence that starts with the other input, but the rule is more complicated and not everyone finds it, and even when they do they can have difficulty running it.) When we start with 254, there is a congenial, compressed, unitary blend, and running the blend makes immediate sense. The blend serves the web of two different sequences, and lets us keep connected to not only the blend but also the two input sequences we started with. In that case, we run in working memory three things rather than two, but running three things is easier than running two because the third thing is a compact blend that connects to and organizes the web involving the other two. Running the blend lets us run the input sequences by inference. More is easier, if the more is packed in a congenial way. More is easier if the more is a packed blend that lets us grasp and manipulate whatever we were trying to hold in mind. More is easier if the blend lets us generate most of the web on the fly instead of having to hold it explicitly in mind. For working memory, more is better if the more comes about by good blending. Blending changes the task; it leverages working memory. On the fly, blending and working memory are a much stronger tool than working memory alone.

Here is a much simpler visual exercise to demonstrate how blending helps us manage a diffuse mental web by creating a compressed, congenial, mathematical blend that serves the decompressed web and makes it possible for us to grasp it. The mental web we are about to look at is overarching, diffuse, and difficult. But the blend is compact and fun. Imagine a line of identical interlocking gears lying flat on a table that stretches for kilometers. If the first gear turns counterclockwise, which way does the 173rd gear turn? The 256th? It is not so easy to reach the right conclusion for these big numbers. Any two gears separated by only one gear must turn in the same direction. So project all the odd-numbered gears in the infinite sequence to the same gear in the blend, the first gear. And project all the even-numbered gears in the infinite sequence to the same gear in the blend, the second gear. Now, in the blend, instead of an infinite number of gears, there are only two interlocking gears, the first turning counterclockwise and the second turning clockwise. In imagination, all the odd gears in the line are projected onto the first gear, and all the even gears are projected onto the second gear. The blend has only two gears, and two directions, but it can be expanded to help us deal with any part of the infinite sequence of gears. The blend organizes a mental web much too diffuse to be held in working memory. From the blend, we can manipulate the mental web of the infinite sequence, and even rebuild it. Now we know that the 173rd gear must turn just like the first gear, which is to say counterclockwise. Using the blend, we know that the 456,251st gear turns counterclockwise, too. And we know that the 256th gear must turn just like the second gear, which is to say clockwise. We know that the 12,345,678th gear turns clockwise, too. *Blending leverages working memory.* Blending makes working memory more powerful. It even saves us from having to carry the contents of working memory around with us all the time. Because of blending, we do not need to hold the entire mental web in working memory. We can reactivate parts of it on the fly as needed by working from the blend (Fig. 2.1).

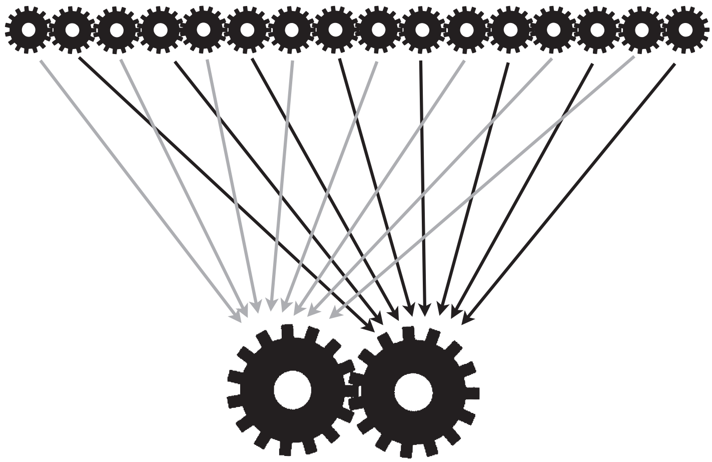


Fig. 2.1 A two-element blend of an indefinitely large mental web

Compression through blending makes it possible for the human mind to go to places it has never gone before. We are living beings, in many ways very much like other living beings. Like plants, we are made of cells. Like mammals, we breathe air. Like primates, we have not only two arms and two legs, but also the common primate brain. In the big picture, what we share with other species is surely the largest part of the human story.

But there are other things that we do not share with other living beings. We do not share photosynthesis with plants. We do not share echolocation with bats. We do not share flight with birds. And they do not seem to share advanced blending with us. Advanced blending provides us with extraordinary flexibility and a unique power for innovation. Mathematical compression and decompression through blending let us, for example, go to sea.

Land is the kind of place we are built to be able to understand. Imagine that something in our field of vision comes straight at us. That is, we look in its direction, keep our gaze fixed, and this object we see stays in the same place in our field of vision, but it keeps getting bigger. We know that we need to get out of the way. That understanding is at human scale; it requires no big thinking. It takes only basic cognition.

Next, suppose that whatever we are looking at on land stays in the same spot in our field of vision but gets smaller, or stays the same size. Put another way, its image subtends an ever-decreasing, or an invariant, angle of our visual field. We do not need to move, because it is either moving away from us or staying at the same distance from us. This understanding is at human scale; it requires no big thinking. It takes only basic cognition.

Since motion is relative—we, in relation to the thing we are looking at—there are actually two or three ways to understand what is going on when we see these objects getting bigger or getting smaller. In one way, the object is moving while we remain stationary. In the other way, the object remains stationary while we are moving.

There is also the possibility that both we and the object are moving. Any of these three scenes might fit what we see when the object seems to get closer. No matter which of these three scenes you are imagining, suppose that the object is staying in the *same spot* in your field of vision and *getting bigger*. In that case, you know you are going to collide with the object. In any of the three scenes, we have the same little physical story: the story of a collision. We are going to collide with the object.

Landlubbers might find it difficult to imagine how ambiguous and uncertain a boat's location can seem to be when one is on it out at sea. In our time, the problem of locating the boat relative to the Earth is solved by signals from global positioning satellites. But only a few decades ago, GPS was not available, and even now GPS can fail, requiring the sailor to use other methods. There are charts, but it can be a very hard job to establish the relationship of the boat's location to the images on the chart, especially out of sight of land, at night, under overcast skies.

Suppose we see another object out on the water—a boat, a buoy, a board, or a green or red or white light. Is it moving? Will we collide with it? There can be a considerable lag between a pilot's action and a vessel's response. The larger the boat, the greater the lag, for the most part. And there is even more bad news: A great deal is going on while a boat sails along a bearing out at sea. Maybe it is not so easy for the sailor to change bearings. Maybe there are obstacles preventing the sailor from going some of those ways. Maybe the conditions are better over here than over there. It might not be so easy to make a course correction later. What should we do?

The sailor can try to compute in working memory all the relative positions of objects on the water and their futures, but that results in a diffuse, distended mental web, reaching over time, space, causation, and agency. Such a web is very hard to hold and manipulate in working memory. There is a lot at stake. A mistake could be fatal.

This at-sea mental web of possible actions and possible consequences stretches far beyond anything at normal human scale. The collision could be an hour away, or 2 hours away. We have to think about it. How can we get a handle on this problem? First, there is the simplest case, in which something is close, straight ahead of us, and getting bigger fast. We have to get out of the way. Move that tiller or rotate that wheel. But suppose the object is not straight ahead of us, or is getting bigger only slowly. We are not moving directly toward it. It is off the port or starboard side of the boat. Now we have to do some blending.

We can start to think about this process by remembering what we know about basic mammalian movement on land. We are very good at understanding movement along a path—a skill we share with other mammals. We are good at picking out in our visual field something that moves fairly quickly along a path: a bird, a fly, or an ant. Often, the object leaves a trace along the path, as when a child rides a bicycle through the beach sand. Activate this idea of something moving along a path and leaving a trace of its movement. At the same time, activate the idea of a boat moving on the ocean. The movement of the boat will take a long time, and we do not see the trace made during all of that time. But if we blend these two ideas—the relatively quick movement along a path that leaves a trace, and the boat's movement—then, in the blend, the boat has motion along a path, and we can see that path all at once. In the blend, we now have a boat, its entire movement along a bearing from past through present to future, and its trace.

Notice what is happening here: In one of the input mental spaces, we see the entire movement and its trace in a very little bit of time—as when the bicycle goes through the sand—but in the other, the amount of time is immense. When we blend these together, we see in the blend something happen quickly that we know takes a long time. This is a compression of time. We perform such compressions routinely. When we look at the calendar, for example, the whole month is right there; the movement from day to day takes no more time than is required for our finger to go from one spot on the calendar to another. In the blend, the movement of the boat that might take hours takes only a few seconds. That is a time compression. And even though we see no lasting trace on the water, since the wake disappears, in the blend, we have the boat's trace mentally available as something to be used.

We are not deluded: When we expand the blend, we realize that it contains a compressed mental representation of time. In the blend, the extended activity of the boat has been blended with our simple notion of quick movement along a path that leaves a trace. The blend uses something with a basic, at-home structure to let us conceive of a far-from-home mental web of ideas.

Now that we have this blend for our boat's movement along its bearing, we can find a way of dealing with the question of whether we will collide with something else we see out there on the surface of the water. First, make not just one, but two of these boat-on-a-path-with-a-trace blends. In each, there is something moving along a course, both under severe time compression. One blend has our boat. The other has whatever object we are looking at, out there on the surface of the water. In fact, it could be that the "course" of the other object, or even our own course, is to stay motionless in one spot on the water. Either we or the other thing could be anchored, for example, or becalmed. But when we are out at sea, we cannot tell just by looking (this may surprise those who do not go to sea) whether one of us is not moving. So, imagine that what we see is a boat out on the water. Then in each of the two boat-on-a-path-with-a-trace blends, there is a boat and a course for the boat along a bearing, and in each blend there is a static line, a trace.

Now, blend again. Blend those two boat-on-a-path-with-a-trace blends into a hyper-blend, so that in the new blend we have two bearing lines. Do those bearing lines intersect? If so, do the two lines, or, more accurately, the two boats, arrive at the intersection at the same time? If so, then there is a collision coming. Move that tiller or rotate that wheel. But how can we think about whether the two boats arrive at the intersection at the same time? That point of intersection could be a long distance from where we are, and we do not know that distance. It could be a long time from now, and we do not know how long. Here, what we know about the human-scale scene of colliding with something or someone comes to our aid. In our human-scale scene, we know that if we keep our angle of vision just the same as we look at the other object, then the other object must move "forward" in our field of vision if it is going to get to the intersection before we do. That is, it must look as if it is "gaining" on us, shortening the distance to the intersection faster than we are. Alternatively (assuming, again, that we keep our angle of vision just the same as we look at the other object), the other object must move “backward” in our field of vision if we are going to reach the intersection first. That is, the other object must look as if it is “losing” because we are shortening our distance to the intersection faster than the other object. If the other ship is moving “forward” in our field of vision, it is beating us to the intersection. If it is moving “backward” in our field of vision, we are beating it to the intersection. In either case, there is no problem, because the two objects will not arrive at the intersection simultaneously.

But if the other object stays in just the same spot in our field of vision and gets bigger, then we are going to reach the intersection at the same time. We are going to collide.

Baseball fielders use this “gaze heuristic” to catch a fly ball: They run toward where they think the ball might be headed, but keep looking at the ball flying through the air with the same angle of vision, and speed up or slow down so as to keep the ball in the same spot in the field of vision, neither advancing nor falling back. That way, they will intersect with the ball, and perhaps catch it.

By blending what we know from our human-scale experience of movement and collision with the vast at-sea web of objects in the distance, we can understand, in the blend, something that we cannot actually see. In the blend, there are intersecting lines and everything else we need to decide whether to deviate from our course. In the case of the gaze heuristic, it may be that instinct has built the right tool into lots of mammals. But what we are interested in here is the way in which advanced blending can make this kind of idea available for scenes far beyond the local area and moment.

For those who are interested in the mathematics, it can easily be shown that the same blending provides the mathematical understanding of the situation as it is usually taught in navigation classes for sailors. The geometry goes like this: We can mentally draw an imaginary triangle in the blend. One vertex is the position of boat A, the other vertex is the position of boat B, and the third vertex C is where the two bearing lines intersect. If both boats arrive simultaneously at C, then each boat, sailing at a constant speed along an unchanging bearing, traverses its leg of the triangle in the same amount of time. In that case, boat A traverses *x%* of its course to C in the same amount of time that boat B traverses the same *x%* of its course to C, and therefore we have, at each moment, a triangle, ABC, and all these triangles are *similar*: They all have the same, unchanging interior angles, and the three sides always stand in the same proportions. So, for all these triangles, the angle—that is, the bearing—from boat B to boat A stays the same. In other words, all these different static triangles can be packed to a blend in which there is *one triangle* that is *shrinking over time*. This blend is, in visual imagination, something direct and human scale. Of course, it is not something one can see in the visual field. The triangle in the blend changes in a manageable amount of time. It can be unpacked to the entire diffuse web of locations in time for both boats.

Will we collide with the other boat? Now we can see the math that can be applied to the blend: As we stand on our boat, we gaze at the distant object and ensure that our gaze is constant relative to our boat, by, for example, picking a point on our gunwale and looking out over it at the object, and not moving our head or our eyes.

If the distant object remains in the same spot in our gaze and grows larger, we are going to hit it, so we should change course.

The "shrinking triangle" blend for determining whether we will collide uses an extremely common pattern of blending: In a big web, find the analogies and disanalogies across lots of mental spaces; compress the analogies and disanalogies into the blend, with the result that the blend has *one* thing that *changes*. There can be very many conceptual inputs in such a web, with analogies and disanalogies connecting them. The analogies are compressed to a *single* entity—in this case, the *triangle*. And the disanalogies are compressed to *change* for that one thing. In this case, the change for the one thing is that the triangle is *shrinking*. The triangle keeps its proportions as it shrinks, and shrinks down to a single point: the point of collision. So if there is a collision (in the blend), then the bearing from boat B to boat A *stays the same* and boat A gets closer to boat B because the leg of the triangle that connects them is shrinking. This implies that boat A subtends a larger angle in the visual field of the sailor on boat B. This situation is called by sailors *Constant Bearing, Decreasing Range.* The mnemonic is Charter Boaters Detest Returning. To the sailor, it means: If there is a constant bearing to the other object, and the range to the other object is decreasing, alter course or the two of you are going to collide.

Of course, if what you want to do is collide with or meet the other object, then the shrinking triangle blend is exactly what you want.

By using this blending web, the sailor creates a manageable mental scene—the shrinking triangle of doom. It can be easily grasped. The shrinking triangle blend enables the sailor to make a human-scale decision. That decision can be expanded to manage a diffuse mental web that otherwise could not be mastered within working memory. This is new stuff, on a grand scale, and lets us boldly go where we were not adapted to go. Dogs, wonderfully talented as they are, so flexible that they can learn to work on a boat in blue water, seem to be equipped with the gaze heuristic: Some breeds are good at catching fly balls and Frisbees. But never expect a dog looking at a dot on the sea a long way off to start barking to alert you that you and the dot are going to collide.

The shrinking triangle blend lets us think far beyond human scale. It lets us think at web scale, about a future that is only slightly specified. But in all versions of that future, with different triangles, there is a collision if we have constant bearing, decreasing range, and no change of course. The blend organizes and serves the diffuse mental web. We conceive of the blend, and it allows us to grasp the mental web, to reason in the mental web, and to draw inferences for our present action.

The blend delivers new stuff that is amazing, once we think about it. To begin with, there is of course no triangle except in the blend. Imagining a single triangle on the water is already an impressive compression.

Moreover, this conception of the triangle on the water is not actually an idea of a specific triangle. Instead, the imagined triangle is a potential, generalized triangle, with some constraints. Why is it not a specific triangle? Because the sailor does not know the lengths of any of its legs. Therefore, what she knows is not a triangle, but rather a set of constraints on the relevant (but unknown) triangle. There is an uncountable infinity of triangles fitting those constraints, and an uncountable infinity certainly cannot fit inside working memory, but that uncountable infinity of triangles is all compressed into *one* triangle in the blend.

The next compression we achieve, a compression that is quite different from the compression of all the possible static triangles into one static triangle, is the *shrinking*. Over time, there will be an uncountable infinity of such similar triangles. But each of them is static. There is no shrinking in reality. The shrinking is new stuff that arises in the blend. The new stuff in the blend is not itself shared by the ideas that we call into use to make the blend. In none of our ideas of reality is there a shrinking triangle. But we make one in the blend, and it helps us understand reality. The blend originates a new idea that helps us manage the vast mental web.

It is not even given that there is a triangle that results in a collision. The entire blending web for the collision triangle is held as potentially counterfactual with respect to another blending web in which the two lines that are the courses of the boats either do not intersect or have an intersection that is not a simultaneous location for the boats. It is no problem if the courses intersect—this happens all the time. The only problem is if their intersection is a *simultaneous* location for both boats.

Managing the mental web of events on the water depends upon having a blend at human scale, one we can mentally grasp. From the mathematical blend of the shrinking triangle, we can manage the out-at-sea web it serves.

The evolution of an advanced blending capacity and the evolution of memory capacity could have bootstrapped each other in the evolution of human beings, in two different but related ways:

1. An expansion in working memory—where by “working memory” we mean the capacity to hold information in mind for processing—would have made more mental stuff available to the process of advanced blending. So, other things being equal, an expansion in working memory would have been more useful, fitter, if that blending capacity was already strong enough to handle the new load and deal with the range of new material.
2. Long-term memory might have evolved to provide some mental input spaces to the advanced blending mill that are *not compatible* with the present situation. That is, contents of long-term memory might be incompatible with the present situation, so long-term memory could be a great resource for blending if blending can work with incompatible input spaces. The present situation we inhabit has stuff that is pretty much compatible—after all, it is all right here right now together. So where would a capacity for advanced blending that is superb at blending incompatible ideas get the incompatible ideas? One answer is an evolved long-term memory that is freed from submission to the present situation. In cognitive science, a memory incompatible with the present situation is called “decoupled.” The more capacious the power of long-term memory, the greater the range of the conceptual material it can supply to blending.

In the cases of both working memory and long-term memory, we have an evolutionary bootstrap: An expansion of blending capacity makes it fitter for working memory to expand and for long-term memory to expand; and an expansion of

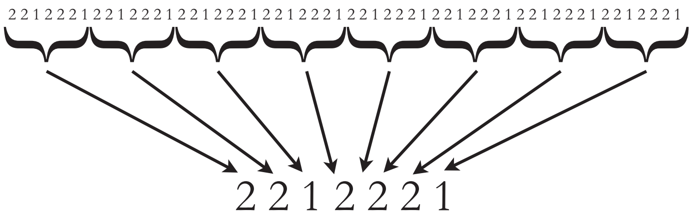


Fig. 2.2 Blending an indefinite mental web to a seven-element cyclic blend

working memory or long-term memory makes it fitter for the blending capacity to expand.

One great difference between our species and all other species is our capacity to manage complex, diffuse mental webs that range far beyond the here and now. Our ability to manage these mental webs depends upon our ability to compress them into congenial, human-scale blends. Here is a snippet of a sequence of integers: 2 2 1 2 2 2 1 2 2 1 2 2 2 1 2 2 1 2 2 2 1 2 2 1 2 2 2 1 …. It can be quite difficult to grasp this sequence, but a first step is to recognize that it is a repetition of 2 2 1 2 2 2 1, like this (Fig. 2.2):

If one memorizes "2 2 1 2 2 2 1" in an auditory loop, it might be possible to write out a lot of this sequence by repeating it and writing as one repeats.

Musicians may recognize that this sequence defines the major diatonic scale: Its numbers give the number of steps (semitones) between notes in the major diatonic scale, beginning on any note. For example, the major scale beginning on C is C D E F G A B C D E F G A B…. (Fig. 2.3)

The number of semitones between C and D is 2; between D and E is 2; between E and F is 1; and so on. This is a spectacular mental compression not just of the major diatonic scale beginning on C, but of all major diatonic scales, regardless of the beginning note. For example, the major scale beginning on G is G A B C D E F# G A B C D E F # ….

Of course, this major diatonic scale has the same repeating pattern as any other: Beginning at G, take two steps to A, two steps to B, one step to C, two steps to D, two steps to E, two steps to $F^{\#}$, and one step to G, and then repeat indefinitely.

Children studying music chant, "Whole Whole Half Whole Whole Whole Half," which is another way of saying "2 2 1 2 2 2 1." This chant triggers a compressed blend that can be expanded to help us understand the major diatonic scale, *any* major diatonic scale.

On the piano keyboard, the sharp and flat notes are the black keys. The key of C major has no sharps or flats—so just white keys. If you start at C, you have to make two steps (first to the black key, second to the white key) to reach D. Then two steps to E, then only one step to F, then two steps to G, then two steps to A, then two steps to B, and then only one step to C: 2 2 1 2 2 2 1, or "Whole Whole Half Whole Whole Whole Half." If you find the keyboard representation confusing, look at the neck of a guitar. Play a C, then move up 2 frets, play a D, and so on up the scale.

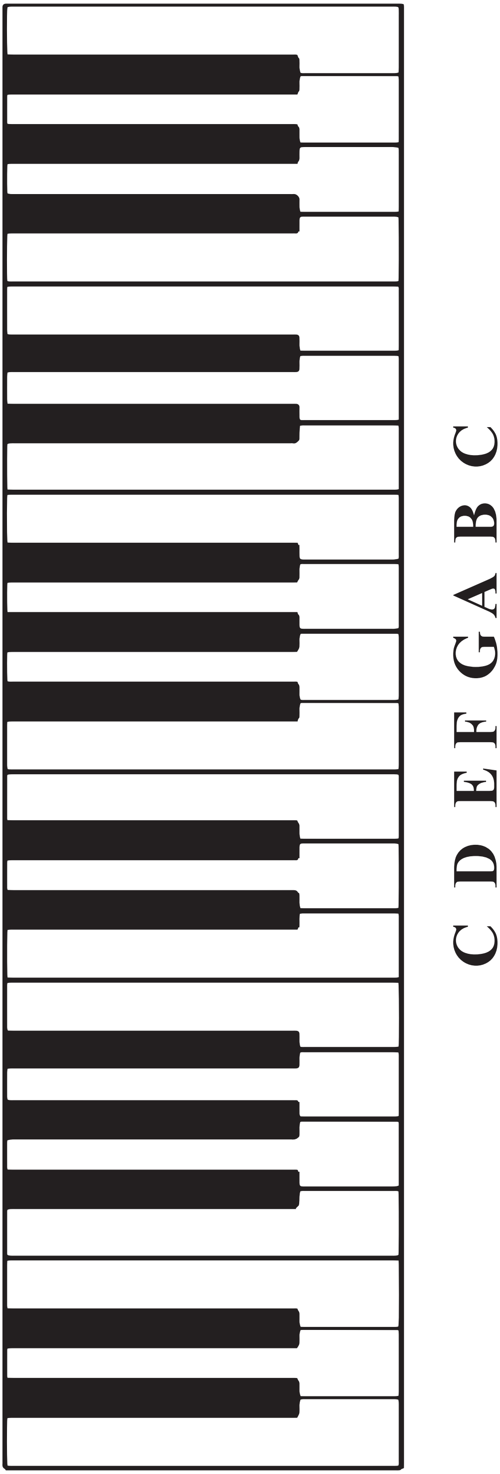


Fig. 2.3 C major scale on a keyboard

But the 2 2 1 2 2 2 1 recurring sequence is good for *any* major scale. If we blend it with a specific beginning note (like C), then the new blend is a more specific compressed blend—compressed not least because it has only seven elements in a repeating sequence, whereas the actual scale itself goes on indefinitely in either direction. This new specific compressed blend can be expanded to give the entire major diatonic scale beginning on that note.

Inspection shows that the sequence that consists of repetitions of the period 2 2 1 2 2 2 1 is the same as the sequence that consists of repetitions of the period 2 1 2 2 1 2 2. And if one starts at the beginning of the period 2 1 2 2 1 2 2, one has the natural minor scale. Accordingly, every major scale has a relative minor scale that has the identical notes, but that begins on the sixth note of the major scale. The C major scale is C D E F G A B C D E F G A B … and its sixth note is A, so the A minor scale has the identical notes: A B C D E F G A B C D E F G …. The difference is only where one starts. Everyone hears and feels that musical difference immediately. Both scales have two steps between their first two notes, but where the major scale has two steps between its second and third notes, the natural minor scale has only one step between its second and third notes.

There is a further compression scheme that creates an effective blend: All the structure of all the major diatonic scales and their relations to all the natural minor diatonic scales can be conceived of as an expansion from a compressed blend known as the "circle of fifths" (Fig. 2.4).

A circle is very much at human scale, and the idea that you go from one spot on the circle to the next spot in the circle by repeating the identical operation is also at human scale. Of course, in music, the keys are not actually arranged in any physical circle, and one does not actually move from one physical location in the circle to another when one "changes keys," but the blend can recruit both the idea of the circle with steps and the idea of moving from one spot of the circle to another.

Teachers of music create further compressed blends to help students reconstruct this blend. "Fat Cats Go Down Alleyways Eating Bread" gives, in the first letters of its words, enough of the structure of the circle to get the student rolling in generating the rest: F C G D A E B. If one remembers the mnemonic phrase, and remembers that the relative minor scale starts on the sixth note of the major scale, one can generate all the relative minors. There are even more powerful mnemonic blends: "BEAD Girls Can't Fight BEAD Girls" provides the sequence of all the major scales all the way around: B E A D G C F B♭ E♭ A♭ D♭ G♭.

A professional musician with a solid formation no longer needs to do all this expansion from the circle of fifths on the fly, because so much has been entrenched in the musician's long-term memory and muscle memory that the musician can call up any scale on its own. But even the professional musician was once unable to call to mind, much less hold in mind, the entire structure of tonality so as to work within it and manipulate it, and needed these compressed blends to get through the music. Blending and memory complement each other in human thought.

It might seem at first blush as if creating the compressed blend and adding it to the mental web for the scale would only increase the mental load and so make it even more difficult to work with the diffuse mental web. On the contrary, the

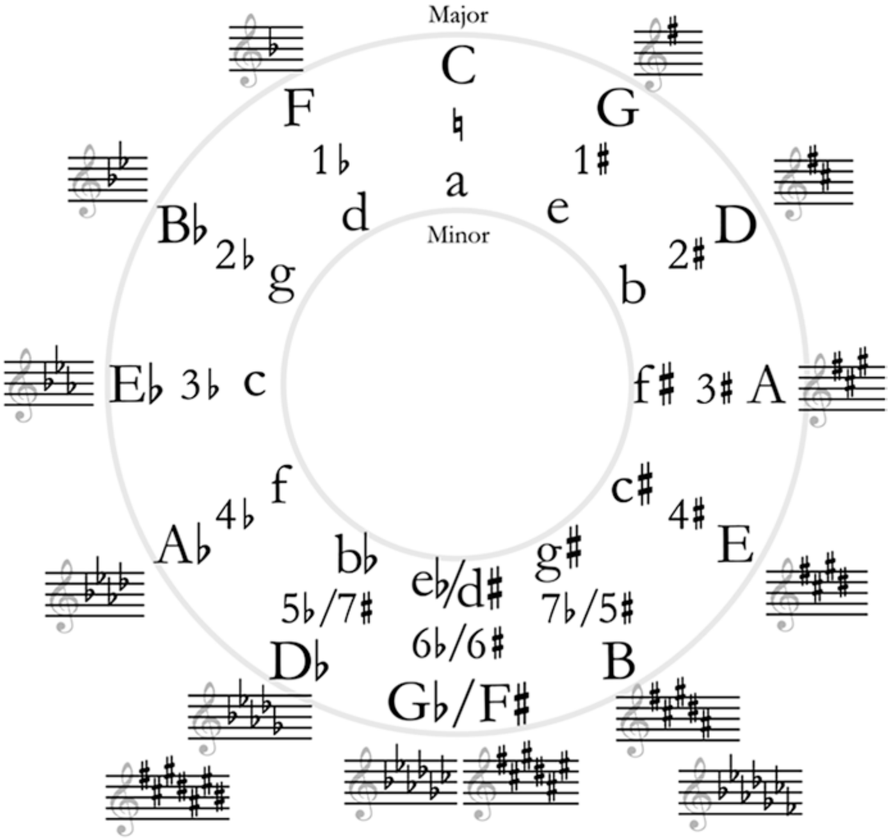


Fig. 2.4 The circle of fifths

osity are part of the human endowment and have been here for a long time, evolutionarily. The main example is subitizing—our ability to make rapid, accurate, and confident judgments of the quantity of items in an array up to three or four. Is the mental number line something like subitizing, or is it a relatively new idea in our descent that originated by blending? Rafael Núñez (2011) argues that the neurobiological and psychological evidence for an abstract, hard-wired, innate mental number line is weak, that the hypothesis of an innate mental number line is implausible, and that Old Babylonian mathematics had no number line. Núñez writes:

> Explicit characterizations of the number line seem to have emerged in Europe as late as the 17th century, and only in the minds of a few pioneering mathematicians. It was apparently John Wallis in 1685 who, for the first time, introduced the concept of number line in his *Treatise of Algebra.* Earlier precursors may have paved the way, such as John Napier with his 1616 diagrams used to define the concept of logarithm. The number line mapping, however, was not a common idea among mathematicians …

> It is important to point out that Wallis's and Napier's texts, intended for readers with advanced knowledge in mathematics, proceed with detailed and careful—almost redundant—explanations of how to generate and use a number line mapping. These explanations are not "formalizations" of the idea of a number line, but rather, they are elaborated presentations of a new meaningful and fruitful idea. The hand-holding narrative, however, is similar to what we see in many elementary school classrooms today, showing just how unfamiliar the idea of a number line was to 17th century mathematicians, let alone to the rest of the majority of illiterate citizens in Europe at that time. Taken together, these facts from the history of mathematics—from Old Babylonia to 17th Century Europe—are simply at odds with the idea of a hard-wired MNL [mental number line] that would spontaneously manifest in *all* humans.

Can this be? Can it be that a concept that seems so natural as the number line is actually a relatively recent achievement of cultural innovation, through blending? Núñez argues that there are human beings alive today in remote indigenous groups who do not have a mental number line: "Uneducated Mundurukú adults dramatically failed to map even the simplest numerosity patterns—one, two, and three—with a line segment, and a high proportion of them only used the segment's endpoints, failing to use the full extent of the response continuum" (Núñez 2011: 655–656).

How we grasp the number line is evidently an open question in cognitive science. I review it here to emphasize that our intuitions about the origins of a very clear idea—such as the idea of the number line—might be very far off base. Although the mental number line seems to us to be inevitable and inescapable, perhaps it originates in cultural time through blending.

Let us take one step further in looking at the ways in which we blend number and motion along a line. In particular, let us look at something that everybody knows arose only very recently, inventively, and among a select few thinkers: the concept of *number* as a *limit*. Brilliant high school students, for example, are often stumped, and argue, touchingly, about whether the infinite decimal .9999 … is a number, and if so which number.

We can think of .9999 … as a number by putting together a particular mental blending web. Imagine a conceptual web consisting, potentially, of an infinity of numbers, each with one more decimal place: .999, .9999, .99999, .999999 …. These are different finite decimal numbers, and there is an infinity of them. How shall we make sense of this web of numbers? It is obviously much too big for working memory to handle by listing all the elements and remembering individually their order. We must do something to compress all this stuff into a tight idea. There are analogies across all the numbers, and disanalogies across all of them, too. If we compress the analogies to a unique element—a point—and compress all the disanalogies to *change* for that element—so the unique point *moves*—then we have in the blend a point that keeps hopping toward the integer 1 but never goes past it. This compressed blend has one entity, a number point on a line, and that entity is changing—it is moving along the number line toward 1. Now working memory is adequate to grasp what is going on. Working memory has now been provided with something that is compressed, manageable, familiar, at human scale, and congenial to the human mind. Working memory can now use that blend as a platform from which to grasp, manipulate, and work on the full web, a little at a time. The job becomes tractable. It is like stacking all the wine bottles in a nice rack.

In the blend, we can now ask about *the point* that is *moving toward* a *fixed point*. Does it grow ever closer to that point and never go beyond it? If so, then we can think of it in the compressed blend as *approaching a limit.*

Advanced mathematics provides much more sophisticated tools for measuring whether something approaches a limit, but in this case we do not need those tools. In this case, we have the very simple compact blend in which each additional decimal place *advances the moving point closer to the fixed number point 1*, and does so for an infinite number of steps. Because, in the blend, the infinite decimal *approaches a limit* that we already take to be a number, we can blend again to create an even greater compression: the infinite decimal can be fused with the limit it approaches. Then, in the blend, we can stipulate that .9999 … is indeed a number, and we know exactly which number it is: .9999 … = 1. In the blend, the infinite decimal is fused with the limit it approaches.

High school students confronted with such analyses sometimes feel that the analysis is just an arbitrary trick, a rabbit out of a hat. The effective but incomplete answer to the high school student is as follows: "Well, if you think .9999 … is less than 1, how much less than 1 do you think it is?" But the more fundamental explanation we should offer to the resisting high school student is that in the discipline of mathematics we have chosen to call .9999 … a number and to fuse it in the blend with the limit "it" "approaches" because such fusions produce a mathematical system that is truly useful both in theory and in practice. Blending is the origin of the idea that .999 … = 1.

We take one last step in the mathematical blending that blends motion and number. This one last step is on the same path, but is known only to those who have studied calculus. Riemann sums and Riemann integrals provide examples of the ways in which blending to a compressed mental space helps us invent mathematical structures, operations, and knowledge. The blending that produces Riemann sums and integrals is clearly a matter of innovation, not of the genetic human endowment. Indeed, very few human beings alive today have the idea of a Riemann integral. Alexander (2011) argues that mathematics as a formal system routinely deploys blending, in an iterative manner, to develop the rich structures of "higher" mathematics, and that mathematics has developed strong controls on the use of blending so as to maintain the rigor of the innovations. His central point is that blending and other such mechanisms are incorporated into the formal structure of the discipline of mathematics.

So let us take a simple look at Riemann sums and integrals as an example of the origin of new mathematical ideas by blending. Riemann sums are sums of the areas of rectangles—that's all. Take a curve in the Cartesian plane, like the one below. What is the area under the curve between two given points on the *x*-axis? We can approximate that area under the curve by fitting adjacent rectangles to the curve, where one side of each rectangle lies on the *x*-axis and each rectangle has the same width. The width is the "domain" of the rectangle. The height of the rectangle can be taken in any of several usual ways: The height can be the height on the leftmost point in the domain, the height on the rightmost point in the domain, the average of those two heights, the maximum height over the domain, or the minimum height over the domain. For our purposes, it doesn't matter. In the figure below, we choose the height of the rectangle to be the height on the leftmost point in the domain (Fig. 2.5).

When we divide the domain of the function into n equal subintervals like this, we have specific values for the endpoints of the subintervals, and a specific, invariant value for the width. All n rectangles under the curve have the same width. The formula for the summed area of the rectangles is then as given in Fig. 2.6.

This just says that the Riemann sum (S) is the sum of all the rectangles, computed as follows. We divide the horizontal distance on the x-axis under the curve into n equal widths. This gives us $n + 1$ equally spaced points along the x-axis: $x_0, x_1 \ldots x_n$. For the first rectangle, the one on the left, take its height, which is just the value of the function at x_0. We write that $f(x_0)$. Now take its width, which is just $x_1 - x_0$. Because the area of a rectangle is the height multiplied by the width, the area of the rectangle is the product $f(x_0)(x_1 - x_0)$. Now move to the rectangle that sits between x_1 and x_2, and take its area, and all the way down to the rectangle that sits between x_{n-1} and x_n. Now add up all the areas. That's all there is to it. More technically, the invariant width of the rectangles is $(x_{i+1} - x_i)$. So the area of the rectangle sitting on the subinterval that starts at x_i and ends at x_{i+1} is just the width times the height of the rectangle, namely, the product of $(x_{i+1} - x_i)$ and $f(x_i)$. The sum of all these rectangles is the Riemann sum, and it is our approximation of the area under the curve.

To obtain a more accurate approximation, we increase the number of subintervals, that is, we make n larger, thereby shortening the width of the rectangles. When we do so, we have a second Riemann sum. But then we can repeat that process again and again. We can increase the number of the rectangles an infinite number of times, just by making n larger each time, each time narrowing the width of the rectangles.

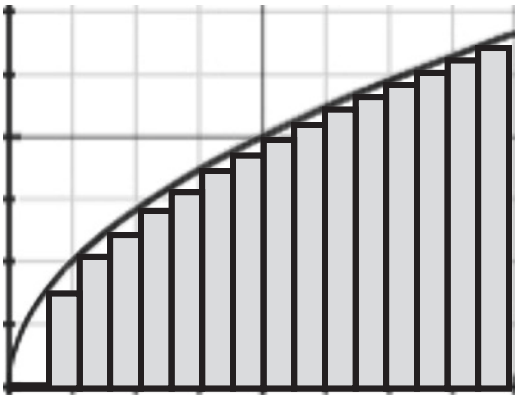


Fig. 2.5 A Riemann sum = the sum of the areas of rectangles under the curve

Fig. 2.6 Formula for a Riemann sum

$$S = \sum_{a}^{b} f(x_i)(x_i - x_{i-1})$$

As we increase the number of rectangles so that it goes beyond just a few, we get a mental web of lots of rectangles much too big to manage in our thinking.

But now, we use standard mathematical blending, including the compression it provides. In fact, we use once again the very common general blending template we have already seen: Blend the analogies together into an identity and the disanalogies together into a change for that identity. In this case, take all those Riemann sums and compress them to *one* Riemann sum in the blend. In that way the analogies across all these input Riemann sums are compressed to *uniqueness*: We have one Riemann sum in the blend. And the disanalogies across all those input Riemann sums are compressed to *change* for that one unique entity in the blend: The Riemann sum changes; it *approaches a limit.* Note that we say "the Riemann sum": language here marks that we have created a compressed blend, and we can refer directly to the compressed blend. When we say "the Riemann sum," no one responds "What do you mean, '*the*' Riemann sum? Every time you increase n by 1, you have a new and different Riemann sum, with a different number of rectangles, and probably a different value. Which one do you mean?" Instead, we know that "*the* Riemann sum" refers to the compressed entity in the blend, the one that "changes," the one that "approaches a limit" by "moving" along "the number line." We decompress this "one" Riemann sum in the blend to indefinitely many in the inputs. Analogy and disanalogy across the web are compressed to the blend, where we now have a unique element that changes. This general blending template is a strong tool of cognition, used widely, no matter what we are thinking about. That is why it is available to higher mathematics in the first place.

The blend has new stuff, that is, something that cannot be projected to it from any of the input mental spaces in the web. A blend is almost never just a cut-and-paste reassembly of elements from the inputs. We run and develop the blend mentally, creating new stuff in the blend. In this case, the crucial new stuff in the blend is the *limit*. This limit becomes a new mathematical structure: a Riemann *integral*. Blending provides us with the origin of this idea. If a and b are the endpoints of the domain for which we want to measure the area under the curve defined by the function f, we write the Riemann integral, like that in Fig. 2.7.

We say that "in the limit," we get the exact area under the curve. This blending approach generalizes over any number of dimensions. We are already getting well beyond the general mathematical knowledge of even educated people, so we will stop with the Riemann integral. But mathematicians will be instantly able to rattle off hundreds of examples of such new stuff from blending, in algebra, geometry, analysis, set theory, and logic.

Blending is flexible, systematic, and principled. It is our essential mental tool of compression. But it can produce many different blending webs. They do not separate into just a few kinds. There is no taxonomy or partition of the products of blending.

Fig. 2.7 Riemann integral

$$\int_a^b f(x)\ dx$$

Still, certain general patterns of blending arise so often that they have been given names. They are reference points in the theory of blending that stand out from the crowd. If we want to emphasize that one of the input mental spaces to a blend is already a blend, which is often the case, we call the resulting blend a *hyper-blend.* If we want to emphasize that what is being blended are a common mental frame and a mental space that has exactly the kind of stuff to which the frame is built to apply, we call it a *simplex* web. As an example of a simplex web, consider the statement "Paul is the father of Sally." Obviously, the kinship frame is built to apply to people. In the blend, Paul is blended with father and Sally with daughter.

If we want to emphasize that the input mental spaces to the blend all share a mental frame, or more generally share the same organizing structure, we call it a *mirror* web. The name comes from the loose idea that the input spaces all *mirror* each other in their main organization. When we want to emphasize that the analogy and disanalogy relations across the input mental spaces are blended to *change* for a *unique* element in the blend, we call it a *change* web.

Some names for blending webs are more specific. If we want to emphasize that agents who do not interact in the input mental spaces are blended to interact in the blend, we call it a *fictive interaction* web (Pagán Cánovas and Turner 2016). An example of a fuller fictive interaction web, a *fictive communication* web, would be the web in which the woman has a conversation with her younger self. A blending web that creates something in the blend that *repeats* is called a *cycle* web. When two input mental spaces have strong conflicts in their organizing structure but one of them controls the organization of the blend, we have used the name *single-scope* web. This comes from the loose notion that in such a web, one is "looking" mentally mainly through one of the input mental spaces. That input mental space is a lens on the organizing structure of the blend. But *single-scope* webs very quickly and easily become what are called *advanced* blending webs, in which both of the organizing structures of the input mental spaces contribute to the organization of the blend, and the blend has new stuff of its own. *Advanced* blending webs have also been called *double-scope* or *vortex* webs.

There are particular general blending patterns that have achieved strong status in one culture or another. For example, Pagán Cánovas (2011) has shown, in "The Genesis of the Arrows of Love: Diachronic Conceptual Integration in Greek Mythology," an article notable for its sensitivity to the role of historical context, how general blending templates underlie a new idea in Greek mythology. He writes, "No symbol from ancient Greek culture seems to have been more successful than the arrows of love." There is a very common general blending pattern, the *Event-Action* blending pattern, in which we blend an event with an action—the action being one that would have led to the event. The result is that something from the input mental space for the event becomes, in the blend, an actor performing an action that leads to the event. We say, "Time is the best doctor." *Time* is causally related to the event of healing. In the blend, Time becomes an actor, a person, or a doctor, who performs an action that is causally related to the event of healing. This *Event-Action* blending pattern is at work in Death, the Grim Reaper. Death, the general cause of a category of events, becomes an actor, a person, or a reaper. Pagán

Cánovas explains that the Eros, the Archer blend, in which Eros shoots someone with an arrow to cause love, is another example of this Grim Reaper pattern. He calls this general blending pattern *Abstract Cause Personification*. He finds another general blending pattern in classical antiquity—the *Erotic Emission* blending pattern. He locates another, specific blending web: Apollo the Archer, or, as he quips, "Death the Grim Archer." His analysis shows how the general blending templates *Abstract Cause Personification* and *Erotic Emission*, the Greek archaic idea of love as a punishment, and the idea of Apollo the Archer all blend to create the arrows of love. He writes: "A process of conceptual integration, taking place probably through several centuries of Greek culture, shaped and refined the religious symbol … This magnificent blend … achieves human scale by compressing the multiple causes, effects, and participants of the erotic experience into a clear story of divine emission" (2011: 573–574).

Many such patterns have been located and analyzed, but it is important to remember that they overlap and can be used simultaneously and that blends constantly arise that do not fit into any of these particular boxes. Blending is an operation with principles and constraints, and it creates a great variety of blending webs. It produces compressions. Blending and compressions turn out to be fundamental tools for mathematics.

Since antiquity, it has been recognized that the human body and brain are small, local, and limited. So is working memory, for no matter how capacious our working memory, human thought outstrips it very quickly, requiring us to find some way to transform what we want to think about into something that can be managed within the limits of working memory.

It has also long been recognized that one of the great open scientific questions—perhaps the greatest—is how people are able to transcend the limits of the body and the brain to achieve immense conceptual sweep, to attain a scope of thought so expansive that many observers have taken it as evidence of our connection to divinity. Philo of Alexandria (c. 20 BCE–40 CE) wrote:

> How, then, is it natural that the human intellect, being as scanty as it is, and enclosed in no very ample space, in some membrane, or in the heart (truly very narrow bounds), should be able to embrace the vastness of the heaven and of the world, great as it is, if there were not in it some portion of a divine and happy soul, which cannot be separated from it? For nothing which belongs to the divinity can be cut off from it so as to be separated from it, but it is only extended. On which account the being which has had imparted to it a share of the perfection which is in the universe, when it arrives at a proper comprehension of the world, is extended in width simultaneously with the boundaries of the universe, and is incapable of being broken or divided; for its power is ductile and capable of extension (Philo 1854–1890, section 90, pp. 264–265).

Like many others, Philo of Alexandria recognized the daunting scientific problem: A local human brain—which is what he means by "membrane"—in a local human body in a local human place manages to think with vast scope. Human thought runs over times, places, causes, agents, and every other sort of distributed meaning. Philo, again like many others, offers an explanation: Human beings are partly divine, having been touched by divinity. They retain something of this divinity. Since divinity spans everything, we accordingly have a scope of thought that would otherwise lie beyond us. Plato proposed something a little different: The human soul lived and thought under supernatural conditions before birth. What we are doing when we think and learn with such vast scope is just remembering what we knew before birth. We remember, by gists and piths, some of the sweeping knowledge we had before being born.

Another range of proposals has it that this sweeping knowledge is given to us by awesome messengers—muses, oracles, ravens who circle the world, aliens, and ancient astronauts, all of which, of course, are products of blending.

In our scientific age, we have moved away from supernatural and divine explanations of the sort Philo offers. We are the originators of our vast and new ideas, including mathematical ideas, and blending is the tool we mostly use.

References

Alexander, J. (2011). Blending in mathematics. *Semiotica,* 187): 1–48. ISSN (Online) 1613-3692, ISSN (Print) 0037-1998, doi: https://doi.org/10.1515/semi.2011.063.

Dehaene, S., Izard, V., Spelke, E., and Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. *Science* 320*:* 1217–1220.

Núñez, Rafael. (2011). No innate number line in the human brain. *Journal of Cross-Cultural Psychology* 42(4): 651–668.

Pagán Cánovas, C. (2011). The genesis of the arrows of love: diachronic conceptual integration in Greek mythology. *American Journal of Philology* 132(4): 553–579.

Pagán Cánovas, C. and Turner, M. (2016). Generic integration templates for fictive communication. In: E. Pascual and S. Sandler (eds.), *The conversation frame: Forms and functions of fictive interaction*, pp. 45–62. Amsterdam: John Benjamins.

Philo (called Judæus or "of Alexandria"). (1854–1890). On the principle that the worse is accustomed to be always plotting against the better, (*Quod Deterius Potiori Insidiari Soleat),* chapter 7 of Yonge, Charles Duke, translator, *The works of Philo Judæus.* London: H. G. Bohn.

Turner, M. (1991). *Reading minds: The study of English in the age of cognitive science*. Princeton: Princeton University Press.

Turner, M. (2014). *The Origin of Ideas: Blending, Creativity, and the Human Spark.* New York: Oxford University Press.

Weyl, H. (1952). *Symmetry.* Princeton: Princeton University Press.

Chapter 3
How Technology Has Changed What It Means to Think Mathematically

Keith Devlin

Early Mathematics

Assigning a start date to mathematics is an inescapably arbitrary act, as much as anything because there is considerable arbitrariness in declaring which particular activities are or are not counted as being mathematics.

Popular histories typically settle for the early development of counting systems. These are generally thought to have consisted of sticks or bones with tally marks etched into them. (Small piles of pebbles might have predated tally sticks, of course, but they would be impossible to identify confidently as such in an archeological dig.) The earliest tally stick that has been discovered is the Lebombo bone, found in Africa, which dates back to around 44,000 years ago. It has been hypothesized that the (evidently) human-carved tally marks on this bone were an early lunar calendar, since it has 29 tally marks (though it is missing one end, that had broken off, so the actual total could have been higher).

Whether the ability to keep track of sequential events by making tally marks deserves to be called mathematics is debatable. "Pre-numeric numeracy" might be a more appropriate term, though the seeming absurdity of that term does highlight the fact that you can count without having numbers, or even a sense of entities we might today call numbers.

Things become more definitive if you take the inventions of the positive counting numbers, as abstract entities in their own right, as the beginning of mathematics. The most current archeological evidence puts that development as occurring around 8000 years ago, give or take a millennium, in Sumeria (roughly, southern Iraq). Various generations of clay object tally systems led eventually to sophisticated schemas of iconic markings on clay tablets that I (and others) suggest we would today call numerals.

K. Devlin (✉)
Stanford University, Stanford, CA, USA
e-mail: devlin@stanford.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_3

To be sure, such an interpretation imposes a modern lens on a much earlier society, so it requires some justification. I recount the full story—pieced together from the archeological evidence—in my book *The Math Gene* (Devlin 2000: 48–49), but here is the general outline.

Initially, the Sumerians used small clay objects as tokens to represent goods, one shape of token for a jar of oil, another for a bale of wheat, another for a goat, and so on. A person's wealth at any one time was represented by the collection of tokens they had, which were kept in sealed clay envelopes held by a village elder (an early form of banking). When two individuals traded, they would go to the elder who would break open their envelopes and transfer tokens according to the transfer of goods, and then seal their "accounts" in fresh envelopes. In time, to facilitate easy checking of accounts prior to a trade, the elders adopted the practice of pressing each token onto the wet clay before placing it inside and then sealing the contents. The outside of each envelope thus carried markings that recorded the contents. The system thus had three components: actual goods, clay tokens that represented those goods (and hence were in one-to-one correspondence with the goods), and markings impressed on the clay exterior that represented the tokens inside (and hence were in one-to-one correspondence with the tokens).

The next step was the realization that there was no need for the clay tokens. In this case, the clay sheet did not have to be folded into an envelope. All you needed was a sheet of clay and one token of each kind to make the markings on the clay. In today's parlance, those markings would be called numerals (albeit, initially, one kind of numeral for each kind of good). Our present-day abstract numbers came into being as the mental ghosts of the tokens that used to be locked inside clay envelopes.

This perception definitely accords with our current concept of numbers, where numerals represent numbers and numbers count things in the world, though how the Sumerians thought of the process is not knowable to us. What we can say, however, is that if we look back in time to find a practice that accords with our current framework of

NUMERALS represent ABSTRACT NUMBERS represent OBJECTS

then the earliest known example is the Sumerian system of

MARKINGS IN CLAY represent CLAY TOKENS represent OBJECTS

When the clay tokens were eliminated, you arrive at a situation where the two frameworks are essentially the same.

A Sumerian might have said, "There used to be clay tokens in the middle." Today, we might say, "We postulate the existence of abstract entities called numbers in the middle." This modern-day mental shift of regarding the absence of some entity as the presence of some abstraction would surely have made no sense to the Sumerians 8000 years ago. So we cannot claim that the Sumerians had our modern concept of number. But from a functional perspective, that's exactly what they had.

Of course, having (counting) numbers is a far cry from having any form of arithmetic beyond the simple addition and subtraction that was implicit in their earlier manipulations of the clay tokens. So it barely counts as mathematics. Nevertheless, it provides a meaningful time stamp when mathematics first arose and what the first math comprised. (In Devlin 2000, I argue that the brain's *capability* to do mathematics was coevolved with the *capacity* for language, at least 70,000 years ago, but that's not the same as having a mental activity we can classify as mathematics.)

If you look for arithmetic (counting numbers with addition and multiplication) as the earliest genuine mathematics, the best current archeological evidence is the Ishango bone, found in Africa and dating back to around 20,000 BCE. The markings on this tally stick suggest some knowledge of multiplication.

From around 2000 BCE onwards, there is clear evidence of mathematics, with the Egyptians, the Babylonians, the Chinese, the Indians and the Greeks all developing some form of arithmetic, leaving behind multiplication tables inscribed on clay tablets or written on papyrus.

Around the same time, those ancient societies also developed early forms of geometry, extending mathematics from the recognition and study of patterns of number to include also the recognition and study of patterns of shape. In both cases, the driving force for these new ways of thinking was the solution of practical problems: trade and commerce in the case of arithmetic and land apportionment for geometry. (The word "geometry" comes from the Greek *geo metros*, meaning earth measurement.) The focus was primarily on computation—numerical computation and geometric computation, respectively, though in the case of geometry we usually refer to it in terms such as "procedural execution" or "construction" rather than "computation."

Then, starting with Thales of Miletus around 500 BCE, Greek mathematicians introduced the concept of mathematical proof, a process to establish the truth of a particular mathematical assertion, starting with a small collection of precisely stated assertions (called "axioms") and proceeding by the step-by-step application of precisely formulated logical rules of deduction. During the period from around 500 to 300 BCE, Greek mathematicians studied both arithmetic and geometry from this theoretical perspective, culminating in the publication of Euclid's famous work *Elements* around 350 BCE.

This development resulted in a classification of the discipline of mathematics into two broad categories that continues to this day: pure mathematics, where the emphasis is on establishing mathematical truth by means of formal (or at least rigorous) proofs, and applied mathematics, where the goal is to find answers to practical questions, those answers often, but by no means always, being numbers.

While that classification can be useful, it can also be misleading. For one thing, the two categories overlap massively. But more to the point, the distinction obscures the point that, whether or not the goal is to prove a theorem or to obtain an answer to a problem (say, solve an equation to obtain a numerical answer), what the mathematician actually does is compute—in the broader sense of that word mentioned earlier, which includes, in addition to step-by-step numerical calculation, processes such as step-by-step geometrical construction, step-by-step algebraic derivations, and step-by-step construction of a logical proof.

The Growth of Mathematics

By the time the nature of present-day mathematics was (essentially) established by the start of the Current Era, the scope of mathematics had already grown to encompass fractional arithmetic (quotients of counting numbers), integer arithmetic (positive and negative whole numbers), rational arithmetic (positive and negative fractions), real arithmetic (the concept of "real number" coming from measurement rather than counting), and trigonometry (combining geometry and real arithmetic).

In the ensuing two millennia, mathematics continued to expand still further, with new branches of the discipline being developed: algebra, probability theory, differential and integral calculus, mathematical logic, real analysis, complex analysis, differential equations, algebraic number theory, analytic number theory, topology, differential geometry, and more. (Several of these "new" branches had their origins much earlier; for instance, although historians typically ascribe the birth of calculus to Isaac Newton and Gottfried Leibniz in the seventeenth century, some of the key ideas were known to Archimedes around 250 BCE.)

Some of these domains are highly abstract, dealing with mathematical entities well beyond everyday cognitive experience. Nevertheless, regardless of whether the goal was to prove a theorem or calculate (in some manner) an answer, what mathematicians spent the bulk of their time doing was computation—developing and executing procedures of various kinds. Unless you were competent in executing computational procedures, you could not do mathematics. In fact, in the more recent times of systemic education, without mastery of calculation you could not obtain a credential in mathematics.

This dominance of computation was the case throughout the 2000-year development of mathematics up until the 1960s (of which more in due course).

As more and more new branches of mathematics were introduced, it was not just that the objects mathematicians computed on that changed; there were also changes in the way those objects were represented and in the manner in which the computations were carried out.

The most familiar new representation, and arguably by far the most significant in terms of broad impact, is the place value, Hindu-Arabic system for representing and computing with positive whole numbers using just ten symbols 0, 1,…, 9. Developed in India in the first few centuries of the Current Era, it was adopted and extended by Arabic- and Persian-speaking traders, who extended the numerical procedural rules (algorithms) for performing arithmetical calculations to include logical procedures. One of those logical procedures they called *al-jabr*, the Arabic term from which we get the modern Western name for that form of procedural reasoning: "algebra."

Today, we associate the word *algebra* with procedural, symbolic manipulation and reasoning, but that association is largely as a result of the invention of the printing press in the fifteenth century. Although the use of abstract symbols is as old as anything we would today call mathematics, until the fifteenth century, when mathematicians wrote up their work to be copied and distributed (on parchment or later paper), they wrote everything in natural language, with the only abstract symbols being numerals and symbols for the operations of basic arithmetic. This was the case for the many mathematics texts written in and around Baghdad in the ninth, tenth, and eleventh centuries, and the even greater number of books written in Italy (in particular) in the thirteenth and fourteenth centuries.

The reason why mathematics was written in prose was to ensure accuracy of any copies made. Books were duplicated by hand copying, by monks in monasteries in the case of the initial copies of a new work, thereafter by readers making their own copies. The most common way to learn mathematics or study a new mathematical technique was to borrow a copy of an appropriate book and slavishly make a copy of the manuscript, without pausing to understand it or work through the written examples. Then, after returning the original, the learner would slowly work through their newly created personal copy, writing symbolic expressions and drawing diagrams in the margins as they did so, in order to assist with their understanding. Since the 1960s, historians working in the archives in Italy have discovered hundreds of fourteenth-century manuscripts that were evidently created in that way.

Clearly, if a book made use of symbolic mathematical expressions, which would likely be unfamiliar to the monk or the learner making the copy, there would be a high likelihood of copying errors. And as anyone learning mathematics quickly discovers, just one symbolic error can cause a beginner significant difficulties. To avoid this, authors of mathematics books spelled out everything in words and numerals. Even the first ever algebra textbook, written by the Persian mathematician al-Khwarizmi in the ninth century, contains no symbolic equations.

With the introduction of the printing press, however, the situation changed dramatically. Because of demand, mathematics texts were among the very first books to be put into print. With printed books, the process of learning mathematics from a text changed from writing symbolic expressions in the margin to help understand the prose as you progressed through the text, to writing prose remarks and short notes in the margin to elucidate the printed symbolic expressions.

In other words, the cognitive challenge of distilling a prose description of a problem and its solution down to the bare structure and logic (going from concrete to abstract) changed to be the very opposite: taking a symbolic representation of a problem and its solution and creating a mental image—turning the symbols into a story (going from abstract to concrete).

The ability to accurately reproduce symbolic mathematical expressions—and diagrams—that came with the printing press not only changed mathematics learning, it also greatly accelerated the growth of mathematics. The steady development of new branches of mathematics (the algebra, probability theory, differential and integral calculus, mathematical logic, real analysis, complex analysis, differential equations, algebraic number theory, analytic number theory, topology, differential geometry I listed earlier, and others) involved an overall increase in abstraction.

For example, arithmetic and geometry begin with the abstraction of patterns in the world (number and shape, respectively); number theory studies patterns of numbers (patterns of mathematical abstractions); algebra (high school algebra, that is) looks at patterns of arithmetic (patterns across mathematical procedures); and so on.

Such is the complexity and the degree of abstraction of the majority of mathematical patterns studied over the past several centuries that to use anything other than symbolic notation would be prohibitively cumbersome. And so the more recent development of mathematics has involved a steady increase in the use of abstract notations.

The introduction of symbolic mathematics in its modern form is generally credited to the French mathematician Francois Viète in the sixteenth century.

The Nineteenth-Century Mathematical Revolution

During the nineteenth century, mathematicians tackled problems of ever greater complexity, and in so doing they occasionally found that their intuitions were inadequate to guide their work. Counterintuitive (and occasionally paradoxical) results made them realize that some of the methods they had developed to solve important, real-world problems—particularly where calculus was involved—had consequences they could not explain. For instance, the Banach-Tarski theorem says that you can, in principle, take a sphere and cut it up in such a way that you can reassemble it to form two identical spheres each the same size as the original one. Because the mathematics is correct, the Banach-Tarski result had to be accepted as a fact, even though it defies our imagination.

Faced with such "paradoxes," mathematicians had to accept that there are occasions when certainty is achieved only through the mathematics itself. In order to have confidence in discoveries made by way of mathematics, but not verifiable by other means, they had to be sure that the definitions of the mathematical entities and concepts the reasoning depends on are sound and unambiguous, and that the mathematical reasoning itself is correct. To achieve this end, they turned the methods of mathematics inwards, and used them to examine the subject itself.

By the middle of the nineteenth century, this introspection culminated in the adoption of a new and different conception of mathematics, where the primary focus was no longer on performing calculations or computing answers, but formulating and understanding abstract concepts and relationships.

Led by pioneering mathematicians such as Lejeune Dirichlet, Richard Dedekind, Bernhard Riemann, and David Hilbert, there was a shift in emphasis from doing to understanding. Mathematical objects were no longer thought of as given primarily by formulas, but rather as carriers of conceptual properties. Proving something was no longer a matter of transforming terms in accordance with rules—a form of calculation—but a process of logical deduction from concepts. [To be sure, it is possible to view the process of logical deduction as another form of calculation. When you do so you arrive at the branch of mathematics known as formal logic. Indeed, you can do it in using abstract symbols, which results in the subject known as symbolic logic. But this is a side issue for another day.]

In terms of the mechanics of doing mathematics, things did not outwardly appear to have changed; indeed, the entire shift came about as a result of turning those very mechanics inwards onto the abstract entities of mathematics itself. But in the minds of mathematicians, things *had* changed. By the start of the twentieth century, mathematics was primarily about understanding, not calculation.

For example, prior to the nineteenth century, mathematicians were used to the fact that a formula such as $y = x^2 + 3x - 5$ specifies a function that produces a new number y from any given number x. Then Dirichlet said to forget the formula and concentrate on what the function does in terms of input-output behavior. A function, according to Dirichlet, is any rule that produces new numbers from old. The rule does not have to be specified by an algebraic formula. In fact, there's no reason to restrict attention to numbers. A function can be any rule that takes objects of one kind and produces new objects from them. This definition legitimizes functions such as the one defined on real numbers by the rule

$$\text{If } x \text{ is rational, set } f(x) = 0; \text{if } x \text{ is irrational, set } f(x) = 1:$$

For such a function, the notion of "calculating values of the function" makes no sense. It is not possible to graph the function. The questions mathematicians asked about abstract functions, not specified by a formula, focused on their behavior. For example, does the function have the property that when you present it with different starting values it always produces different answers? (This property is called *injectivity*.)

This abstract, conceptual approach was particularly fruitful in the development of the new subject called real analysis—the rigorous underpinnings of calculus, which had been a mathematical Holy Grail since calculus was invented by Isaac Newton and Gottfried Leibniz in the seventeenth century. In real analysis, mathematicians studied the properties of continuity and differentiability of functions as abstract concepts in their own right. French and German mathematicians developed the "epsilon-delta definitions" of continuity and differentiability, that to this day cost each new generation of advanced calculus mathematics students considerable effort to master.

Again, in the 1850s, Riemann defined a complex function by its property of differentiability, rather than a formula, which he regarded as secondary.

The residue classes defined by the Karl Friedrich Gauss were a forerunner of the approach, now standard, whereby a mathematical structure is defined as a set endowed with certain operations, whose behaviors are specified by axioms.

Taking his lead from Gauss, Dedekind examined the new concepts of ring, field, and ideal, each of which was defined as a collection of objects endowed with certain operations.

And so on, continuing to this day.

Like most revolutions, the nineteenth-century shift in focus had its origins in times long before the main protagonists came on the scene. The Greeks had certainly shown an interest in mathematics as a conceptual endeavor, not just calculation; and in the seventeenth century, calculus co-inventor Gottfried Leibniz thought deeply about both approaches. But for the most part, until the nineteenth century, mathematics was viewed primarily as a collection of procedures for solving problems. To twentieth-century (and today's) mathematicians, however, brought up entirely with the postrevolutionary conception of mathematics, what in the nineteenth century was a revolutionary new conception of mathematics is simply taken to be what mathematics is. The revolution may have been quiet, and to a large extent forgotten, but it was complete and far reaching.

Outside the professional mathematical community, however, there were few signs of a revolution at all. For most scientists, engineers, and others who make use of mathematical methods in their daily work, things continued much as before, and that remains the same today. Computation (and getting the right answer) remains just as important as ever, and even more widely used than at any time in history.

Mathematics in the Digital Age

If we view the development of Hindu-Arabic arithmetic as the first revolutionary change in the way mathematics is done, then the second change in mathematics praxis of comparable magnitude would be the introduction of symbolic mathematics in the sixteenth century—facilitated in large part by the introduction of the printing press a century earlier.

I would argue that there has been just one further shift in praxis that qualifies as a major revolution. It began in the 1960s with the introduction of the electronic calculator followed by the graphing calculator, and culminated with the appearance of computer algebra systems (*Mathematica*, *Maple*, and others) running on personal computers, in the late 1980s.

For the entire history of mathematics up until the computer age, you had to be good at calculation to get into mathematics, including (in more recent times) acquiring qualifications in the subject, and you had to be good at calculation in order to do or apply mathematics. [By calculation, I mean the execution of any procedure or algorithm.] Moreover, prior to the digital age, if you developed or used mathematics in your career, almost all your time was spent doing calculations.

That is why most people, even to this day, think that mathematics essentially *is* calculation. Yet it is not, and many mathematicians from the ancient Greeks onwards were aware of the distinction, though even they spent most of their time calculating. But the ready availability of first computers and then electronic calculators in the 1960s removed the need for humans to perform numerical calculations.

Because of the electronic calculator, when I arrived at university to study mathematics in 1965, I did not need to make use of the fluency at arithmetic I had developed through many years of school education. (Indeed, over the ensuing decades my arithmetic prowess gradually lost its edge through under-use.) On the other hand, I did have to spend a great deal of my undergraduate career as a mathematics major mastering a whole range of algorithms and techniques for performing a variety of different kinds of numerical and symbolic calculations, geometric reasoning, algebraic reasoning, and equation solving. I had to. In order to solve many problems, I had to be able to crank the algorithmic and procedural handles. There was no other way. There were no machines to do it for me the way the calculator in my pocket performed arithmetic calculations for me (faster, with virtually no errors, and for far more—and larger—numbers than I could handle in my head or with paper and pencil).

That remained the case for the early part of my career as a mathematician. But then, in October 1987, Steven Wolfram released the first version of his massive computer algebra package *Mathematica*. The name "computer algebra system" was an inherited baggage from early attempts to automate mathematical calculation, which totally under-represents what Wolfram's program can do. Quite simply, it can execute pretty well any mathematical procedure, in any branch of mathematics.

Soon after, Canadian developers released *Maple*, and a number of other products came out that do similar things. These products did for almost all of mathematics what the electronic calculator did for arithmetic: they made it obsolete as a human skill (other than for educational purposes, of which more later).

For the first time in history, being able to perform calculations was no longer a necessary requirement for using mathematics. This highlighted the distinction, always there but invisible to most people, between the routine parts of using mathematics (executing procedures) and the creative parts. (I'll discuss later the uses of systems like *Mathematica* in pure mathematics, i.e., the formulation and proof of theorems.)

For a few years, products like *Mathematica* and *Maple* were used mainly in university departments of mathematics, physics, and engineering. They were expensive and challenging to use, and ran only on upper end personal computers. But with the release of *Wolfram Alpha* in 2009, the power of *Mathematica* became available in a cloud-based application that could be accessed (for free) from any PC, tablet, or smartphone. Moreover, *Wolfram Alpha* had a simple user interface that makes it possible to execute pretty well any mathematical procedure with as much ease as using an electronic calculator.

The simplest way to get a sense of how *Alpha* works is simply access it with a Web browser and explore for a while. The point relevant to this essay is that it makes it possible for people to use mathematics without having expertise in any particular topic or procedure. (I'll come later to exactly what knowledge is required to do this.)

The arrival of *Wolfram Alpha* has changed forever the way people can use mathematics. More than that, it has made it possible for people who cannot (or believe they cannot) execute formal mathematical procedures—for example, solving a quadratic equation, to take a particularly simple case—to make effective use of mathematics. Today, having a mastery of calculation is no longer the price anyone has to pay to use mathematics.

To help people understand what it is like to use mathematics in today's world, I often draw an analogy with the world of music. To be a mathematician in the pre-*Alpha* era was akin to mastering many instruments in an orchestra. You had to master the arithmetic instrument, the geometry instrument, the trigonometry instrument, the algebra instrument, the calculus instrument, and so on. The more mathematical instruments you mastered, the greater your power as a mathematician. But using mathematics today is more akin to being a conductor of the orchestra. To conduct that orchestra well, you have to know what all of those instruments are capable of, and you surely need to gain experience with a number of them, at least one of them fairly well (ideally more than one). But there is no need to be world class in any of them. The instruments are what "do all the work." As conductor, you have to know how and when to make them work together, deciding which one(s) to use for each purpose as you progress through the symphony.

Actually, a symphony orchestra is too big for the analogy to work for any one math problem; it's more like a small ensemble. But you get the picture. And for sure, there are enough different mathematical tools out there that they definitely constitute an orchestra, and a large one at that. Indeed, *Wolfram Alpha* alone is orchestra scaled, since it encompasses all the mathematical methods that are typically taught at universities at undergraduate and graduate levels—and a lot that are not.

Clearly, with mathematics being done that way, the experience of using mathematics is very different than it was throughout the entire previous history of mathematics. And gone is the need to be good at any kind of calculation. Mathematicians today do not need to be able to calculate quickly or accurately; indeed, they never do that. The detailed execution of any formal procedure or algorithm is now done by machines. They do it considerably faster than humans ever could, and they make far fewer errors (essentially none). Moreover, they do it with far bigger data sets. For example, mathematics students of my generation learned how to solve linear equations and handle matrices and determinants for two, three, and maybe four variables, and if required could go beyond that to five or six or so, maybe a bit higher. But today, many optimization problems solved routinely by computer packages have thousands or millions of variables. No human could ever cope with that.

So does any mathematics student have to be able to handle any kind of linear equation, matrix, or determinant, and if so to what extent? Mathematics educators are still assessing the pedagogic implications of the digital revolution in mathematical praxis, but the general consensus for that particular example is that mastery of 2 and 3 variables is sufficient, with the learner being able to get correct answers in the two-variable case and solve three-variable examples without worrying too much if they make slips. Of course, making an error when dealing with a real-world problem can be a big deal, sometimes having catastrophic consequences, but the computer system that actually executes the procedure won't make that mistake. The goal of mathematics teaching today is not execution; it is understanding. The conductor of any orchestra, musical or mathematical, has to have a deep understanding of what each instrument does, what it is capable of, and when and how to make use of it, but mastery of an instrument is not necessary.

Notice that it is not mathematics that has changed in the digital age, though there have been changes in the form of new branches of mathematics that resulted from the growth of computer technology (fractal geometry, for example). That caveat aside, however, what has changed is the way people use it. Since mathematics itself is largely unchanged, to understand a new mathematical result is essentially the same challenge it always was. It is mathematical praxis that has changed. And with that change in praxis has come a change—or rather, there is an emerging process of change—in what it takes to become a mathematician.

Being able to calculate quickly, efficiently, and accurately used to be essential; now it is not required. In place of that skillset (which took most people considerable time and effort to master, with many dropping by the wayside in the process) is a new set of skills. Those new skills are in fact much closer to those in the humanities or the creative arts than most people yet realize (or in some cases are willing to contemplate). [In fact, my personal view is that they are now practically indistinguishable, but that's for future generations to judge. Mathematics, I would argue, is no longer a special case. From the perspective of mathematical cognition, I believe that the modifier "mathematical" is no longer necessary; it's just (human) cognition. What distinguishes mathematical praxis is the *what* to which human cognition is applied. That's all.]

Whatever childhood (or adult) experiences arouse an individual's interest in pursuing mathematics, being able to master the art of calculation (i.e., executing any formal procedure except in the rudimentary form required to gain sufficient understanding) is no longer a prerequisite. To be sure, you have to be intrigued by the very idea of formally specified abstractions and context-free reasoning. Not everyone will see mathematics as having appeal. But then, few among us can see the attraction in everything our fellow humans decide to pursue either. From the human perspective, it's not so much that today's digital mathematical tools have added something to the discipline; rather, they have removed what for many was an obstacle.

Experimental Mathematics

So far, my focus has been on the use of mathematics in the world. That may be unusual in a mathematical commentary (which this essay is), but using mathematics to solve real-world problems is what the vast majority of mathematicians do. Admittedly, in many cases such individuals don't call themselves mathematicians, since that word tends to be reserved for the few who focus on pure mathematics (as I did for the first 20 years of my career). The essence of pure mathematics was captured perfectly by Euclid in his famous geometry and number theory text *Elements*, written around 350 BCE: the formulation and proof of precise statements (theorems) about mathematical abstractions.

By and large, it's fair to say that, for most pure mathematicians, the core activities today are much the same as they have always been. The most important tool remains paper and pencil, or perhaps a blackboard. (Mathematicians overwhelmingly prefer a chalkboard to a white board, for ease of frequent erasing—the outsider's perception that mathematicians hardly ever make mistakes is as far from the truth as could be; pure mathematicians engaged in research make errors all the time. Errors frequently lead to new ideas.)

In fact, paper-and-pencil math was the key even for the famous, first major inroad of computer technology into pure mathematics: Kenneth Appel and Wolfgang Haken's 1976 proof of the four color theorem. (For any map drawn on a plane, four colors suffice to color the regions so no two with a stretch of common border are colored the same.) Their proof was obtained by familiar paper-and-pencil-assisted mental reasoning, with a twist that their argument left them having to check that 1,936 different possible (specific) configurations of adjacent regions (mini-maps) could be so colored.

Had they been faced with just three or four special cases, or maybe even a dozen or so, Appel and Haken would surely have done everything by hand. But almost 2,000 cases was far too big a task. (The problem of finding a coloring for each one was also time consuming; the method was simply to examine all possible combinations of colors and see if one worked.) Instead, they wrote a computer program to go through all those configurations and find (by exhaustive search) an admissible coloring for each one. When, after over a thousand hours of computing (using 1976 technology), the program had generated colorings of all the special mini-maps, the four color theorem, first conjectured 124 years earlier in 1852, was declared proven.

With the Appel and Haken case, the computer was not really doing any of the logical reasoning. The two mathematicians simply outsourced to a computer a mundane task that could have done by hand, were it not for the number of cases involved. (The number of cases necessary to examine was later reduced to 1,476. A later proof by another team required only 633 special configurations to be examined; but that is still too many for a human to do.)

Well, that last paragraph is not entirely true; at least, it's not the whole truth. There is another aspect to the story that should be included. Viewing their proof as a classical mental construction, with a computer being used only to cope with a large amount of data, is valid if you focus only on the final proof. In terms of process, Appel and Haken actually used the computer as an experimental tool to help them arrive at the set of special configurations they used for the final search. That aspect of their work, often overlooked, proved to be an early instance of what is now regarded as a whole new area of mathematical research: experimental mathematics (Borwein and Devlin 2008).

Experimental mathematics is the name generally given to the use of a computer to run computations—sometimes no more than trial-and-error tests—to look for patterns, to identify particular numbers and sequences, and to gather evidence in support of specific mathematical assertions that may themselves arise by computational means, including search.

But the truth is mathematicians have always engaged in experimental activities. Had the early mathematicians in ancient Greece and elsewhere had access to computers, it is likely that the word "experimental" in the phrase "experimental mathematics" would be superfluous; the kinds of activities or processes that make a particular mathematical activity "experimental" would be viewed simply as mathematics.

True, the carefully crafted image of mathematics presented in published papers and textbooks gives no indication of "experiments." Mathematicians' published works consist of precise statements of true facts, established by logical proofs, based upon axioms (which may be, but more often are not, stated in the work). But if you examine the private notebooks of practically any of the mathematical greats, you will find page after page of exploratory calculations, trial-and-error experimentation (symbolic or numeric), guesses formulated, hypotheses examined, and so forth. Famous mathematicians such as Pierre De Fermat, Carl Friedrich Gauss, Leonhard Euler, and Bernhard Riemann spent many hours of their lives carrying out (mental) calculations in order to ascertain "possible truths," many but not all of which they subsequently went on to prove.

Indeed, the experimental part of mathematics is precisely what mathematicians enjoy! As Gauss wrote to his colleague Janos Bolyai in 1808, "It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment."

Gauss was very clearly an "experimental mathematician" of the first order. For example, his numerical analysis—while still a child—of the density of prime numbers led him to formulate what is now known as the prime number theorem, a result not proved conclusively until 1896, more than 100 years after the young genius made his experimental discovery.

It was when mathematicians started using computers to carry out the exploratory work that the massive role played by calculation and experimentation came to the fore. What makes modern experimental mathematics different (as an enterprise) from the classical conception and practice of mathematics is that the experimental process is regarded not as a precursor to a proof, to be relegated to private notebooks and perhaps studied for historical purposes only after a proof has been obtained. Rather, experimentation is viewed as a significant part of the mathematical enterprise in its own right, to be published, considered by others, and (of particular importance) contributing to our overall mathematical knowledge.

In particular, this shift in perception gives an epistemological status to assertions that, while supported by a considerable body of experimental results, have not yet been formally proved, and in some cases may never be proved.

On the other hand it may also happen that an experimental process actually yields a formal proof. For example, if a computation determines that a certain parameter p, known to be an integer, lies between 2.5 and 3.784, that amounts to a rigorous proof that $p = 3$. There have been instances of this. (See Borwein and Devlin (2008) cited earlier.) More common has been when insights gained by an experimental investigation have been sufficient for mathematicians to develop classical proofs. This has happened a number of times with proofs in number theory, theory of minimal surfaces, geometry, and other areas (op. cit.).

In terms of the topic of this essay, mathematical praxis, the acceptance of experimental mathematics as a recognized branch of pure mathematics provides us a clear instance of an area of pure mathematics where praxis has been changed by computer technology.

Among the daily activities of an experimental mathematicians are:

- Symbolic computation using a computer algebra system such as *Mathematica* or *Maple*
- Data visualization methods
- Integer-relation methods, such as the PSLQ algorithm
- High-precision integer and floating-point arithmetic
- High-precision numerical evaluation of integrals and summation of infinite series
- Iterative approximations to continuous functions
- Identification of functions based on graph characteristics

In terms of my earlier orchestra analogy, where mathematicians in the past spent many hours carrying out hand calculations, the pure mathematician working in an experimental fashion today is simply a mathematician who conducts an ensemble of a particular set of computational instruments. In the case of experimental mathematics, the computer revolution certainly changed how some pure mathematicians work; moreover it did so in essentially the same way it did for applied mathematicians.

Mathematics Education

Given the significant change to the way mathematics is done in the world today, how do math educators prepare their students for life in that world?

So far, many don't. By and large, math classes around the world today operate in much the same way they did in medieval times, often using what are essentially the same textbooks, albeit with computers and other digital technologies sometimes playing an auxiliary role. To a large extent, this is because of resistance to change among some teachers, and (often strong) opposition to change from parents and education administrators who are not familiar with the degree to which the mathematical landscape has been transformed.

This was illustrated dramatically in the United States by strong opposition to the Common Core State Standards, released in 2009 to provide guidelines as to what skills were required for today's world. While poor implementation of the standards—by education boards and the developers and suppliers of textbooks and other educational materials—can fairly be blamed for some of the complaints, the pushback went well beyond that, to opposition toward the basic principles of modern mathematics the standards are based on in order to prepare future citizens for life in the modern world. Why was this? What was being missed?

Once again, an everyday analogy might be helpful here. Florence Cathedral, completed in 1436, took 140 years to build. It is universally acknowledged to be one of the world's most beautiful large buildings. So too is Sydney Opera House, completed in 1973. Yet, for all it has comparable size, it took a mere 14 years to build. How was it possible to build Sydney Opera House ten times faster than Florence Cathedral?

After all, the basic principles of large building construction are essentially the same. The laws of physics did not change. Aesthetic principles are broadly the same.

What changed, of course, are the tools available. Late-twentieth-century architects and construction engineers had very different tools at their disposal from their forebears in the thirteenth to fifteenth centuries. With different tools available, they needed very different skillsets. Whereas medieval builders had to do many things by hand—or at the very least using hand tools—modern builders "conduct an orchestra of different construction tools." Different tools require different skillsets.

Analogously, until the final decade of the twentieth century, mathematics educators had to ensure that students graduated with basic number skills and the ability to perform mathematical reasoning using (in particular) those number skills. Using those basic number skills required good calculation skills, with a premium put on speed and accuracy. But with today's digital tools, the need for calculation has been removed. Instead, today's graduate needs to be able to make good, efficient, constructive, and accurate use of the vast array of mathematical tools now available. Different tools require different skillsets.

Being able to use those new mathematical tools does not require training in any particular one of them (which is just as well, since they evolve and change with considerable rapidity). Rather it requires understanding the basic concepts and principles of mathematics that underpin them.

The key word in that last sentence is *understanding*. For example, in the days when only people could perform calculations, it was important that arithmetical algorithms were as efficient as possible. What are nowadays called the "classical algorithms" of arithmetic were developed and honed over many centuries to do just that. The brilliance of the Hindu-Arabic number representation is that it facilitated the creation of such algorithms. It was not necessary to understand numbers in any deep way—indeed, studies showed that few people really understood the place-value system—or to understand how the algorithm works or why it was constructed the way it was. You just had to master the (few) basic rules and apply them carefully; something that practically anyone could achieve, given sufficient repetitive practice. It was, to all intents and purposes, mindless arithmetic—sometimes amusingly rendered with intentional ambiguity as "meaningless arithmetic." [ASIDE: In terms of *computation*, the shift from the use of an abacus (either the European marked board and pebbles or the Chinese beads-on-wires equivalent) to written Hindu-Arabic arithmetic was largely a wash; one learned mechanical procedure was replaced by another. The real *benefit* of the written system—and it was a huge benefit—was that the written algorithms left an audit trail, enabling anyone to check, and perhaps correct, a calculation after it was completed.]

In contrast to mindless arithmetic, the arithmetic algorithms used in the more progressive schools today have been designed to optimize not speed or accuracy, but understanding (of numbers, the place value system, and the basic ideas of arithmetic). Not surprisingly, those algorithms are, from a getting-the-answer perspective, nothing like as efficient as the classical algorithms that twentieth-century students had to master. That's one of the reasons parents and education administrators opposed the Common Core, which supported the use of algorithms optimized for understanding. But in so doing, they were missing the key point. Today, we have ubiquitous, cheap machines that do the calculations for us. In fact, we don't need to buy specialized machines at all. The smartphone we carry around with us serves that purpose by way of cheap, if not cost-free, apps. A crucially important mathematical skill today is the ability to "conduct the orchestra of those calculation tools."

One of the core skillsets that mathematics educators identified is *number sense*. The most frequently cited definition of this notion is due to Gersten and Chard (1999):

> Number sense "refers to a child's fluidity and flexibility with numbers, the sense of what numbers mean and an ability to perform mental mathematics and to look at the world and make comparisons."

Interestingly, the notion of number sense was originally formulated as a way for educators to help students with special needs master the basic arithmetic that, at the time, still dominated mathematics instruction. With (belated) recognition that the need for calculation had faded away with the increasing availability of computational devices, however, educators began to recognize the relevance, and power, of the notion in navigating the world of twenty-first-century mathematics. [It may be instructive to recast Gersten and Chard's definition in terms of music, to see what is required to be a good orchestral conductor.]

How do you acquire that high-level number sense? The answer is the same way people always did: through lots of practice. What is different is that instead of the practice being structured to achieve speed and accuracy, the goal is to produce understanding. That requires reflective practice, not the rote repetition that can, at least in some people, result in fast, accurate computation—albeit not remotely as fast or accurate as a free app on your smartphone!

The change from society's need for calculation skill to the new need of the higher order number sense may seem revolutionary, and indeed it is. But it is at heart just today's iteration of a series of revolutions that have occurred throughout mathematics' history. Other skills that are essential for today's mathematics developer or user are the ability to recognize and construct logically sound arguments (and recognize unsound ones); the ability to make smooth, efficient use of the digital tools that are available (conducting the orchestra); and increasingly the ability to work well in teams. Since mathematics began, mathematicians have calculated and reasoned logically with the basic building blocks of the time. Today's procedures (that have to be executed) turn into tomorrow's basic entities (on which you operate).

For instance, in the ninth century, the Arabic-Persian-speaking traders around Baghdad developed a new, and in many instances more efficient, way to do arithmetic calculations at scale, by using logical reasoning rather than arithmetic. Their new system quickly became known as *al-jabr*, after one of the techniques they developed to solve equations.

Whereas arithmetic operated on numbers, algebra (as we now call it) is a form of calculation that (essentially) operates on classes of numbers. (That's where the variables come in.) When the sixteenth-century French mathematician François Viète introduced symbolic algebra, those classes of numbers were the new building blocks, on which it was possible to study the operations of arithmetic, and more general forms of operations.

In each case, advances in mathematics were introduced to make mathematics more easy to use and to increase its application.

The rise of modern science (starting with Galileo in the seventeenth century) and later the Industrial Revolution in the nineteenth century led to still more impetus to develop new mathematical concepts and techniques, though some of those developments were geared more toward particular groups of professionals.

Calculus provides a good example. In differential calculus, functions are no longer viewed as rules that you execute to yield new numbers from old numbers, but higher level objects on which you operate to produce new functions from old functions, new building blocks on which to reason.

Today, entire computations can be treated as mental building blocks. If and when those computations are run (on a machine), you may end up with a number, a graph, or some other output. But until then, the (human) mathematician reasons about them as entities in their own right. (It does not necessarily feel that way, but functionally that's what is going on.)

To conclude this section, I'll present a simple arithmetical puzzle to illustrate the kind of mathematical thinking processes that today's more progressive mathematics teachers sometimes use to help their students develop. Because of its simplicity, it's easy to miss the key issues, but for all that simplicity it captures the spirit of how today's mathematicians work.

The puzzle is of the kind you often find in cheap puzzle books or on puzzle websites. In this case, however, my goal in presenting it is not for you to get the right answer. Rather, it is for you to solve it *as quickly as you can*—ideally *instantaneously*. The reason is to try to get some insight into what the human brain can do with ease, so that educational emphasis can be put on enhancing the brain's capacity to do mathematics when working in the "orchestral conducting" fashion of today's professionals (rather than wasting time trying to train the brain to perform calculations, which your smartphone app can do much faster and more accurately).

Here then is the puzzle:

PROBLEM: A bat and a ball cost $1.10. The bat costs $1 more than the ball. How much does the ball cost on its own? (There is no special pricing deal.)

How did you do? The most common answer people give *instantly* to this problem is that the ball costs 10¢. That answer is wrong (and many realize that is the case soon after their mind has jumped to that wrong number). What leads many astray is that the problem is carefully worded to run afoul of what under normal circumstances is an excellent strategy. [So if you got it wrong, you probably did so because you are a good thinker with some well-developed problem-solving strategies—problem-solving "heuristics" is the official term, and I'll get to those momentarily. So take heart. You are well placed to do just fine in twenty-first-century mathematical thinking. You simply need to develop your heuristics to the next level.]

Here is, most likely, what your mind did to get to that 10¢ answer. As you read through the problem statement and came to that key phrase "cost more," your mind said, "I will need to subtract." You then took note of the data: those two figures \$1.10 and \$1. So, without hesitation, you subtracted \$1 from \$1.10 (the smaller from the larger, since you knew the answer has to be positive), getting 10¢.

Notice you did not really perform any calculation. The numbers are particularly simple ones. Almost certainly, you retrieved from memory the fact that if you take a dollar from a dollar-ten, you are left with 10¢. You might even have visualized those amounts of money in your hand. Notice too that you understood the mathematical concepts involved. Indeed, that was why the wording of the problem led you astray! What you did is apply a heuristic you have acquired over many financial transactions and most likely a substantial number of arithmetic quiz questions in elementary school. In fact, the timed tests in schools actively encourage such a "pattern recognition" approach. For the simple reason that it is fast and usually works!

We can, therefore, formulate a hypothesis as to why you "solved' the problem the way you did. You had developed a heuristic (identify the arithmetic operation involved and then plug in the data) that is (a) fast, (b) requires no effort, and (c) usually works. This approach is a smart one in that it uses something the human brain is remarkably good at—pattern recognition—and avoids something our minds find difficult and requiring effort to master (namely, arithmetic calculation).

Of course, primed by the context in which I presented this particular problem, you probably expected there to be a catch. So, after letting your mind jump to the 10¢ answer, you likely took a second stab at it (or, if you were anxious about "getting a wrong answer," made this your first solution) by applying an algorithm you had learned at school. Namely, you reasoned as follows:

Let x = cost of bat and y = cost of the ball. Then, we can translate the problem into symbolic form as the two equations: $x + y = 1.10$, $x = y + 1$.
Eliminate x from the two equations by algebra, to give $1.10 - y = y + 1$.
Transform this by algebra to give $0.10 = 2y$.
Thus, dividing both sides by 2, you conclude that y = 5¢.
And this time, you get the correct answer.

You may, in fact, have been able to carry out this procedure in your head. When I was at school, I could do algebraic manipulations far more complicated than this in my head, at speed. But, truth be told, since I started outsourcing arithmetic to machines over three decades ago, I have lost that skill, and now have to write down the equations and solve them on paper. (This is a confirmation, if any were needed, that arithmetic calculations do not come naturally to the human brain. Over the years, as my mental arithmetic skills have declined, my pattern recognition abilities have not diminished, but on the contrary have dramatically improved, as I learned—automatically, through exposure—to recognize evermore fine-grained distinctions.)

Whether or not you can do the calculation in your head, it is of course entirely formulaic and routine. Unlike the first method I looked at (a heuristic that is fast and usually right), this method is an algorithmic procedure, it is slow (much slower than the first method, even when the algebraic reasoning is carried out in your head), but it always works. It is also an approach that can be executed by a machine. True, for such a simple example, it's quicker to do it by hand on the back of an envelope, but as a general rule it makes no sense to waste the time of a human brain following an algorithmic procedure, not least because even with simple examples it is familiarly easy to make a small error that leads to an incorrect answer.

But there is another way to solve the problem. It is typical of the ways professional mathematicians vocalize their solutions when asked to do so. Like the first method we looked at, it is a heuristic, hence instinctive and fast, but unlike the first heuristic method it always works.

This third method requires looking beyond the words, and beyond the symbols in the case of a problem presented symbolically, to the quantities represented. Though I (and likely other mathematicians) don't visualize it quite this way (in my case it is more of a vague sense of size), Fig. 3.1 more or less captures what the pros do.

As we read the problem, we form a mental sense of the two quantities, the cost of the ball on its own and the cost of the bat plus ball, together with the stated relation between them, namely that the latter is $1 more than the former. From that mental image, where we *see* that the $1.10 total consists of three pieces, one of which has size $1 and the other two of which are equal, we simply "read off" the fact that the ball costs 5¢. No calculation, no algorithm. Pure pattern recognition.

This solution is an example of number sense in action, the critical twenty-first-century arithmetic skill I discussed earlier. It's hard to imagine how a computer system could solve the problem that way.

The acclaimed Australian (pure) mathematician Terrence Tao has called those three ways of solving the bat-and-ball problem, respectively, pre-rigorous thinking, rigorous thinking, and post-rigorous thinking. In a post in his blog *What's new* (https://terrytao.wordpress.com), titled "There's more to mathematics than rigour and proofs," in which he introduced those terms, he was discussing the way professional mathematicians solve abstract problems in pure mathematics. The formal, symbolic, rigorous description you see in papers and books comes primarily at the end, he notes, to check that the solution is logically correct, or at various intermediate points to make those checks along the way. But the key thinking is post-rigorous.

In the case of solving real-world problems, the pros almost always turn to technology to handle any algebraic deductions. In contrast, though pure mathematicians

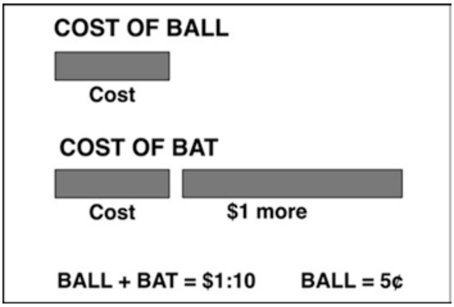


Fig. 3.1 A "professional's" mental representation of the bat-and-ball problem

sometimes do use those technology products as well, they often find it much quicker, and perhaps more fruitful in terms of gaining key insights, to do the algebraic work by hand. But in all cases, they go beyond the numerals and the symbols and reason with the semantic entities those linguistic elements represent.

One of the big questions facing mathematics teachers today is how do we best teach students to be good post-rigorous mathematical thinkers.

In the days when the only way to acquire the ability to use mathematics to solve real-world problems involved mastering a wide range of algorithmic procedures, becoming a mathematical problem solver frequently resulted in becoming a post-rigorous thinker automatically. But with the range of tools available to us today, there is a good reason to assume that, with the right kinds of educational experiences, we can significantly shorten (though almost certainly not eliminate) the learning path from pre-rigorous, through rigorous thinking, to post-rigorous mathematical thinking. The goal is for learners to acquire enough effective heuristics.

To a considerable extent, those heuristics are not about "doing math" in the traditional sense. Rather, they are focused on making efficient and effective use of the many sources of information available to us today. But before anyone throws away their university-level textbooks, it's important to be aware that the intermediate step of mastering some degree of rigorous thinking is probably essential.

Post-rigorous thinking is almost certainly something that emerges from repeated practice at rigorous thinking. (See, for example, Willingham 2010.) Any increased efficiency in the education process will undoubtedly come from teaching the formal methods in a manner optimized for understanding, as opposed to optimized for attaining procedural efficiency, as it was in the days when we had to do everything by hand. Stay tuned!

Figure 3.2 provides a graphical summary of Tao's categorization of the three kinds of mathematical thinking we can bring to problem-solving.

In addition to providing a perspective on the three phases each one of us has to go through to become a proficient mathematical (real-world) problem solver, Tao's classification also provides an excellent summary of three historical stages of mathematical thinking as it has evolved over the past 10,000 years or so, from the invention of numbers in Sumeria, where the mathematical thinking of the time was accessible to all, through three millennia of formal mathematics development, where many people were never able to understand it or make effective use of it, and now into the third phase, where, because of technology, mathematical thinking can once again, I believe, be accessible to all.

As noted above, we do not know the degree to which people have to master rigorous thinking to become good post-rigorous thinkers, but Willingham (2010) and others present evidence to suggest that stage cannot be bypassed. Still, given today's technological toolkit, including search, social media, online resources like Wolfram Alpha and Khan Academy, and a wide array of online courses, it is surely possible to master much of the rigorous thinking you need "on the job," in the course of working on meaningful, and hence motivational and rewarding, real-world problems.

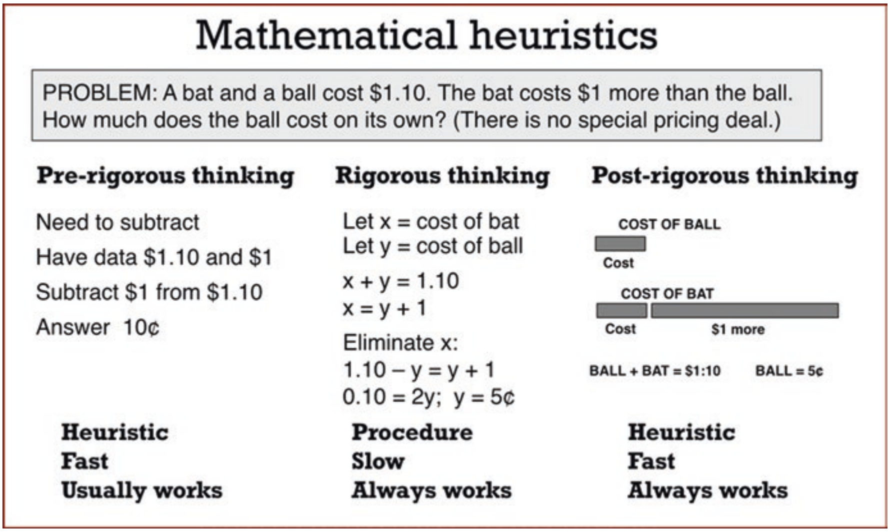


Fig. 3.2 Tao's categorization

This is not to say that there is no further need for teachers. Far from it. Very few people are able to become good mathematical thinkers on their own. Newtons and Ramanujans, both of whom achieved great things with just a few books to guide them, are extremely rare. The vast majority of us need the guidance and feedback of a good teacher.

But, whereas the process of doing mathematics was, until a quarter century ago, dependent on being able to perform calculations of various kinds, a skillset that the brain does not find naturally and requires considerable training and practice, given the readily accessible calculation tools at our disposal, mathematical praxis today consists largely of using the brain in a manner it finds far more natural: analogical reasoning, rather than the logical reasoning previously required.

The Symbol Barrier

Heuristics-driven, post-rigorous thinking is—or at least, should be—the goal of today's mathematics educators, in order for tomorrow's mathematics users to be able to make full and good use of all the available technology tools. Perhaps then, digital technologies themselves can provide new ways to develop (or help develop) those new skillsets. That, in fact, has been the focus of much of my own research over the past few years. The approach I have taken goes back to some groundbreaking social science research conducted almost 30 years ago.

In the early 1990s, three researchers, Terezinha Nunes (at the University of London, UK), Analucia Dias Schliemann, and David William Carraher (both of the

Federal University of Pernambuco in Recife, Brazil) embarked on an anthropological study in the street markets of Recife. With concealed tape recorders, they posed as ordinary market shoppers, seeking out stalls being staffed by young children (between 8 and 14 years of age, it turned out). At each stall, they presented the young stallholder with a transaction designed to test a particular arithmetical skill. The purpose of the research was to compare traditional instruction (which all the young market traders had been receiving in school since the age of 6) with learned practices in context. In many cases, they made purchases that presented the children with problems of considerable complexity.

What they found was that the children got the correct answer 98% of the time. “Obviously, these were not ordinary children,” you might imagine, but you’d be wrong. There was more to the study. Posing as shoppers and recording the transactions was only the first part. About a week after they had “tested” the children at their stalls, the three researchers went back to the subjects and asked each of them to take a pencil-and-paper test that included exactly the same arithmetic problems that had been presented to them in the context of purchases the week before, but expressed in the familiar classroom form, using symbols.

The investigators were careful to give this second test in as nonthreatening a way as possible. It was administered in a one-on-one setting, either at the original location or in the subject’s home, and the questions were presented in written form and verbally. The subjects were provided with paper and pencil, and were asked to write their answer and whatever working they wished to put down. They were also asked to speak their reasoning aloud as they went along. Although the children’s arithmetic had been close to flawless when they were at their market stalls—just over 98% correct despite doing the calculations in their heads, and despite all of the potentially distracting noise and bustle of the street market—when presented with the same problems in the form of a straightforward symbolic arithmetic test, their average score plummeted to a staggeringly low 37%.

The children were absolute number wizards when they were at their market stalls, but virtual dunces when presented with the same arithmetic problems presented in a typical school format. The researchers were so impressed—and intrigued—by the children’s market stall performances that they gave it a special name: they called it “*street mathematics.*”

As you might imagine, when the three scholars published their findings (Nunes et al. 1993) it created a considerable stir. Many other teams of researchers around the world carried out similar investigations, with target groups of adults as well as children, and obtained comparable results. When ordinary people are faced with doing everyday math regularly as part of their everyday lives, they rapidly achieve a high level of proficiency (typically hitting that 98% mark). Yet their performance drops to the 35–40% range when presented with the same problems in symbolic form.

It is simply not the case that ordinary people cannot do everyday math. Rather, they cannot do *symbolic* everyday math. In fact, for most people, it’s not accurate to say that the problems they are presented in paper-and-pencil format are “the same as” the ones they solve fluently in a real-life setting. When you read the transcripts of the ways they solve the problems in the two settings, you realize that they are doing completely different things. (Nunes and her colleagues present some of those transcripts in their book.) Only someone who has mastery of symbolic mathematics can recognize the problems encountered in the two contexts as being "the same."

In my 2011 book *Mathematics Education for a New Era* (2011), I referred to the problem Nunes et al. discovered as the "symbol barrier." Much of my work since that book was published has been to try to develop technological learning tools that set out to break the symbol barrier, by presenting mathematical puzzles (in mathematics education language, they are complex performance tasks) in a manner similar to the kinds of mental representations that arose in my above discussion of post-rigorous thinking for the solution to the bat-and-ball puzzle.

To do that, I contacted some colleagues I had met while consulting for an educational technology company, and together we co-founded a small development studio (subsequently named BrainQuake) to design and build such tools.

Though each of BrainQuake's puzzles (three have been released to date) is built around particular mathematical concepts (integer arithmetic, linear growth, and proportional reasoning, respectively, for the first three puzzles we created), they are not designed to teach or provide practice in the basic skills on which they are built (though engaging with the tools undoubtedly does provide additional practice in those requisite skills). Rather, the goal is to develop number sense and general problem-solving ability.

Because the primary target audience is middle-school mathematics students, the mathematical puzzles we developed are presented as challenges in a video game (called *Wuzzit Trouble*), to maximize engagement, but that aspect is not relevant to this discussion. What is relevant is that they provide an alternative, more learner-friendly interface to mathematical thinking and (multistep) problem-solving than do the traditional symbolic presentations.

The manipulable digital objects in BrainQuake's learning products provide direct representations of mathematical concepts, breaking the symbol barrier. Students (players) solve puzzles entirely within the application itself, by manipulating digital objects, instead of writing and manipulating symbols on a page. The (multistep) solutions students have to develop to solve all but the most elementary puzzles are logically identical to the steps they would carry out to solve the puzzle in classical symbolic form. But the experience of doing so is dramatically different. So much so, that hundreds of thousands of children in the age range of 14–16 have, for instance, successfully solved systems of simultaneous linear equations in up to four unknowns, subject to optimizing their solution to meet various constraints on the solution. See Fig. 3.3.

Figure 3.3 shows two representations of the same problem. On the right is a classical symbolic representation of a problem requiring the student to solve a system of simultaneous linear equations in two unknowns, subject to various constraints. The student is also asked to try to find solutions that are optimal in two ways (parts 2 and 3 to the question). On the left, the same problem is presented as a mechanical puzzle dressed up as a quest to free a caged creature (a *Wuzzit*) from a trap, by rotating, one at a time, two small cogs to turn the large wheel. When the player turns the

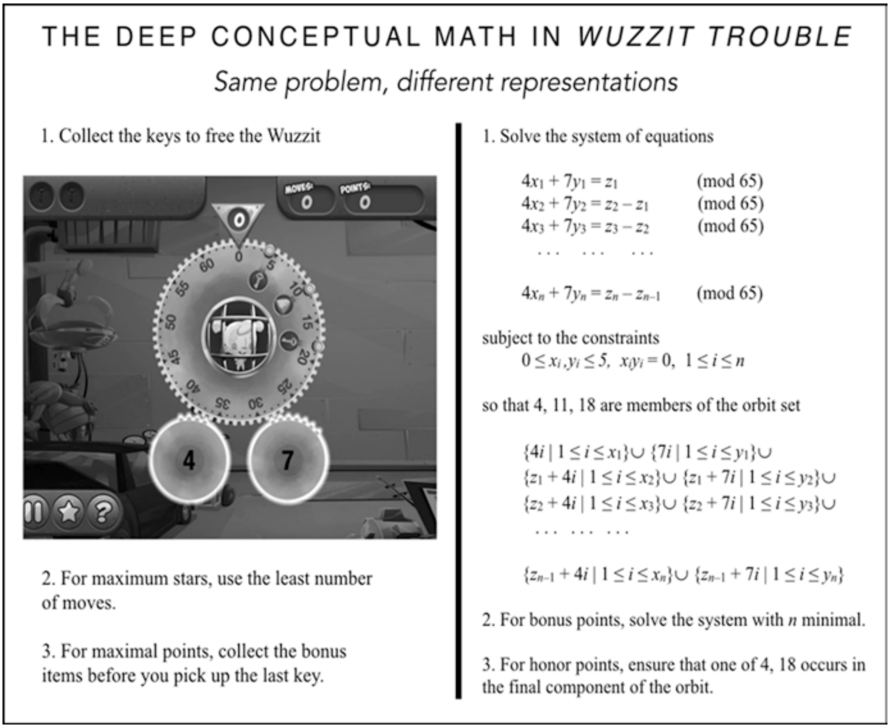


Fig. 3.3 Breaking the symbol barrier

cogs to rotate wheel to bring one or more of the three items to land beneath the origin marker at the top, the player acquires the item. Acquisition of both keys frees the Wuzzit and the puzzle is solved. (The equations have been solved.) Maximum stars are awarded if the player solves it in the fewest possible number of turns (part 2 of the question). Part 3 asks the player to collect the bonus item on the wheel before the last key is acquired.

To be sure, the system of equations on the right is not a standard one. Rather, it is precisely the system of equations that corresponds to solving the puzzle on the left. But the purpose of the puzzle is not to develop the ability to solve systems of symbolic linear equation; the goal is to develop number sense. In this case, the solution of systems of linear equations is simply the mathematical topic chosen as a vehicle to do that. [BrainQuake has produced another version of the puzzle that is stripped of the game features but carries the gears mechanism and the corresponding symbolic equation representations side by side, so the student can see both develop in tandem, thereby explicitly linking the two representations.]

The *Wuzzit Trouble* puzzles have from one to four drive cogs, which means that the mechanism provides a mechanical representation of systems of linear equations in up to four unknowns. See Fig. 3.4.

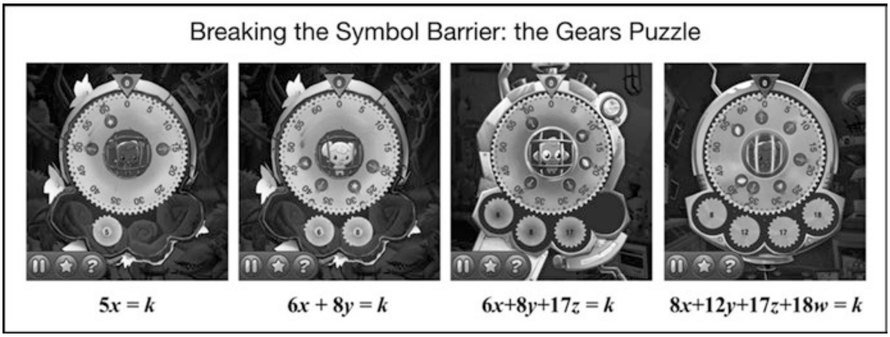


Fig. 3.4 The *Wuzzit Trouble* gears puzzle

Two independent university classroom studies (one in the United States, the other in Finland: Pope and Mangram 2015; Kiili et al. 2015) showed that use of the game *Wuzzit Trouble* for as little as 120 min of play spread over 4 weeks in 10-min bursts at the end of math class produced significant improvements in student number sense, as measured by a written pre- and post-test in the first study, and by both a written test and another digital math game as pre- and post-evaluations in the second. Thus, we know that this approach works.

[BrainQuake is one of a handful of educational technology developers that have adopted this approach. Other products of note are *DragonBox Algebra*, *MotionMath*, and MIND Research Institute's *ST Math*.]

The use of alternative, nonsymbolic representations clearly provides an alternative approach to developing number sense, breaking the symbol barrier that can cause so many problems for learners. Of course, for students who wish to go on to further study or a career in STEM, number sense alone is not sufficient. There remains the problem of leveraging the problem-solving skills acquired in a nonsymbolic fashion to master the traditional symbolic representations, which is a necessary skill for STEM areas. This process is known as "concreteness fading," and has already been studied by others (e.g., Goldstone and Son 2005). It is a special case of education's notorious transfer problem. Technology can help, and as already mentioned BrainQuake has started to develop such tools. But at present this is still work in progress, after completion of which efficacy studies will have to be conducted.

References

Appel, K. and Haken, W. (1976). Every planar map is four-colorable. *Bulletin of the American Mathematical Society*, 82(5): 711–712.

Borwein, J. and Devlin, K. (2008). *The computer as crucible: An introduction to experimental mathematics.* A K Peters/CRC Press

Devlin, K. (2000). *The math gene: How mathematical thinking evolved and why numbers are like gossip*. New York: Basic Books.

Devlin, K. (2011). *Mathematics education for a new era: Video games as a medium for learning*. Boston: A K Peters.

Gersten, R. and Chard, D (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *Journal of Special Education* 33 (1): 18–28.

Goldstone, R. and Son, J. (2005). The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences* 14(1): 69–110.

Kiili, K., Devlin, K., Perttula, A., Tuomi, P. and Lindstedt, A. (2015). Using video games to combine learning and assessment in mathematics education, *International Journal of Serious Games*, October 2015.

Nunes, T., Schliemann, A., and Carraher, D. (1993). *Street mathematics and school mathematics*. Cambridge: Cambridge University Press.

Pope, H. and Mangram, C. (2015). Wuzzit trouble: The influence of a digital math game on student number sense. *International Journal of Serious Games*, October 2015, 5–22.

Willingham, D. (2010). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom*. New York: Jossey-Bass.

Chapter 4
Machine Versus Structure of Language via Statistical Universals

Kumiko Tanaka-Ishii

Introduction

Big data has revealed new facts about natural language, including that human language universally follows certain statistical properties. These properties are observed using methodologies developed in the statistical physics domain, with application to large-scale data. The particularity of the findings is that the properties hold unexceptionally for any texts across languages and time, and even beyond, including infant utterances, music performances, and programming languages.

The crucial problem of these properties is that the reasons why they hold remain unknown. There are mathematical approaches to this problem, as will be summarized later in this chapter, and they suggest a possible common nature underlying any linguistic sequence. These approaches provide explanations about what language could be and what it is not, through models of a linguistic sequence. They consider a linguistic sequence as a string of mathematical symbols, however, irrespective of the meaning conveyed, and therefore they have no grounding of what language is in a humanistic sense. The aim of this chapter is thus to provide a conjecture on the signification of the statistical properties of language from a more humanistic viewpoint, according to previous philosophical conjectures.

The most important understanding we gain from these properties is that linguistic sequences seem likely to be produced through a certain typical behavior, of which we all are unconscious. We produce language, typically trying to be meaningful, but consecutive acts in language production seem to be bound by a predefined mechanism. Because the outcome has a mathematical feature, one way to describe

K. Tanaka-Ishii (✉)
Research Center for Advanced Science and Technology,
University of Tokyo, Tokyo, Japan
e-mail: kumiko@cl.rcast.u-tokyo.ac.jp

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_4

this property underlying linguistic behavior would be the term *machine*, meaning that language production has some aspect of being mechanical. Putting it this way suggests considering the relation of the statistical universals with the philosophical notion of a *machine*, as developed in Deleuze and Guattari (1972) which was preceded by Guattari (1969). Examples of those concepts, however, have mainly been provided via analogy only, and the concepts have remained devoid of concrete forms of what exactly they could be.

This chapter thus attempts to bridge recent scientific findings about language and concepts of modern philosophy. Doing so contributes to grounding the scientific findings in the humanities. The chapter also provides scientific facts that correspond to the philosophical concepts of a structure and a machine. Trying to understand what the statistical universals are would lead us to consider what a machine and a structure could be in relation to language.

Statistical Universals Underlying Human Language

The statistical universals observed in linguistic data appear as power laws, in which the logarithms of two statistical variables are related linearly. As the laws are only visible in log-log space, the properties are only observable with at least a certain large amount of data.

The most famous such property is Zipf's law (Zipf 1965). Figure 4.1 shows a Zipf's law plot of Wittgenstein's *Philosophical Investigations* (PI). For every word kind in a given text, its frequency f is plotted (vertical axis) with respect to its rank r in order of frequency (horizontal axis). Then, $f(r) \propto r^{-1}$, and it becomes apparent that the plot is aligned linearly, with a power-law exponent close to -1 in log-log space. Although this is an approximate result—at the beginning there is a curved region, and then the exponent is almost but not quite -1—this global tendency with an exponent of almost -1 is shared across human texts, without exception. Moreover, this tendency appears in common for language data besides text: in speech, including child-directed speech; in any activities related to human language, such as music; and in computer program code.

Similarly to Zipf's law, the other global statistical universals form power laws that take the form of $y \propto x^a$ for a constant a, where x and y are two variables measured for a linguistic finite sample from a source such as text. A power function has the characteristic of being invariable with respect to the scale of the data, and it thus represents some degree of self-similarity underlying a phenomenon. This can be explained by extending x through multiplication by a constant value λ. Then, $y \propto (\lambda x)^a \propto x^a$, so the relation between x and y is invariant with respect to the size transformation.

A power law thus signifies that a system is *self-similar*. The notion of self-similarity can be intuitively grasped from the Koch curve shown in Fig. 4.2. In the Koch curve, the whole includes itself as a part at a smaller scale. Because of this scale invariance, the power laws known to hold for a system are considered to possess the property of

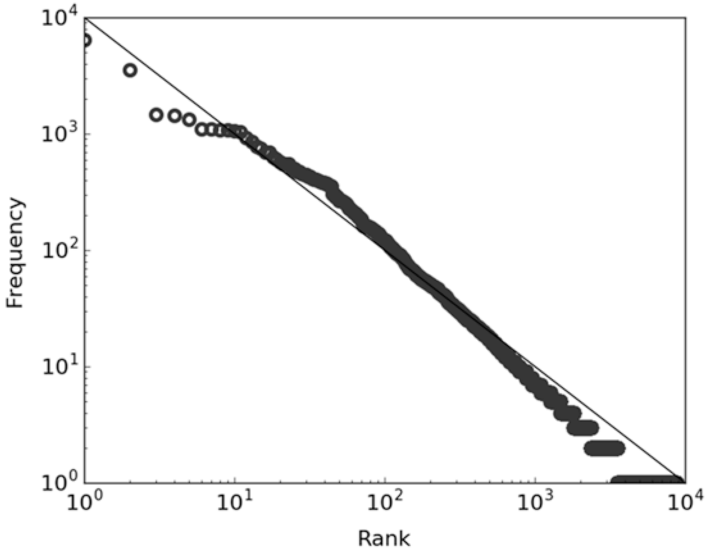
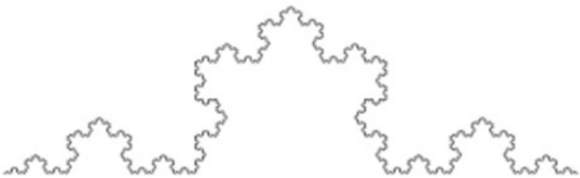


Fig. 4.1 Zipf's law plot of Wittgenstein's *Philosophical Investigations*. [The dots plot the actual rank-frequency data, whereas the black line shows a theoretical Zipf's law plot with an exponent of −1]

Fig. 4.2 Koch curve as an example of a self-similar shape



being *scale-free*, and such power laws are also called *scaling properties*. These properties indicate that the system is statistically self-similar, meaning that the scale invariance holds almost but not quite as cleanly as in the case of the Koch curve. In the case of Zipf's law, the self-similarity implies the fact that no matter how large a text is, a large part of the vocabulary occurs only once.

Several other power laws are known to hold for natural language, but they have relations with each other, and one law can often be derived (mathematically) from other laws. Among them, another important law that describes a different property than Zipf's law is the fluctuation underlying word appearance.

Any word in a text has the tendency to occur in a clustered manner. For example, Fig. 4.3 represents part of *Philosophical Investigations*, with a bar indicating the occurrence of words included in a set S consisting of some of the rarest words occur-

Fig. 4.3 Bursty occurrences of rare words (1200th–1500th words of *Philosophical Investigations*)

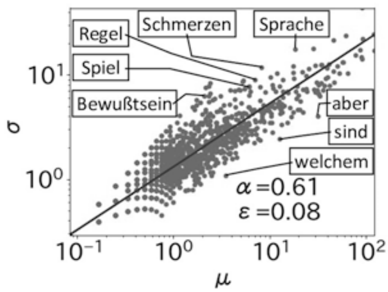


Fig. 4.4 Taylor's law plot of *Philosophical Investigations*

ring in PI.[1] The figure shows how the bars occur in a clustered, bursty manner: a certain region might have many bars, but another might have hardly any. Such bias in occurrence positions characterises language and cannot be expected for a uniformly random sequence. As such fluctuation can occur anywhere, for any word and for any set of words, the entire linguistic sequence makes the whole system self-similar.

There are various methods to measure the degree of fluctuation underlying language. Recently, by using Taylor's law, another power law reported widely for natural and social phenomena (Taylor 1961; Smith 1938), such clustering behavior was shown to be universal also in language (Tanaka-Ishii and Kobayashi 2018). Figure 4.4 shows an example Taylor's law plot for *Philosophical Investigations*. Here, one point corresponds to a word kind, and the points are scattered around the diagonal regression line in the middle. The figure is in log-log space, with the horizontal axis indicating the mean, and the vertical axis indicating the standard deviation of word occurrences within a 5620-word chunk of the given text.[2] Note that the standard deviation of word occurrences measures the degree of fluctuation, and the most fluctuating words—which are often keywords—appear the furthest above the regression line: some examples are annotated in the figure, including keywords such as *Spiel* and *Regel*.

In Tanaka-Ishii and Kobayashi (2018), this property was examined for over 1350 texts across 14 languages, including texts taken from Project Gutenberg, as well as news articles, child-directed speech utterances, music recordings, and program source code. Without exception, all texts showed the power-law behavior, with a different exponent from that of a uniformly random sequence. All the real data exhibited fluctuation phenomena of larger degrees, depending on the category. The clustering behavior underlying any text hence can be said to be a universal property.

[1] The figure shows a sequence of 300 words starting from the 1200th word. The set of rare words here consists of the rarest words up to 1/512 of the document length.

[2] This setting is arbitrary, and the result does not change for any chunk length.

The power laws of Zipf's law and Taylor's law, therefore, describe the universal nature underlying language. Similar laws are reported to hold in a wide variety of natural and social systems. For example, Zipf's law holds for the rank of a village's size and its population, the rank of income and the corresponding income amount, and many other phenomena.

Likewise, Taylor's law holds for cases such as stock price data and phenomena related to crops, habitats, and meteorology (Eisler et al. 2007). Power laws are common to large-scale systems surrounding us, and one example of these is language. Despite this prevalence, the causes of such power laws are unknown, thus driving us to conjecture on what language is, given these scientific facts.

Layers of Universals

A language universal is a common property that holds throughout all languages,[3] and the question of what a language universal could be has been an important theme in the domain of linguistics. Comrie (1989) categorized approaches to studying such universals as either empiricist or rationalist; among representatives of the latter approach is the work of Chomsky. Chomsky's universal grammar (Chomsky 2015) proposed a universal model of human grammar by elaborating his phrase structure grammar (Chomsky 1957). He considered the human linguistic faculty to be largely inborn, and thus he proposed rationalist models. Because his grammatical models are mathematical, they have influenced not only possible theories of language but also other fields, such as building computer program compilers. For language, however, Chomsky's theories have been controversial, especially from the empiricist perspective. For example, Tomasello (2001, 2005) studied the nature of language via infant prelinguistic utterances and primatology and presented counterexamples to Chomsky's theories.

Therefore, the main approaches to study language universals that have been widely accepted in linguistics are empiricist. As language is both syntactic and semantic, there are corresponding empiricist approaches of each kind. From the semantic viewpoint, Swadesh attempted to list the common words that exist in any language (Swadesh 1971). This resulted in various lists such as the Swadesh list. Unfortunately, the relevance of his intention has been undermined, because it is a difficult question to consider whether a word in one language corresponds with another word in a different language, given debates related to the meanings of words. In contrast, studies of syntactic universals have continued until today and were originated by American structural linguists. One representative is Harris (1970), who, for example, showed the mechanism that bridges between phonemes and morphemes (Harris 1955). As another example, Greenberg (1963) indicated a correlation tendency underlying word order: in particular, the basic word order of

[3] There are considered to be several thousand languages on the earth, with the exact range depending on how we define a language.

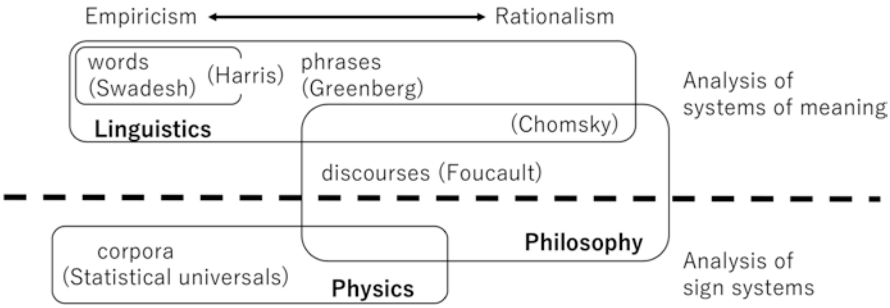


Fig. 4.5 Different layers of universals with respect to language

SOV correlates strongly with the modifier-modified order. Many of these nontrivial laws, however, do exhibit large exceptions. Nevertheless, the findings show the range of variation among natural languages around the globe and have flourished into projects such as the World Atlas of Language Structures, in which all the features of languages on the earth are collected by professional linguists.

Such variety of studies show that there are layers of granularity with respect to linguistic units in the quest for language universals. Figure 4.5 shows different layers, including Swadesh's words, and phrases in the works of Greenberg and Harris. Roughly speaking, linguistics' primary interests lie in these basic linguistic units such as words and phrases.

By extending the size of the target unit of language, the study of universals evolved beyond linguistics. For example, at the level of discourse, Foucault (1969) analyzed large archives across different fields and sought the common consequence of how thoughts evolve running through different fields. Interestingly, he emphasized the rarity and variation among utterances playing roles in the formation of a train of thought. Likewise, Zipf's law describes how a large portion of words remains rare, while fluctuation is the property that concerns mathematical variance. Therefore, what Foucault noticed through rarity and variation not only might derive from human rationality but could also be more deeply rooted in the nature of language.

The statistical laws of natural language consider a target unit larger than Foucault's, situated at the layer of a *corpus*. The difference between a *corpus* and Foucault's *field* is that a corpus is an archive of texts that can range across fields. The statistical universals of language are investigated with large-scale archives of newspapers and texts that cover multiple fields. Such studies only became possible with big data and powerful computers, and therefore the statistical universals constitute a new, state-of-the-art understanding that we have become able to gain only recently.

At the level of linguistics, the universals characterize semantics and grammatical structures, but outside linguistics the distinction between semantics and syntax becomes obscure, and the universals have been denoted under the name of *structure*. For example, Foucault sought the *structure* underlying human knowledge, in which analysis is conducted within the sphere of meaning and signification. Later, we will return to the philosophical notion of this *structure*. At the level of a corpus, words are analyzed even without considering the meaning. Every sign is comparable to a molecule, and text is considered like a gaseous body. The methodology, therefore, is naturally that of physics, specifically statistical physics, or more precisely complex systems theory. Note that laws are primarily an important concept in the domain of physics, e.g., Kepler's laws of planetary motion, the combined gas law, and so on. Power laws of language are closer to these kinds of physical laws than to Greenberg's language universals, and therefore statistical universals of language are discussed in the domain of physics rather than in linguistics.

Nevertheless, texts are written primarily for communication, and thus, by nature, to produce meaning. At the level of linguistics, universal findings would show the possible range of words and phrases. At the level of discourse, the resulting universals would show the range of the human symbolic sphere. At the level of the molecular view of texts, however, the analysis itself is conducted beyond any meaning that a text would produce. What then would the universals underlying the corpus imply?

As mentioned at the end of the previous section, similar power laws have been reported for many other natural and social systems. For all these laws, the underlying reasons for them are unknown. The only understanding that we have gained from the statistical universals of language is that the linguistic system is possibly one of those systems. We have the tendency to consider that language is a human system, thus originating in intelligence and being exclusive to humankind, but the findings thus far suggest that language is yet another system spawned from natural systems.

When we perform linguistic acts, we are not aware that our language is seamless with social and natural systems. We never conduct linguistic actions by aiming to produce global statistical universals. Instead, what we do is to choose the next word or next sound, typically trying to be meaningful, and the accumulation of such actions inevitably leads to the global properties. The universals at the corpus level suggest that there is some *behavior* underlying human linguistic acts. Given the fact that the same statistical properties are apparent in language-related activities even beyond language, there is some *behavior*, or some mechanism, underlying human handling of a sequence of signs. There must be some particular acts that we perform that imply the statistical universals of language. The rest of the chapter considers this unsolved question of what kind of behavior we follow.

Mathematical Reasoning on Statistical Universals

Because the appearance of the statistical universals is mathematical, a natural consequence is to also seek reasons for them in a mathematical way within the field. A mathematical cause would only remain a model, of course, and therefore unveiling the source of the behavior will require a leap in neuroscience. Still, a mathematical explanation might provide some conditions that would lead to power laws.

An explanation to deduce Zipf's law is possible via optimization. Mandelbrot (1952) demonstrated how minimizing the cost per amount of information would mathematically imply a linguistic sequence with the rank-frequency relation being the power law. By defining the cost to be linear with respect to the length of a word, the total cost C of a text is defined, given the probability of every word $p(w)$. Letting the information amount (Shannon entropy) be H, minimization of C/H implies that $p(w)$ should be proportional to the power of the word's rank.

This formulation of Mandelbrot was constructed so that the problem could be mathematically solved. The optimization function considering the cost per information amount was already somewhat arbitrary. Moreover, the optimization scheme could only explain Zipf's law and not the other laws related to fluctuation. An extension to include other conditions would have instantly made the problem intractable to solve.

Above all, Mandelbrot's idea suggests that we optimize text globally throughout. In other words, for an entire given sequence, the probability of production is globally optimized. We wonder, however, whether this is the case. As mentioned above, the statistical properties roughly hold even for an infant's utterances. Do we not choose which word to use at every moment? Instead, it seems more appropriate to seek a process that is more consecutive.

One such approach is a stochastic process. Miller (1957) demonstrated how monkey typing would generate Zipf's law. Given a set of characters, consider a monkey that randomly types one character after another and then a space. Because a space separates words, let the space be hit with a certain probability, and otherwise let every character be hit uniformly. A monkey would type a sequence of characters and then a space, another sequence of characters and a space, and so on. Such a simple procedure is mathematically proven to result in a rank-frequency distribution exhibiting Zipf's law. Miller's argument suggests the possibility of underlying behavior caused by something random, even if it is not as simple as monkey typing.

Of course, examination of monkey typing in detail shows its limits, and it does not correspond to what we do in human language production. It requires further constraints to exhibit Zipf's law. Moreover, monkey typing continues randomly, so it does not reproduce fluctuation.

One important another stochastic process is the Simon process, defined as follows (Simon 1955). The process generates one word after another along time t. The process initiates at $t = 1$ with one arbitrary word. Given the sequence already produced, a speaker chooses the next word by the following rule:

- With a constant probability α, choose a new word.
- Otherwise (i.e., with probability $1-\alpha$), choose a word randomly from the past sequence.

For example, suppose that, at $t = 1$, the sequence started with [a], and at $t = 5$, the sequence had developed into [a, a, b, a, b]. The above two rules stipulate what to do at $t = 6$. With probability α, a new word is chosen (such as c or z: any word besides a or b); otherwise, a word from [a, a, b, a, b] is chosen. Note that in this latter case, previous elements are chosen in proportion to their frequencies. In this example, a is chosen with probability 3/5, whereas b is chosen with probability 2/5. Therefore, the more frequent an element is, the more it is chosen, leading to the phenomenon of *the rich getting richer*.

Mathematically, this process is proven to generate Zipf's law (Barabási and Albert 1999; Mitzenmacher 2003) but not Taylor's law, although it does reproduce another statistical universal of *long memory*, which has a relation with Taylor's law (Eisler et al. 2007). It therefore still has problems in reproducing fluctuation (Tanaka-Ishii and Kobayashi 2018).

Even if the Simon process itself cannot reproduce all the statistical universals, a stochastic process is far easier to study than optimization. Therefore, the search continues for a process that can generate a sequence exhibiting the statistical universals known for language. A further consequence will be summarized later, in terms of reproducing both laws presented above.

Overall, we have seen two different approaches in the scientific domain to understand the reason for production of the statistical universals of natural language: optimization and a stochastic process. Because of its formulation, optimization is limited to explaining a specific law. That study approach does not tell us that we optimize at every word production. In contrast, stochastic processes are flexible and have the potential to illuminate what lies behind our linguistic acts. Thus far, at least, we have seen that our linguistic actions are neither monkey typing nor a Simon process. As mentioned before, even if we obtain a process that reproduces all the statistical universals, it can never be said that such a process is the actual linguistic process. The quest could reveal, however, what qualities are necessary in consecutive linguistic acts.

These attempts towards reasoning on statistical universals suggest that behind our linguistic acts could lie a machine-like procedure. The stochastic process considered above is indeed mechanical, to the extent that a computer can produce it. A machine performs a defined task that does not require a moment of thought for every action. The consecutive actions of a machine are predefined, and it only fulfills its mission. Such action lacking a moment of thought is contradictory with respect to language production, because language production is about thinking. I am not suggesting that thought is absent in producing a meaningful sequence for communication, of course. Rather, I would like to indicate that the statistical universals of language show that behind the production of a linguistic sequence lies some certain but unknown behavior, which can be compared to a machine.

This raises the philosophical term *machine* for a further conjecture on statistical universals. The term *machine* is a key notion in postmodernist philosophy, often contraposed to the notion of a *structure*. If a generative model were to correspond to a *machine*, then the notion of a *structure* would be deemed to correspond to the statistical properties. We are interested in how such correspondence would add to better interpretation of what statistical universals and their generative models are. Before proceeding further, the next section summarizes the philosophical notions of a machine and a structure. Then, the statistical properties are considered in the context of postmodernist philosophy.

Postmodernist Machine Versus Structure: A Brief Summary

In analyzing a humanistic phenomenon, such as "language," "unconsciousness," or even components such as "discourse," briefly, a *structure* analyzes the phenomenon by describing underlying common properties, whereas a *machine* does so by describing the mechanism to produce the phenomenon.

Structuralism is a philosophical trend established by Saussure (1911) and intended to describe the characteristics of a phenomenon as a structure underlying interconnected elements within a holistic system. As Saussure was a linguist and applied his theory primarily to language, a typical structuralist analysis applies to a system of meaning, as shown at the right-hand side in Fig. 4.5. The target of analysis was soon extended to include phenomena in which the meanings of elements are not as clear as in typical texts. The first representative attempt was that of Lévi-Strauss (1962), who indicated that the relation among ethnic tribes interconnected via marriage can be formalized using a mathematical group. Later, Kristeva (1969) indicated the importance of structural analysis of semiotic systems in general, situating symbolic systems of meaning as a subset of semiotic systems. Since then, various semiotic systems have been analyzed structurally (Nöth 1990; Trifonas 2015). In parallel to these semiotic approaches, Lacan (1998) introduced structuralism to psychoanalysis, advocating that our unconsciousness has a structure.

All these works attempted to extract the common nature of the target phenomenon in terms of structure. A structure is meant to describe the common characteristics behind phenomena, and in this sense it tends to presuppose the static appearance of a specific characteristic. If multiple samples of phenomena are characterized by some common feature, then this common feature is extracted as the structure. Because this analytical process is inductive, for natural and social phenomena, the universality of the structure is often a matter of degree, as in the case of a structural linguist's universals, as mentioned above. The universals often depend on both the data and the methodology for observing the phenomena.

This presupposition of static, common properties underlying phenomena, however, led to criticism of structuralism, as the most important viewpoint in philosophy is that texts are singular. As an alternative, the notion of a *machine* was raised, because it highlights the singularity of a phenomenon and the dynamism behind it, as will be conjectured further in the following section. The notion of a machine as a contraposition of a structure became prevalent in Deleuze and Guattari (1972). The notion is fundamentally linked with physiology, as shown by the term *desire machine*. For example, we eat and then digest, and such a consecutive procedure works as if all the physiological parts with various functionalities all together perform a particular task, beyond our consciousness. Deleuze and Guattari described how such a machine-like mechanism is not only about physiology but could also be a social outcome.

This notion of a machine did not first appear in Deleuze and Guattari (1972). In a preceding short essay, Guattari (1969) contraposed the structure and machine in the title. It seems that human behavior driven by some mechanism, such as drumming one's fingers on a hard surface, has been a fundamental train of thought in the field of psychiatry. Guattari focused on the idea to understand various psychiatric pathologies and even extended it to broadly consider various phenomena. He indicated that, behind any structural system, there is a machine, and this is applicable also to language, as follows from Guattari (1969) [p.322]:

> The voice, as speech machine, is the basis and determinant of the structural order of language, and not the other way round. The individual, in his bodiliness, accepts the consequences of the interaction of signifying chains of all kinds which cut across and tear him apart. The human being is caught where the machine and the structure meet.

This quote suggests that Guattari considered language as a machine that generates the *structural order of language*. Later, in Guattari (1979), he elaborated more extensively how language could be produced through a machine. Some ideas resonate with whatever underlies statistical universals, like comparing words to *molecules*. Guattari's primary interest was psychiatry, and he mainly considered social factors with respect to machines. In the introduction, however, he referred to Thom (1974)[4] and approved the possibility of the inevitable influence of physics and biological nature.

The concrete shape of a machine, as an outcome of the overall conjecture, is a *rhizome* (Deleuze and Guattari 1998). A rhizome is a plant organ acting as a kind of stem and root at the same time, seen widely among plants such as weeds and forbs. Whereas a stem grows upwards and a root grows downwards, a rhizome extends horizontally close to the surface and grows by generating and connecting with other rhizomes. Rhizomes are known to connect easily to rhizomes of other individuals by grafting.

Deleuze and Guattari (1998) demonstrated how various social and natural systems can be analogically considered as rhizomes. They set out three principles of a rhizome: connectivity, homogeneity, and multiplicity. Just as language and signs were discussed (Deleuze and Guattari 1998), language was situated as a kind of rhizome in their thoughts. Indeed, from the perspective of the three principles, the rhizomatic nature of language is apparent. For example, a technical term in the field of botany such as *rhizome* can easily be grafted onto the field of philosophy and transformed further into a daily term. Language therefore encompasses various words with different roots, thus forming a system with multiplicity. Such a system can be achieved by the homogeneous nature of signs, flexible media that can easily be used with any other, and the system appears by representing the signification. Comparing language to a rhizome is therefore deemed reasonable, although it remains an analogy.

[4] Thom, a mathematician, was a Fields Medal winner and the founder of catastrophe theory. He modeled language by using his theory.

Statistical Universals as the Structure of Singular Samples

In science, the contraposition of the structure and machine would seem to imply that descriptions via a structure and via a machine should ultimately be the same, as they describe the same phenomenon.

If a machine operates under a particular setting, and if it is a stochastic machine, then it will produce a set of samples with a static mean and variance. A machine thus describes how a sample space is produced. In contrast, a structure describes the property characterized by the machine, with a mean and variance of the property. This scientific view of structure versus machine is depicted in the upper panel of Fig. 4.6. The figure shows some samples observed at a given time (usually now), and a corresponding generative model. Here, a sample is reproducible by the machine, being replaceable with similar samples. If the machine changed, then the resulting structure would change, and this change would be observable through different sample spaces. Overall, the approaches from either a structure or a machine should thus ultimately be equivalent.

On the other hand, in philosophy, the most important aspect of the entire discussion is the singularity of a sample. Among many other possibilities, every sample is *the* sample, being irreplaceable, unreproducible, rare, unique, and singular. In contrast to a mathematical set of samples, in which all possible samples are readily projected to form a space, the sphere of possible reality is neither tractable nor existent. There are only some real samples, characterized as rare, that do not easily form the shape of a sphere. The interest of philosophy is to describe the nature that threads through these rare samples. This nature is called the *structure* in

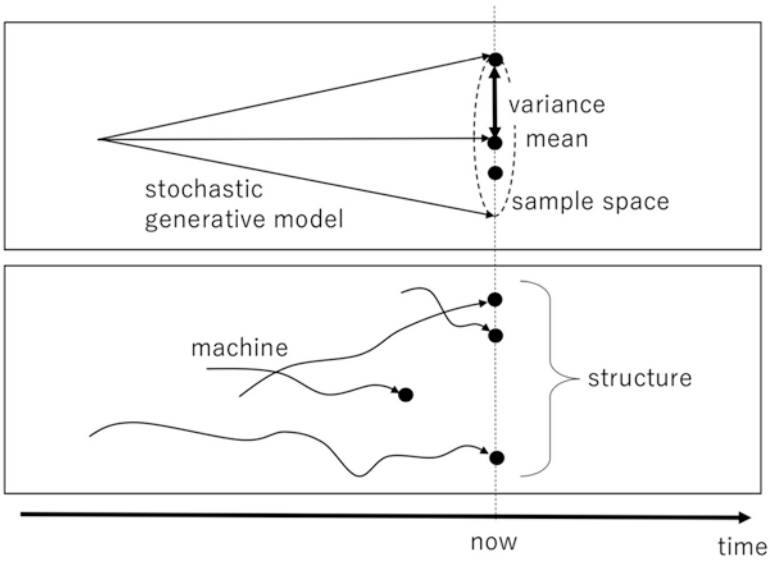


Fig. 4.6 Structure vs. machine in the fields of science (upper panel) and philosophy (lower panel)

structuralism. The method of structural analysis is obliged to generate abstraction among samples, however, even though each sample is singular. The notion of singularity and the spirit of abstraction underlying the analysis are contradictory, and instead the alternative notion of a *machine* was proposed.

With the method of a machine, because a (stochastic) machine with the same settings can already generate different samples, with different settings, the outcome could be completely different. Moreover, when the nature of the phenomenon changes over time, the consequent phenomenon at a later time would become mixed of different nature, making it difficult to analyze. With a machine, the change can be analyzed as a change in the machine, which explains the dynamism of the phenomenon. A *machine* therefore is a generative view of a process, but in philosophy it is meant to highlight the singular dynamism necessary to reach a phenomenon. This view is depicted in the lower panel of Fig. 4.6.

The scientific and philosophical methods with respect to a structure and a machine are thus deemed to approach phenomena from the opposite extremes of abstraction and singularity, respectively. The scientific method considers the target through abstraction, typically in the form of mathematical tools, by disregarding the singularity of phenomena. This leads to the equivalent analysis results for approaches based on a structure and a machine. In contrast, the philosophical method considers primarily instances and is therefore limited in abstraction. Structuralism has been devoid of means to describe an abstract structure, probably in part because of this emphasis on singularity. Structuralist debates often rely on analogy, as criticized by Sokal and Bricmont (1998). The consequences of analysis in terms of a structure and a machine thus differ for this philosophical approach.

With this view, it could be conjectured that statistical universals show the structure underlying singularity. The singularity of a text is primarily formed by rare words. Zipf's law shows the number of rare words that characterize a text. The vocabulary infinitely increases as the context evolves in a sample text, and this increase is described by Zipf's law. Taylor's law indicates how these words are assembled along the evolving sequence. The statistical universals show that there is an exclusive nature: a concrete, tractable structure among rare events. The laws indicate a self-similar structure among these rare words that results in the form of power laws.

It is important to note that, in this sense, the statistical universals provide an advanced understanding about language, at the meeting point of the two approaches of abstraction/science and singularity/philosophy. The statistical universals could be possible concrete evidence of what structuralism aimed to describe, namely, the universal structure underlying the singularity of texts. Unlike mathematical notions appearing in previous philosophical conjectures, the statistical universals are not analogies. Whereas other language "universals" often have exceptions, the statistical universals are strictly universal and apply to any text instance. The description is explicit, in the form of a power law, and not by analogy. To the best of my knowledge, thus far, the nature of the singularity underlying texts has been explicitly described in the form of a structure only via these statistical universals.

The problem of fully understanding this nature, however, has not been solved, as mentioned previously. Mathematical generative models of language do not fulfill all the statistical universals. From a *machine* perspective, we want a model that fulfills all the statistical universals, thus explaining the singularity underlying language.

Two Mechanisms that Almost Reproduce Statistical Universals

Several possibilities have been studied in the quest for procedures that reproduce the statistical universals of natural language. Different procedures have been tested using computers to verify whether the consequences of stochastic generation can lead to the statistical universals of natural language. For example, given a Simon process, a long sequence can be generated stochastically following the definition mentioned above. The resulting sequence can then be tested as to whether it follows the statistical universals of natural language, and in fact such sequences were found not to reproduce fluctuation (Takahashi and Tanaka-Ishii 2019). Similarly, it is possible to test whether a generative process fulfills the statistical universals. Currently, there are two mechanisms that almost reproduce the known statistical universals.

The first is a neural network. The scale of neural computing has greatly increased, and recent neural network systems have improved the mechanisms that handle memory and context over a longer term. Given a context consisting of a certain number of preceding words, the network structure as a whole predicts the subsequent word through vast, nonlinear computation of vectors by using multiple layers of matrices, whose parameters are estimated in order to match the input with the desired output. It is known that different nodes of the neural structure take responsibility in representing different semantic targets (Le et al. 2012), and thus the network learns to handle the given context. The resulting neural network can stochastically produce words after words. Before neural language models, no language model was capable of reproducing all the statistical universals, but state-of-the-art neural models do to some extent (Takahashi and Tanaka-Ishii 2019). Current neural networks are effective for various language engineering problems (such as automatic translation), partly because they are successful in assimilating language to reproduce statistical universals.

Compared with natural language texts, however, neural networks are still limited in two aspects, which are represented by the statistical universals: the capability to produce new instances, and the limitation of capturing fluctuation. Zipf's law indicates that the number of words should increase infinitely with respect to the data size. With current neural networks, however, the dimension is cut off to be finite. As for the degree of fluctuation, neural networks are still limited in handling context. Context represents that a sequence can change and evolve, and such behavior is not achieved in current architectures. Precisely, a text generated by a neural network cannot produce enough fluctuation. Furthermore, current neural networks lack ways to evolve after training.

The second candidate mechanism is a stochastic procedure based on a complex network (Tanaka-Ishii and Kobayashi 2018). Complex networks (Barabási 2016) are studied as a subfield of complex systems theory, and various power laws are known to hold for those networks, too. A linguistic word sequence also forms a complex network, for example, by considering every word to form a node; a branch between nodes is then formulated when one word appears following another word. One simple way to produce a sequence on such a network is by a random walk. Given a network structure, a procedure stochastically conducts walks on the graph and produces an output at every node visited during a walk. Introduction of new instances is integrated naturally by visiting new nodes, or the network structure can even be extended during a walk. With this mechanism, Zipf's law follows, if the network structure is readily constructed to exhibit the power law. Moreover, fluctuation is naturally implemented through the random walk, because it allows leaving and coming back to a node, and thus nodes in proximity have the possibility to be repeated. Therefore, a random walk on a complex network system can reproduce clustering phenomena (Tanaka-Ishii and Kobayashi 2018).

There are still two limitations related to this approach, too. Tanaka-Ishii and Kobayashi (2018) showed that Zipf's and Taylor's laws are met by some very specific kinds of random walks. As mentioned before, however, there are other laws related to fluctuation, and they are not so well met by these kinds of walks. It is nontrivial to construct a mechanism to readily reproduce all the statistical universals with this approach. The second limitation concerns the relevance of the random walk. A random walk can be controlled probabilistically, but we surely do not "randomly walk" to produce language. The notion seems to move back to a Markovian view of language. The alternative, however, is not trivial and remains an open question. Relatedly, this approach does not produce a language model that serves for language processing. It does not have the notion of input, nor does it have parameter tuning to learn a real text.

The neural language models have been mainly studied in the engineering domain, whereas the random walk on a complex network has been considered more in the physics domain. They are not, however, independent of each other. Neural networks are gigantic in scale, but the effective links of the network are considered limited after training. The structure of such a network is deemed to form a complex network, similar to the basis of a random walk. Attempts to tackle the problems of both methods could lead to a neural network using the structure of a complex network. Currently, the state of the art for both mechanisms is related to large-scale complex networks, and therefore the following section refers to both by the generic term *complex network machines.*

Complex Network Machines Versus *Rhizomes*

Complex network machines are state-of-the-art forms that almost produce statistical universals. As discussed in the previous section, their limitations relate to statistical universals, which serve not only for understanding the nature of language but also for highlighting the weak points of language models.

Because a machine is the source of a structure, this section conjectures on the nature of the limitations of these complex network machines with respect to singularity, from a machine perspective. Again, note that the philosophical *machine* in this chapter is contraposed to the notion of a *structure*, which is deemed to correspond to the statistical universals. Above, it was conjectured that statistical universals describe the universal nature underlying the singularity of linguistic text samples.

The Simon process, as described previously, is a representative *nonstationary* mathematical process. Stationarity defines a property underlying a time sequence, to be unchanging, and that succeeding elements are stochastically predictable. A neural network is still limited in producing such nonstationary behavior despite its recurrent construction, and once its training converges it almost stops evolving. A random walk on a complex network that evolves, however, can be non-stationary. From this perspective, a future model that combines the two mechanisms is deemed important.

Moreover, the two mechanisms are far larger than the Simon process. They are based on large-scale networks, and the prediction performed is far more complex than that of a toy algorithm. Because of this scale, a sample generated by either of these mechanisms is rare, in the sense that the probability to produce a sample becomes very small. Singularity is thus partly secured by making the machine large in scale. Singularity is not only about being rare in terms of possibility, however, and a sample should be a necessary, inevitable choice excluding all other samples. Both complex network machines do not produce samples characterized by any necessity. In future models, however, one way to grant this necessity could be through *optimality* among other possibilities in a context.

In contrast in philosophy, a state-of-the-art machine proposed in philosophy is the rhizome, introduced previously. The outcome of a rhizome is a complex network. Some works indeed studied the quantitative nature of plant rhizomes as a structure (Armstrong 1983; Majrashi et al. 2013) and showed how they form complex networks. From this common structure of a complex network, the two complex network machines are deemed as becoming almost successful in fulfilling the statistical universals partly because they have acquired rhizomatic characteristic. Even in a limited manner, the two mechanisms by nature possess the three rhizomatic principles of connectivity, homogeneity, and multiplicity. First, neural networks are connective, in the sense that neural network parameters determine the degree of connectivity among branches, until convergence through training. Neural networks are also homogeneous, because different aspects of the target are each represented by a node and are connected as a network. Lastly, they have multiplicity, to consider various targets. A random walk on a complex network also meets all the criteria, as the base structure is a complex network, i.e., a rhizome.

Nevertheless, the qualities of complex network machines are still limited, compared with the philosophical notion of a rhizome. One of a rhizome's most important characteristics is that, although it was introduced as an example of a machine, it constructs a network *structure* via a mechanical procedure. In other words, a rhizome is a structure and a machine at the same time. In comparison, our complex network machines are almost machines and structures at the same time, but the structure and machine are still distant from each other. A neural network starts from a predefined dense network structure and terminates by adjusting the structure in the form of a complex network. After this training, it functions as a machine. As for a random walk, the network is built as a structure via a mechanical process, but its overall mechanism is still premature, and how to put it to use still requires study. These limitations in attaining complete deconstruction characterize the state-of-the-art scientific understanding about language.

Concluding Remarks

An overview of the different concepts considered in this chapter appears in Fig. 4.7. The left panel shows the notions of the science of language. We first considered universals, with the focus on statistical universals. By comparing with other universals derived from structural linguistic thought, the characteristics of the statistical universals were highlighted. To tackle the problem of trying to understand why such properties hold, possible explanations were studied via stochastic processes. This included various simple mechanisms and recent larger scale procedures. The analysis thus far is characterized by means of abstraction in which samples are reproducible. The right panel shows the corresponding notions in postmodernist philosophy, characterized by the singularity and the dynamics behind it. The contraposed notions of a structure and a machine are summarized as shown in the figure. A structure corresponds to studies of universals that highlight the nature of singularity. On the other hand, a machine represents possible dynamics that explain the appearance of the structure. The previous section argued to what extent complex network machines could correspond to the postmodernist notion of machines. A rhizome was situated as the form deconstructing a structure and a machine, and the comparison between a rhizome and the state-of-the-art mechanisms of language science was debated.

Deleuze and Guattari (1991) mentioned how science takes the place of philosophy. Complex systems theory had already been studied in the 1970s, in parallel to

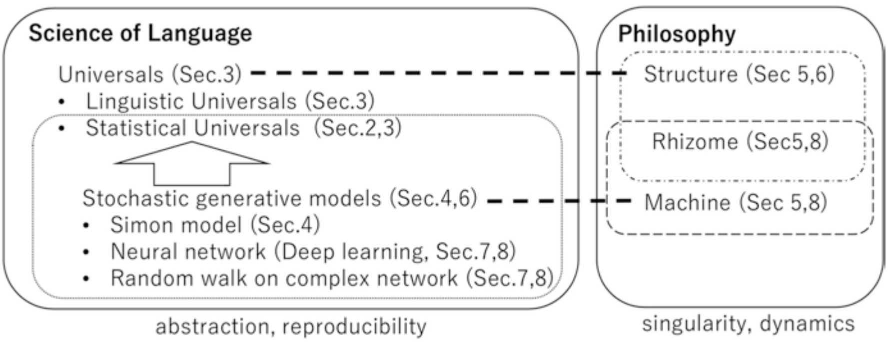


Fig. 4.7 Correspondence between language science and philosophical notions

those authors' writings. Although the term rhizome remains an analogy, it provides a good prescientific philosophical insight to help reconsider the position of science and engineering. The question remains, however, as to what kind of machine/structure is a rhizome. Furthermore, how can it be implemented on computers as a language model? The domain of philosophy does not provide any concrete answers to these questions. To make a step forward, we need a detailed, concrete form of what kind of structure language has and what kind of machine can produce it. The methodology of mathematics and physics is a possible path to better understand language. To this end, it is important to study the structure underlying the nature of language, possibly characterized by singularity. In general, we do not know the basic properties of the linguistic data that we produce, nor do we have much science about big data. The quest is a possible path towards knowing what a rhizome as language is and how to implement it.

Acknowledgment I sincerely thank Professor Kyo Kageura of University of Tokyo for his valuable comments on the first draft of this chapter.

References

Armstrong, R. (1983). On the quantitative theory of resource partitioning in rhizomatous perennial plants: The influences of canopy structure, rhizome branching pattern, and self-thinking. *Ecological Society of America*, pages 703–709.

Barabási, A.-L. (2016). *Network science*. Cambridge: Cambridge University Press.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, pages 509–612. 286.

Chomsky, N. (1957). *Syntactic structures*. The Hague/Paris: Mouton.

Chomsky, N. (2015). *The minimalist program*. Cambridge: MIT Press.

Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago Press.

Deleuze, G. and Guattari, F. (1972). *L'Anti-eedipe*. Paris: Editions de Minuit.

Deleuze, G. and Guattari, F. (1991). *Qu'est-ce que la philosophie?* Paris: Editions de Minuit.

Deleuze, G. and Guattari, F. (1998). *Mille plateau*. Paris: Editions de Minuit.

Eisler, Z., Bartos, I., and Kert'esz, J. (2007). Fluctuation scaling in complex systems: Taylor's law and beyond. *Advances in Physics*, pages 89–142.

Foucault, M. (1969). *L'arch'eologie du Savoir*. Paris: Gallimard.

Greenberg, J. H. (1963). *Universals of language*. Cambridge: MIT Press.

Guattari, F. (1969). *Machine et structure*. *La decouverte*. Originally published in a book form from Editions François Masp'ero in 1972.

Guattari, F. (1979). *L'inconscint machinique, essais de schizo-analyse*. Editions Recherches.

Harris, Z. S. (1970). *Morpheme boundaries within words: Report on a computer test*, pages 68–77. Reidel Publishing Company. Originally appeared in *Transformation and Discourse Analysis Papers*, 73 in 1967.

Harris, Z. S. (1955). From phoneme to morpheme. *Language*, pages 190–222.

Kristeva, J. (1969). *Semeiotiké - Recherches pour une sémanalyse*. Paris: Edition du Seuil. English translation: *Desire in language: A semiotic approach to literature and art*, Oxford: Blackwell, 1980.

Lacan, J. (1998). *Le séminaire de Jacques Lacan*. Polity. Livre V, Formations of the Unconscious, translated by Russell Grigg.

Le, Q., Ranzato, M., Monga, R., Devin, M., Chen, K. an Corrado, G., Dean, J., and Ng, A. (2012). Building high-level features using large-scale unsupervised learning. In *International Conference on Machine Learning*.

Lévi-Strauss, C. (1962). *La pensée sauvage*. Paris: Poche.

Majrashi, A., Bin Bakar, B., Moneruzzaman Khandaker, M., Nasrulhaq Boyce, A., and Muniandy, V. (2013). Fractal analysis of rhizome growth patterns of scirpus grossus l. on peat and paddy soils. *Bulgarian Journal of Agricultural Science of Agricultural Academy*, 19(6): 1319-1326.

Mandelbrot, B. (1952). An informational theory of the statistical structure of language. *Proceedings of Symposium of Applications of Communication theory*, pages 486–500.

Miller, G. (1957). Some effects of intermittent silence. *American Journal of Psychology* 70: 311-314.

Mitzenmacher, M. (2003). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics* 1(2): 226-251.

Nöth, W. (1990). *Handbook of semiotics*. Bloomington: Indiana University Press.

Saussure, F. de (1911). *Cours de linguistique générale*. London: Pergamon. Notes taken by E. Constantin. Translated into English by R. Harris, 1993.

Simon, H.A. (1955). On a class of skew distribution functions. *Biometrika* 42(3/4): 425-440.

Smith, H.F. (1938). An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agriculture Science*, 28(1).

Sokal, A. and Bricmont, J. (1998). *Fashionable nonsense: Postmodern intellectuals' abuse of science*. Brockman.

Swadesh, M. (1971). *The origin and diversification of language*. London: Routledge. published in 2006.

Takahashi, S. and Tanaka-Ishii, K. (2019). Evaluating computational language models with scaling properties of natural language. *Computational Linguistics*. in press.

Tanaka-Ishii, K. and Kobayashi, T. (2018). Taylor's law for linguistic sequences and random walk models. *Journal of Physics Communications* 2(11).

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature* 732: 189-190.

Thom, R. (1974). *Modéle mathématiques de la mophogenèse: recueil de textes sur la théorie des catastrophes et ses applications*. Union générale édition. Translation "Mathematical models of morphogenesis", published from Halsted Press, New York.

Tomasello, M. (2001). *The cultural origins of human cognition*. Cambridge: Harvard University Press.

Tomasello, M. (2005). *Constructing a language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.

Trifonas, P. P. (ed.) (2015). *International handbook of semiotics*. Springer. World Atlas of Language Structures (visited in 2018). WALS Online.

Zipf, G.K. (1965). *Human behavior and the principle of least effort: An introduction to human ecology*. Hafner.

Chapter 5
Number Work: Recovering the Original Complexity of Learning Arithmetic

Brent Davis

Pedagogical Impasses

- "When I was brainstorming different ways of saying 'subtract' with my Grade 2s, one of the children got angry that 'make smaller' was on the list. He argued that '*making smaller* can't be *subtracting* since five will still be five no matter how small you make them'."
- "I teach Grade 5. Last year, when we were looking at the formula for circumference of a circle, $C = 2\pi r$, one student knocked everything sideways when she asked, 'If π goes on forever, how can you times it by 2?'"
- "I found it difficult to get my 7th-graders to measure angles. They can't seem to figure out how to use their protractors properly, no matter what I do."
- "My students in Grade 8 struggle with subtracting integers. They can follow the rule, but no one seems to get why 'adding the opposite' makes sense."
- "For me, the sticking point is algebra. Students can't see the difference between an unknown and a variable. Let's say, for example, we have '$2x = 8$' and '$2x = y$,' my kids would be confused what x as a symbol means and does. If I said '$2x$' and '$2p$,' it wouldn't occur to them that the x and p could be the same. Oh … and when they look at a graph, some students don't seem to see the continuous lines, just the points that line up with the whole numbers."
- "A Math 30-1 [Grade 12] student asked me why we can't imagine imaginary numbers. I think that might a big part of the reason I decided to move to junior high this year."

These brief descriptions of "pedagogical impasses" were offered by teachers in response to the prompt, "Tell us about a time you were teaching mathematics in which you found yourself stymied by something a student said or did."

B. Davis (✉)
University of Calgary, Calgary, AB, Canada
e-mail: brent.davis@ucalgary.ca; abdavi@ucalgary.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*, Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_5

Those teachers, all from the same school, are taking part in a longitudinal project aimed at improving the mathematics learning experiences of their students in grades 1 through 8. Now in its 7th year, the project has multiple strands. The principal efforts are focused on enacting more consistent lesson formats across grades, developing reporting strategies that afford parents nuanced information on their children's efforts and understanding, and attending to research from mathematics education and the cognitive sciences, especially research into teachers' disciplinary knowledge of mathematics. Regarding the last of these goals, since the start of the project, each school year has begun with a professional learning session focused on some aspect of mathematics for teaching. Through most of the project, a committee comprising a few lead teachers and myself selected the specific topic and structured the session. During the latest phase of the project, we tried a different tack. Aiming for something more participatory and emergent, our planning consisted of nothing more than the invitation for narratives of pedagogical impasses, sent out a few days before the session.

The notion of "pedagogical impasse," as developed here, isn't entirely new. There have been rich examinations in the mathematics education research literature of moments in learning and teaching when momentum is lost. In the 1990s, the phrase "epistemological obstacles" rose to some prominence as a means to account for many such events. As Sierpinska (1994: xi) defined them, epistemological obstacles are:

> ways of understanding based on some unconscious, culturally acquired schemes of thought and unquestioned beliefs about the nature of mathematics and fundamental categories such as number, space, cause, infinity, … inadequate with respect to the concept in history, and remained somehow "implicate," to use Bohm's term, in its meaning.

According to this description, and as will become evident in my analyses, all of the pedagogical impasses noted above appear to have been rooted in epistemological obstacles. But, that said, I believe there to be an important difference between obstacles and impasses. Based on Sierpinska's description, an epistemological obstacle is a cultural pothole for learners—a describable, identifiable aspect, a potential block to understanding that can be flagged and anticipated by teachers. By contrast, a pedagogical impasse is more amorphous. It arrives as a sensation of being lost, a feeling that something is amiss paired with an inability to home in a specific issue. Phrased differently, a pedagogical impasse typically raises questions for which the answers tend to be epistemological obstacles.

Certainly, that's how most of the few dozen teachers who participated in the project experienced the impasses they related. Every one of them arrived at the session with at least one tale of frustration. And, in every case, the telling of the impasse drew nods of familiarity and smiles (or sighs) of sympathy from the other teachers.

But at no time did these narratives trigger discussions of likely epistemological obstacles—which, in fact, prompted me to grow more and more uneasy as the session unfolded. Indeed, I couldn't hold back from asserting something that I thought should be obvious to all: across multiple concepts and multiple grades, every incident had to do with the same obstacle: they all revolved constrained understandings of *number*. At least, to my ear, they did.

Mathematics for Teaching

Following Davis and Renert (2014: 9), mathematics for teaching is framed in our project as:

> a way of being with mathematics knowledge that enables a teacher to structure learning situations, interpret student actions mindfully, and respond flexibly, in ways that enable learners to extend understandings and expand the range of their interpretive possibilities through access to powerful connections and appropriate practice.

Less formally, within the project, we talk about mathematics for teaching as "what an expert needs to know to think like a novice." Dipping into the novice–expert literature (Ericsson et al. 2006), this characterization is invoked to highlight a defining feature of expert knowledge across domains—namely, the expert's ability to recognize immediately when a concept is appropriate, without conscious mediation, no matter the situation. Novice understanding, however, tends to be much more piecemeal, deliberate, and context bound. The differences arise in the fact that experts have had time and opportunity to integrate diverse instantiations, applications, and representations into consolidated, coherent wholes. For novices, concepts often lack such coherence, and so different instantiations of the same concept can be experienced as unconnected. Consequently, situations sometimes arise in which novices cannot reconcile instances fitted to a concept, whereas an expert might not be able to distinguish among them. I interpret the pedagogical impasses introduced at the start of this chapter in precisely such terms—that is, as moments in which teachers were unable to disentangle elements of their expert knowledge of "number" in order to make sense of learners' inabilities to find coherence across not-yet-connected experiences. Teachers' consolidated understandings deafened them to the disconnects in their students' interpretations.

That mathematical concepts are regular sites of pedagogic struggle is unsurprising. As has been argued and researched by phenomenologists and cognitive constructivists for more than a century, concepts are not static forms or unified wholes which can be shared among knowers. Rather, they evolve across experiences and interpretations that are specific to individuals. Hence, pedagogical impasses such as those presented above should not be met as mistakes, but as divergent construals.

Indeed, as signaled at the end of the previous section, I experienced my own pedagogical impasse as a session leader. For me, it seemed obvious that everyone was talking about number. However, when I artlessly said so, the shaking heads and hasty objections made it clear that few, if any, participants appreciated that my summative interpretation could be fitted to more than a few of the impasses. Recognizing the foolishness of my naked assertion—that is, noticing that I had occasioned a pedagogical impasse in a session devoted to making sense of pedagogical impasses—I pulled it back. But I followed it up with the suggestion that the group might consider a "concept study" of number as one of the year's major themes.

As developed by Davis and Renert (2014), concept study is a participatory methodology that blends analytic foci of *concept analysis* (e.g., Usiskin et al. 2003) with the collaborative structure of *lesson study* (e.g., Fernandez and Yoshida 2004). Concept studies are intended as "moments of collective didactical transformation—that is, opportunities to work together to re-form concepts in ways that render those concepts more accessible to learners" (Davis and Renert 2014: 39). In terms of experts and novices, concept studies are collaborations in which experts analyze their now-consolidated understandings of a concept with a view toward recovering the varied experiences and discrete instantiations associated with learning that concept. Among the documented additional benefits of such engagements, participants typically demonstrate enhanced awareness of how multiple mathematical concepts are interlinked within a single-grade level and how single concepts evolve across multiple-grade levels. In addition, such activities can support dispositions toward "open definitions," that is, treatments of concepts that are sufficient to the needs of a specific level of understanding or application, but that anticipate possibilities of future elaboration. By way of immediately relevant illustration, "Numbers are for counting" would be a closed definition in the primary grades, whereas "Numbers can be used for counting" would be more open. One is locked to specific use; the other includes a hint of other possibilities.

As the focus here is on the insights gleaned through concept study rather than its actual processes, I end this section with only brief descriptions of some key activities of our yearlong examination of *number*. Our opening activity was reading and discussing the three opening chapters of Lakoff and Núñez's (2000) *Where mathematics comes from*, focusing in particular on their "four grounding metaphors of arithmetic"—namely, OBJECT COLLECTION, OBJECT CONSTRUCTION, USING A MEASURING STICK, and MOTION ALONG A PATH. These active, body-based notions, they argued and demonstrated, provide sufficient ground to derive and deploy increasingly complex mathematical constructs, ultimately rendering even the most abstract formulations comprehensible. Of course, our discussions stopped well short of such considerations, as we focused much more on the four grounding metaphors (discussed in the next section) than on mathematical logics.

Oriented by that reading, participants reviewed classroom resources such as textbooks, teachers' guides, manipulatives, and games, aiming to identify the metaphor(s) that are foregrounded for learners. This work was accomplished in grade-based group settings. The teachers gathered for themselves the following questions as they examined the artifacts used to frame their classroom practice:

- How are numbers used? What matters are they used to address? What situations are they used to model? What are numbers used to interpret or denote in diagrams, models, and other spatial representations (e.g., as counts of discrete sets, as dimensions, as locations on number lines)?
- What vocabulary is used when using numbers to compare or calculate (e.g., are the concepts of "greater" and "addition" framed in terms of how *many more*, of how *much more*, of how *much bigger*)?
- What applications are used to illustrate and extend the concept at hand (e.g., when discussing "multiplication," is it illustrated in terms of combining like sets, of making sequential leaps, of generating areas)?

- Which number sets are being used and/or made available (e.g., counting numbers, fractions, rational numbers, integers)?
- Is the vocabulary consistent with the images and applications? Are multiple metaphors being invoked? Are there discontinuities or inconsistencies in meaning evident—that is, situations that might trip up a novice?

The teachers realized very quickly that answering these sorts of questions is not easy, as it involves constant interrogation of personal understandings and assumptions. Discussion of this demand prompted one additional question, which proved to be particularly effective in orienting and enabling their efforts:

- What, if any, pedagogical impasses come up when you teach this topic?

More than any of the others, this query enabled teachers' noticing of diverse interpretations of number as it prompted them to pause and wonder about how manipulative materials and recommended lessons might have inadvertently exposed novices to multiple, unreconciled metaphors.

Once the teachers had sufficient time to generate preliminary analyses of the tools and resources in their classrooms, they provided grade-by-grade reports. Unsurprisingly, every group noted inconsistencies and slippages—that is, instances in which inappropriate metaphors were invoked through images or vocabularies, and thus opening possibilities for pedagogical impasses. At the same time, a clear pedagogical arc across Lakoff and Núñez's four grounding metaphors came into focus.

The work just described occupied most of the professional learning time set aside for mathematics. Through the year, we revisited the pedagogical impasses that the teachers brought to the August session, looking to answer the question of just how useful more nuanced understanding of number might be for teachers. I return to those impasses presently.

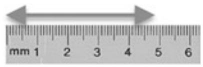
Resolving Some of the Impasses

Some more fine-grained detail of Lakoff and Núñez's four grounding metaphors of arithmetic would be useful before getting into the teachers' follow-up discussions about their reported teaching impasses.

Owing to my focus here on the concept of number within school mathematics, I limit the analysis to entailments for conceptions of number afforded by the grounding metaphors. Some additional entailments for topics beyond number are presented later, but it's important to note that the webs of association and the mathematical power that arises in these webs vastly exceed what is offered here.

In Table 5.1, I interpret the four grounding metaphors in terms of the sorts of number-related questions that learners might ask or be asked. Each is phenomenologically distinct—that is, each invokes a specific cluster of experiences, gestures, and associations. In turn, each calls forward a distinct sense of number.

Table 5.1 Four instantiations of number, associated with Lakoff and Núñez's four grounding metaphors of arithmetic

Lakoff and Núñez's grounding metaphor	Associated metaphor of number	Matter addressed (situation modeled)	An instantiation of "5"
OBJECT COLLECTION	NUMBER AS COUNT	How many?	
OBJECT CONSTRUCTION	NUMBER AS SIZE	How big?	
USING A MEASURING STICK	NUMBER AS LENGTH	How long?	
MOTION ALONG A PATH	NUMBER AS LOCATION	Where?	

As I develop through the analyses that follow, these metaphors proved sufficient for making sense of the pedagogical impasses presented at the start of the chapter. However, they aren't sufficient to span every encounter with number in school mathematics.

Grade 2, Making Smaller

- "When I was brainstorming different ways of saying 'subtract' with my Grade 2s, one of the children got angry that 'make smaller' was on the list. He argued that '*making smaller* can't be *subtracting* since five will still be five no matter how small you make them'."

From the vantage point of different metaphors for number, there's a fairly obvious and highly likely interpretation of this learner's quandary. It would appear that the child was thinking about number strictly in terms of cardinality—NUMBER AS COUNT. Thus, when presented with a remark that was about size—that is, about something that can be made smaller—he applied it to an aspect of the situation that fitted the notion. For him, the actual count of things could not be subjected to the physical process of making things smaller, but the things that were counted could be. In terms of literal meanings, the child was using the entwined notions of *size* and *make smaller* consistently, and the teacher was not. That doesn't mean that the teacher should have immediately perceived the inconsistency, however. Rather, as with most expert knowers, she was likely locked into what Rorty (1991) called a "literalized metaphor" or "dead metaphor—one that has lost its original figurative power by being subsumed into the grander web of associations. It was an instance of coherent expert knowledge that erased the rough inconsistencies of its roots.

In the moment of the teacher's recounting of this pedagogical impasse, I couldn't resist asking what she did next. She responded, "Oh, I just made things worse. I said something like, 'We don't reduce the *size* of the objects, we reduce the *number* of objects'."

> "How'd he respond to that?" I pressed.
>
> "He said something like, 'But if you make the five smaller, it'll still be a five too.' Other kids were starting to laugh, and I didn't want to dig the hole any deeper, so I put a question mark beside 'make smaller' and promised to come back to it later. … I never did get back to it in that lesson. And it never came up again."

What is of especial interest to me in this teacher's follow-up remarks is the indication that, in fact, the child was not working with a rigid, singular interpretation of number. At least two meanings are evident, NUMBER AS COUNT and "number as numeral"—and both the things counted and the numeral, frustratingly, can be reduced in size without altering five-ness. However, only one is a metaphor (i.e., serves to link one category of experience to another) and so only one affords meaning that might be mathematically useful.

The teacher also noticed that the student invoked two different interpretations of five in his final remark. As she explained in a later session, that realization was what prompted her to end the exchange. Ironically, an explicit invocation of a different interpretation of number compelled her to foreclose on an interaction that was triggered by an implicit invocation of an alternative interpretation.

Grade 5, Doubling π

- "I teach Grade 5. Last year, when we were looking at the formula for circumference of a circle, $C = 2\pi r$, one student knocked everything sideways when she asked, 'If π goes on forever, how can you times it by 2?'"

At first hearing, this child's observation that "π goes on forever" might seem to be indicative of the metaphor, NUMBER AS LENGTH. It isn't an instance of this metaphor, however, because π in this interpretation would be the length of the interval between 0 and 3.14159… on a number line. That length can readily be doubled.

What the student was referring to, then, was not the number, but the symbolic representation of the number. While I can't be certain, I would guess that she had applied a NUMBER-AS-COUNT metaphor to the digits in that representation and was troubled by the logical impossibility of applying a digit-by-digit algorithm for multiplication to a number with infinite digits. If correct, then this impasse underscores an issue that came up in the previous impasse. There seems to be a strong disposition among young learners to treat "number" and "numeral" as synonyms—a conceptual move that, I worry, renders number a meaningless operator far too soon in learners' mathematical experiences. I return to this issue later, when I look at interpretations of number beyond those considered by Lakoff and Núñez.

Grade 7, Using Protractors

- "I found it difficult to get my 7th-graders to measure angles. They can't seem to figure out how to use their protractors properly, no matter what I do."

Not a single teacher in the group agreed with me when I first asserted that this impasse had to do with the concept of number. Several countered that the problem was simply that the tool was more complicated than it looked. The multistep process of placing a protractor properly onto an angle, coupled to the demand of choosing the correct scale, renders the skill more an impediment to inquiry than a route to deeper understanding (Fig. 5.1).

However, a key difference of this impasse from the previous two is that it involves a perennial problem—one every middle school teacher I know has encountered. Indeed, participants were confident that the issue would reappear when courses reached the measurement and geometry unit, affording us multiple, staggered opportunities to study the matter.

We thus devoted portions of several of our sessions to digging into the possible contributions of inappropriate interpretations of number to the protractor problem. We began by wondering together what the numbers on the device might mean to novices. When a few teachers took this question to their students, the responses were telling. A majority responded with something like "number of degrees an angle is," and a minority with "how big an angle is." This realization was bolstered by observations of student difficulties when measuring angles that didn't perfectly align with markings on the protractor. Two errors were common. Firstly, students would sometimes count up from the wrong number (e.g., if the leg of an angle fell on 27°, it was not unusual for students to read 153° because they started counting upward from the nearer 150° mark). Secondly, if the leg of the angle felt between unit tick marks, students typically struggled with which whole-number value they should report.

Oriented by these observations, we surmised that students were likely leaning most heavily toward NUMBER AS COUNT, and perhaps somewhat toward NUMBER AS SIZE, rather than the more useful and appropriate NUMBER AS LENGTH. We thus designed a few tasks in which the protractor was introduced in terms of a curved number line, a

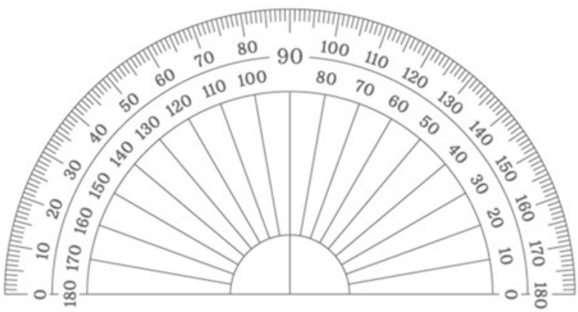


Fig. 5.1 An image of a protractor (included to assist with interpretation of the narrative)

perspective that we hoped might avert the issues noted above highlighting, for example, a meaning of "0" as the starting point for measuring (vs. an interpretation of "nothing" in a count) and offering a more intuitive route into using decimals to estimate angle measures when legs didn't align perfectly with marks. As well, flagging the NUMBER-AS-LENGTH metaphor sponsored an unexpected realization in one sixth-grade class, on the *protractor* itself. It is clearly a length-based notion.

It would be an exaggeration to claim that problems with using a protractor suddenly evaporated. But the tactics were effective. The two teachers who used NUMBER AS LENGTH to frame their entire measurement and geometry units reported dramatic reductions with most of the usual procedural issues, and dramatic increases to the time and energy given to grappling with interesting mathematical issues.

Grade 8, Subtracting Integers

- "My students [in Grade 8] struggle with subtracting integers. They can follow the rule, but no one seems to get why 'adding the opposite' makes sense."

The middle school teachers in this school all made use of a common approach to teaching addition and subtraction of integers in Canadian middle schools, namely using different-colored counters to represent the amounts. For example, going with white for positive and black for negative, his addition statement "−3 + 2" could be interpreted as "3 black chips grouped with 2 white chips." A white and black together make a zero (i.e., +1 + −1 = 0), so when all possible black–white pairs are made and removed, in this instance one black counter remains (i.e., −3 + 2 = −1). Simple.

Things are a tad more complicated for subtraction, however, where a metaphor of SUBTRACTION AS TAKING AWAY is forced onto the situation. That can create problems, especially when minuend and subtrahend have different signs. For example, to accomplish "–3 − 2" using counters, somehow 2 whites have to be extracted from a set of 3 blacks. The common solution is to "add zeroes" (i.e., black–white pairs) to the pile until there are enough of the opposing color to perform the takeaway action. (In the case mentioned, adding two black–white pairs to the original pile of 3 blacks would give a pile with 5 blacks and 2 whites. Removing the 2 whites leaves 5 blacks—or, symbolically, –3 − +2 = −3 + 2(−1 + +1) − +2 = −3 + −2 + +2 − +2 = −5 (Fig. 5.2).)

The link from that chip-based, NUMBER-AS-COUNT representation to an "add the opposite" rule isn't immediately obvious to everyone. Indeed, most of the teachers in our group who were unfamiliar with this strategy saw it as obfuscating rather than

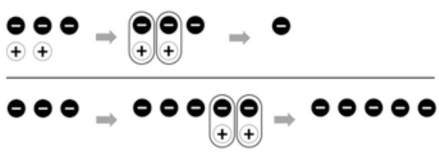


Fig. 5.2 Using number as count to determine "−3 + 2" (shown above) and "−3 − 2" (below)

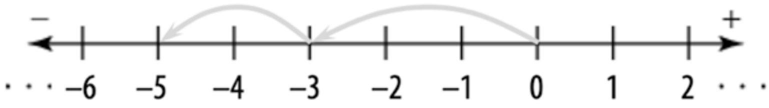


Fig. 5.3 Using number as length to show "−3 + 2" (above) and "−3 − 2" (below)

illuminating. In stark contrast, when a number-line-based approach framing NUMBER AS LENGTH (and, by modest extension, INTEGER AS DIRECTED LENGTH) was offered, the rule "just leaps out at you."

As with the previously discussed impasse, this is one that teachers can typically rely on from one year to the next. Consequently, it was another that could be studied in action, and with similar results. Shifting to a more appropriate metaphor for number didn't do away with the problem, but there were strong indications that a significantly greater portion of the students were able to appreciate the "subtract by adding the opposite" rule as justified and meaningful.

There was also an unanticipated benefit of this approach among the concept study participants. As Lakoff and Núñez highlighted, different grounding metaphors of arithmetic make new number systems available. Irrationals, for example, make no sense when NUMBER IS COUNT, but are readily appreciated when NUMBER IS LENGTH. Our discussion of subtraction of integers helped to drive this point home for participants, with the realization that no modifications were required to accommodate signed rational numbers when addition and subtraction were interpreted as illustrated in Fig. 5.3.

Grade 9, Introductory Algebra

- "For me, the sticking point is algebra. Students can't see the difference between an unknown and a variable. Let's say, for example, we have '$2x = 8$' and '$2x = y$.' My kids would be confused what x as a symbol means and does. If I said '$2x$' and '$2p$,' it wouldn't occur to them that the x and p could be the same. Oh … and when they look at a graph, some students don't seem to see the continuous lines, just the points that line up with the whole numbers."

I almost always cringe internally when someone asks me what I do for a living. When I answer honestly, it's likely that most within hearing range will be eager to communicate their dislike of mathematics—and, worse, that a majority of hearers would have been "really good at math until 6th grade."

These sorts of encounters are so common that mathematics educators actually have a name for them: the "cocktail-party confession." Several years ago, I made an effort to investigate the roots of the phenomenon. After much analysis, I concluded that there are likely two major contributing elements to being "really good at math until 6th grade" … and apparently not after. Firstly, the concept of number becomes much more complicated toward middle school, but learners are rarely made explicitly aware of this shift. As illustrated by the impasses discussed above, many learners enter middle school grades with deeply entrenched—but profoundly inadequate—understanding of NUMBER AS COUNT. This complication becomes perhaps most evident when students are compelled to grapple with new number systems, where they encounter a virtual explosion of interpretations for arithmetic operations. (For instance, there are at least a few dozen instantiations of multiplication at play in a typical grade 6 mathematics curriculum; see Davis 2011.) Small wonder that some learners start to feel like they're missing something.

The second major contributor to the sixth-grade drop-off, in my analysis, is the arrival of full-blown algebra. For learners whose arithmetic is still conditioned by NUMBER AS COUNT, algebra is simply not accessible—as hinted in the multiple elements noted in the above teaching impasse. For many students locked into the NUMBER-AS-COUNT metaphor, the sudden suggestion that letters *not only* substitute for numbers but *are* numbers is simply not a thinkable thought. Further, the presses of algebra toward generalizations and continuities are compromised for learners who dwell in the space of discrete counts. (See Fig. 5.4.) With those sorts of conceptual limitations, it's not surprising that some learners engage with everything that comes after that as meaningless events of rule following.

As signaled in the teacher's description of this impasse, other issues arise (e.g., the conflation of unknowns and variables). For the participants in our group, such issues

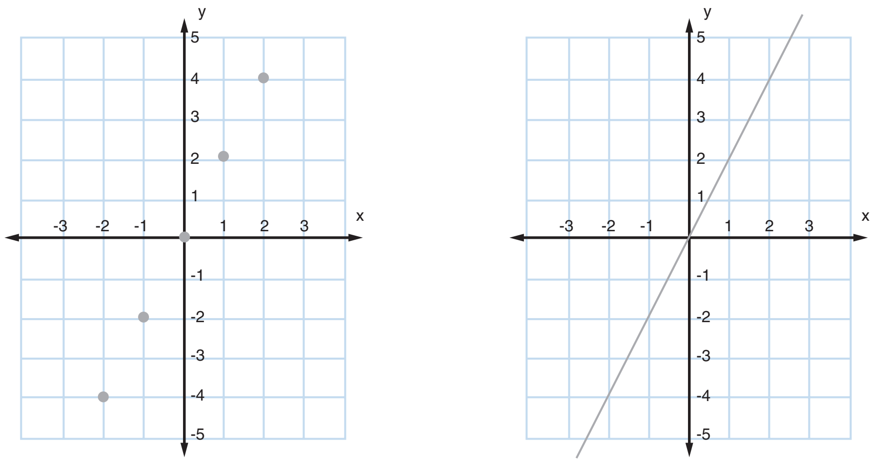


Fig. 5.4 Graphs of $y = 2x$, when x is constrained by NUMBER AS COUNT (left) and when x is continuous (right). (The above teaching impasse suggests that some learners see such graphs as communicating the same information)

served to underscore that the critical question here has to do with identifying which interpretations of number facilitate or enable the leap to algebraic expressions, manipulations, representations, and generalizations. On this matter, the extensive use of number lines and Cartesian grids to represent and interpret algebraic expressions should provide a strong hint: NUMBER AS LENGTH and NUMBER AS LOCATION are clearly the integral elements of the flock of associations that play well together here.

Grade 12, Imagining Imaginaries

- "A Math 30-1 (Grade 12) student asked me why we can't imagine imaginary numbers."

I will defer to Lakoff and Núñez's more nuanced explication on this one, sufficing here to highlight that the metaphor NUMBER AS LOCATION offers a way through this common impasse. Briefly, one must first invoke the commonplace interpretation of "multiplication by -1" as a 180°-anticlockwise rotation of the number line about 0 (mapping a onto $-a$). From that it follows that the square root of -1 (i.e., $(-1)^{1/2}$, or a half of a multiplication by -1) can be interpreted as a 90°-anticlockwise rotation. That rotation generates the complex plane, the horizontal axis of which comprises the real numbers and the vertical axis of which comprises the imaginary numbers. In other words, any imaginary number can be imagined as its location on the vertical axis of the complex plane.

Minding the Gaps

As mentioned earlier, one of the activities in the concept study was a grade-by-grade inventory of the metaphors invoked within the classroom resources. In groups, according to the grades they taught, teachers looked across vocabulary, images, and applications to generate a rough mapping of how numbers were framed for learners at different levels. Their initial impressions are presented in Table 5.2—which, unsurprisingly, suggests an almost exclusive emphasis on NUMBER AS COUNT in the first years of school math, giving some way to a much more varied (and, arguably, conflicted) landscape dominated by NUMBER AS COUNT and NUMBER AS LENGTH by the end of the middle grades.

The group was unsatisfied with this table, however. In fact, they started to express frustration almost immediately when the inventory was undertaken. By the time they were ready to give reports, every person in the room was convinced that the four metaphors we'd listed in the chart, based on Lakoff and Núñez's four grounding metaphors of arithmetic, were insufficient. Other interpretations of number seemed to be at play in school math. In particular, they noted that three frequent encounters with number didn't seem to be included in Table 5.1's categories—namely instances addressing matters of "Which?", "How much?", and "What?".

Table 5.2 Teachers' initial impressions of relative emphases of varied interpretations of number in classroom resources

Grade	Metaphor of number			
	COUNT (How many?)	SIZE (How big?)	LENGTH (How long?)	LOCATION (Where?)
1				
2				
3				
4				
5				
6				
7				
8				

The clear-to-dark shadings indicate absent-to-heavy emphases

After much discussion, the group settled on metaphors of NUMBER AS RANK and NUMBER AS AMOUNT for the first two instances, and a "metaform" (Danesi 2014) of NUMBER AS REIFICATION for the third.

Number as Rank

In our jurisdiction, students are formally introduced to the distinction between using whole numbers to count ("cardinal numbers") and using them to order ("ordinal numbers") early on. Mathematically, cardinal and ordinal numbers can be defined in terms of one another for finite numbers. Experientially, however, they are not the same—which is why the contrast is drawn in elementary school curriculum.

Our group had initially overlooked the ordinal numbers, owing to the naïve assumption of a one-to-one correspondence between our identified metaphors of number and Lakoff and Núñez's grounding metaphors of arithmetic. However, that omission was immediately evident when we delved into curriculum materials. Many lessons in the first few grades were really on the distinction, and one of its primary markers was a shift from questions phrased in terms of "how many" to tasks requiring learners to attend to discrete ranks—involving, for example, levels, positions in groups, and ordered sequences.

We settled on metaphor NUMBER AS RANK to refer to this instantiation. Other options included PLACE and S POSITION, but we worried these were too similar in everyday meaning to the LOCATION. As well, we felt that the notion of RANK better served to underscore the discrete character of ordinal numbers.

Number as Amount

In the early stages of our concept study, the discrete–continuous distinction emerged as a very useful and frequently invoked idea. While all participants would have encountered it somewhere in their histories with formal mathematics, it was received as new by most—and, in fact, was a site of struggle for many.

In an effort to render the distinction accessible, I sidestepped formal definitions and suggested two rules of thumb:

- If the situation involves counting, it's discrete; if it involves measuring, it's continuous.
- If it's grammatically correct to say "fewer" in the situation, it's discrete; otherwise, it's continuous.

While imperfect, those guidelines served us well across the analyses of number behind Tables 5.1 and 5.2. However, they're inadequate around a few applications that are frequently encountered in grade school mathematics, especially ones involving money. As a fourth-grade teacher expressed the issue, during our grade-by-grade inventory of interpretations:

> They're [i.e. situations involving quantities of money] discrete, right? We *count* money. But we don't say, "How *many* does this cost?" We say, "How *much*?" … And we never use "fewer" when we're talking about money.

Her colleague added:

> We noticed kind of the same thing with the way fractions are introduced in the Grade-4 book. Most of the exercises are based on counting—like [holding up the exercise book] this picture where five out of six balls are colored in that asks "What fraction is shaded?" That's not a "How many?" or a "How big?" question, it's a "How much?" question.

Much more time was given to mulling over the matter, but these teachers' remarks seem to sum up an important experiential truth: In many contexts and occasions where questions of "How much?" are asked, number can be deployed as discrete but sensed as continuous. In our analyses, the most common of these situations involve large quantities and/or discrete fractions (including terminating decimals).

Since we were immersed in a discussion of orienting metaphors, it was no surprise that interests turned to identifying an analogy that fitted this situation. Suggestions immediately gravitated to notions of "piling up" and "clustering," at which point some hasty googling of original word meanings prompted us to suspect that others had long ago grappled with a similar issue. It turns out that English has several terms that invoke precisely the same images that the teachers had suggested to address matters of "How much?" such as *amassing* (e.g., a fortune), *amounts* (e.g., owed), and *accumulating* (e.g., parts into a whole). According to the Online Etymology Dictionary (etymonline.com), all three of these have to do with mounding bits into larger unities:

- *Amass* derives from the Old French *à* "to" + *masse* "lump, heap, pile."
- *Amount* derives from the Latin *ad* "to" + *monten* "mountain."
- *Accumulate* derives from the Latin *ad* "to" + *cumulare* "heap up."

The group thus settled on the metaphor NUMBER AS AMOUNT for this new category.

Importantly, there is no suggestion here that NUMBER AS AMOUNT has the same epistemic status as the metaphors for number presented in Tables 5.1 and 5.2. Rather, as hinted by the ambiguity experienced around the discrete–continuous distinction, this metaphor is better seen in phenomenological terms than mathematical terms. It is something encountered in day-to-day applications (mainly situations involving discrete fractions, such as money)—and so while it doesn't appear to be integral to concepts in pure mathematics, it is certainly important in school mathematics. We also concluded that it likely plays an important conceptual role in bridging discrete, quantity-focused (NUMBER AS COUNT) and continuous, magnitude-focused (esp. NUMBER AS SIZE) conceptions of number.

Number as Reification

A somewhat more surprising observation for the teachers was how early and how often classroom resources invoked numbers and posed questions in complete absence of interpretive referents—that is, asked "What?" questions in which numbers were presented as naked operators. Practice exercises devoid of metaphorical anchors were already evident in first grade, and they represented the most common variety by the middle school years. Indeed, the steady increase in proportion of "What?" questions was taken by the teachers as evidence of a systematic process to wean learners from specific, meaningful but necessarily limiting interpretations of number.

In some regards, this progression should have been expected, especially given that the group's analyses of several pedagogical impasses homed in on children's habits of equating "number" and "numeral" in even the lowest grades. Clearly, something is pressing learners toward seeing numbers and things in and of themselves. Nonetheless, we experienced the realization as disconcerting. As was evidenced multiple times in our concept study, a premature compulsion to treat numbers as only symbolic operators—or, worse, as symbols—can debilitate efforts to interpret and extend mathematical concepts. Conversely, failure to elaborate number into a symbolic operator might be similarly debilitating at higher grades, as illustrated by those pedagogical impasses in which learners' interpretations of number didn't keep pace with the increasingly abstract nature of the concepts under study.

There is no quandary here. Humans' understandings of number are both embodied (i.e., rooted in bodily based experiences) and embedded (i.e., called forward in culturally meaningful situations). It is entirely reasonable to expect school mathematics to be structured in a way that draws on and nurtures the former while anticipating and enabling the latter. The issue isn't whether school mathematics should channel learners toward a consolidated concept of number, but how and when it should happen. Such matters, in turn, can only be settled through nuanced appreciations of how integrated concepts emerge and what the integrated concept is expected to do.

Two ideas proved helpful around these concerns. Firstly, one of the participants called attention to Fauconnier and Turner's (2003) research into conceptual blends. While not of immediate pragmatic value to the group, it was affirming to read their characterizations of the emergence of new and more powerful discursive objects through combinations and mash-ups of existing ones. In this regard, we found Danesi's (2014) notion of *metaform* to be very useful. Contrasted with a metaphor, a metaform is an abstract distillation—a fusing that foregrounds common functional elements while suppressing idiosyncratic and potentially dysfunctional elements. In the process, the metaform can be experienced as nonspatial and acausal—as an idea that is unencumbered by the interpretive specificities of a metaphor or a cluster of metaphors. In this sense, a metaform of number would be what Hilbert (1928: 470) dubbed an "ideal object," which "in themselves mean nothing but are merely things that are governed by our rules."

"So basically we should be drawing attention to metaphors to teach a metaform so that students don't have to rely on the metaphors," one participant summed it up, to the general approval of the group. His thought prompted the suggestion from another participant that we should name the metaform under discussion, in order to distinguish it from the clutter of meanings for number than we'd encountered. Ultimately, we settled on NUMBER AS REIFICATION—an imperfect choice, but one useful for underscoring why number is so often engaged as an object.

Three other choices figured prominently in our protracted discussion of what to call the metaform of number: NUMERAL, OBJECT, and OPERATOR. The first two were rejected because although we aimed to flag the "thing-ness" of the metaform, we also wanted to signal its emergent character. The third, NUMBER AS OPERATOR, was initially compelling because of its current prominence in efforts to incorporate computational thinking into school curriculum. In computer-coding contexts, an *operator* is a logical symbol. That is, an operator is not simply a numeral; it represents an action or a process. While that particular meaning seemed fitting, we decided to set it aside because of the explicit and deliberate meaninglessness (in phenomenological terms) of computer- and computation-based number usage.

Revising the Maps

With the three addition interpretations of number distinguished and named, the group undertook to elaborate Table 5.1 into Table 5.3, the contents of which hint at considerably more discussion and debate than I have reported here. I'll leave it to you to explore whatever bits you might find interesting, signaling here that I personally disagree with some significant elements. But I present this evolving analysis in its "current" form, as an indication of participants' ongoing efforts to interpret and represent *number* in manners that enable their classroom practice.

The group also redid their analysis of relative emphases of varied interpretations of number in classroom resources (Table 5.2) to include the three additional meanings (Table 5.4). Whereas adding of the "AMOUNT" and "RANK" columns was uneventful,

Table 5.3 Seven instantiations of number in school mathematics, along with some illustrative entailments

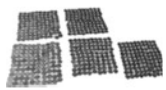
Metaphor or metaform of number	Matter addressed (situation modeled)		Associated grounding metaphor(s) of arithmetic	An instantiation of "5"	How "less" "greater" tend to be expressed		How addition tends to be seen	Some encounters/ contexts/uses	Numbers made available
COUNT	Quantity	*How many*? (discrete)	OBJECT COLLECTION		Fewer	More	Combining sets	Counting; sorting; clustering	Whole N°s; natural N°s; cardinals
RANK		*Which*? (discrete)	OBJECT COLLECTION		Ahead	Behind	Changing rank	Sequencing; ranking; grading	Ordinals
AMOUNT		*How much*? (discrete, but experienced as continuous)	OBJECT COLLECTION & CONSTRUCTION		Less	More	Pooling amounts; amassing	Pricing; accounting; apportioning	Large numbers; discrete fractions
SIZE	Magnitude	*How big*? (continuous object)	OBJECT CONSTRUCTION		Smaller	Larger	Growing; joining pieces	Assembling; sharing; ratios	Continuous fractions
LENGTH		*How long*? (continuous dimension)	USING A MEASURING STICK		Shorter	Longer	Extending; moving farther	Scale-based measuring; traveling	Rational N°s; irrational N°s; integers
LOCATION	Entity	*Where*? (discrete site in continuous space)	MOTION ALONG A PATH		Left of (or lower)	Right of (or higher)	Shift in location	Locating; scheduling; reading time	Real N°s; imaginary N°s; complex N°s
REIFICATION		*What*? (disentangled from physical instantiations)	All of the above	5	<	>	Binary operation	Symbolic manipulation; computing	Any/all of the above

Table 5.4 Teachers' elaborated impressions of relative emphases of varied interpretations of number in classroom resources

Grade	Metaphor or metaform of number						
	COUNT (How many?)	RANK (Which?)	AMOUNT (How much?)	SIZE (How big?)	LENGTH (How long?)	LOCATION (Where?)	REIFICATION (What?)
1							
2							
3							
4							
5							
6							
7							
8							

The clear-to-dark shadings indicate absent-to-heavy emphases

compiling the "REIFICATION" column was especially fraught. As participants analyzed "What?" questions in classroom resources, they frequently struggled with the imagined intentions of textbook authors. Most often, they concluded that "What?" questions weren't framed in a way that enabled and compelled learners to consolidate their evolving conceptions of number. Rather, they most often seemed to be presented as attempts to wean learners off physical referents by ignoring (rather than inviting) processes of differentiation, bridging, and consolidation of varied instantiations. Consequently, for participants, the final column of Table 5.4 points more to "opportunities to develop NUMBER AS REIFICATION" than as actual attempts to prompt learning in that direction.

And So …?

It goes without saying that one might expect a strong emphasis on pragmatics when engaging with educators in a concept study during the school year in one of their classrooms. While our inquiry reached into philosophy, mathematics, cognitive science, and other domains, the gravitational pull of classroom practice ensured that every one of our discussions included considerations of how to teach. Some of these considerations were broader, especially with regard to the arc of the K–12 school mathematics experience. But most were more immediate, ultimately being articulated as seven principles to guide mathematics teaching:

- Whenever dealing with number-related topics, be explicit about what the numbers are being used to do/represent/model and the metaphor(s) at play.

- Phrase explicit statements about "what numbers are" as open definitions. Remind learners that almost all mathematical definitions will be elaborated as their understandings grow.
- When moving from one metaphor to another (e.g., when shifting from an application involving counting to one involving measurement), signal the shift and provide interpretive bridging if appropriate.
- "Multiple representations" and "personal strategies"—two major emphases in the popular professional literature at the moment—should never be encouraged for their own sake. Treat them as opportunities to explore underlying metaphors, as well as obligations to locate varied representations and strategies in the grander matrix of interpretations.
- Explore clusters of association for each grounding metaphor separately. That is, rather than brainstorming synonyms for "subtract," brainstorm synonyms for "subtract" that go along with NUMBER AS COUNT, with NUMBER AS SIZE, and so on—once again providing interpretive bridging if appropriate.
- Be mindful of—and draw attention to—how binary operations can introduce many more categories of interpretation (e.g., familiar interpretations of multiplication include "set of sets" [i.e., COUNT × COUNT = COUNT], "repeated hops" [i.e., COUNT × LENGTH = LENGTH], and "dimension crossing dimension" [i.e., LENGTH × LENGTH = 2d SIZE]; several other combinations will be encountered by all middle school students).
- Be cognizant that the ultimate goals are nimble, consolidated but flexible concepts that are enabled by the richness of diverse interpretations, but unencumbered by the limitations of any singular interpretation.

Our group homed in on some of these principles early on. In fact, the first three were articulated during our discussion of Lakoff and Núñez, in the session that followed our first foray into pedagogical impasses. The other principles arose one at a time, as different aspects of the concept study unfolded.

Not insignificantly, over the course of the concept study, there were also marked shifts in how participants talked about both mathematics and learning of mathematics. On the topic of disciplinary knowledge, an analysis of the year's transcripts revealed that there was a decline in teachers' use of object-based metaphors and a corresponding increase in growth-, evolution-, and systems-based notions to refer to mathematics. Regarding learning, acquisition- and journeying-based metaphors were dominant at the start of the year. While still prominent at the end, they had yielded considerable ground to notions of sense making, construal, and coherence seeking.

There was also a notable shift in teachers' descriptions of their own work. Analyses of their self-references during the session devoted to teaching impasses revealed overriding desires to provide learners with "clear" explanations, "best" illustrations, and "correct" information to "facilitate" learning. Much in contrast, references to such concerns are virtually absent in the transcripts from the end of the year. Instead, as might be inferred from the above list of principles, direct references to teaching were couched in terms of responsiveness and adequacy, as notions of

TEACHING AS FACILITATING (from French *faciliter* "to make easy") gave way to a dominant sensibility of TEACHING AS CHALLENGING.

As for the project's focus on mathematics for teaching, our informal definition—that is, "what an expert needs to know to think like a novice"—proved particularly useful throughout the concept study. Evident in their shifts in thinking on the nature of mathematics and the processes of mathematics learning—and underscored by the realization that both phenomenology and mathematics must be consulted in quests for meanings and metaphors—our inquiry into number clearly and strongly demonstrated that mathematics knowledge and mathematics learning are not separate topics for educators.

References

Danesi, M. (2014). Metaphors and metaforms: Interconnecting language, culture and cognition. In: G. Gobber and A. Rocci (eds.), *Language, reason and education* (pp. 87–102). New York: Peter Lang.

Davis, B. (2011). Mathematics teachers' subtle, complex disciplinary knowledge. *Science* 332: 1506–1507.

Davis, B., and Renert, M. (2014). *The math teachers know: Profound understanding of emergent mathematics*. New York: Routledge.

Ericsson, K. A., Charness, N., Feltovich, P., and Hoffman, R.R., (2006). *Cambridge handbook on expertise and expert performance*. Cambridge, UK: Cambridge University Press.

Fauconnier, G., and Turner, M. (2003) *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.

Fernandez, C., and Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning*. Mahwah, NJ: Erlbaum.

Hilbert, D. (1928/1967). The foundations of mathematics (English translation of Die Grundlagen der Mathematik). In: J. van Heijenoort (ed.), *From Frege to Gödel: A source book in mathematical logic, 1897–1931* (pp. 464–479). Cambridge, MA: Harvard University Press.

Lakoff, G., and Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books

Rorty, R. (1991). *Objectivity, relativism, and truth: Philosophical papers, volume 1*. New York: Cambridge University Press.

Sierpinska, A. (1994). *Understanding in mathematics*. London: Falmer.

Usiskin, Z., Peressini, A., Marchisotto, E.A., and Stanley, D. (2003). *Mathematics for high school teachers: An advanced perspective*. Upper Saddle River, NJ: Pearson.

Chapter 6
The Body of/in Proof: An Embodied Analysis of Mathematical Reasoning

Laurie D. Edwards

Mathematics is a way of using the mind with the goal of knowing the truth, that is, of obtaining certainty

(William Byers)

Mathematics is, in part, a search for structure and certainty, and the primary practice of mathematicians toward that end consists of creating and communicating mathematical proofs. The purpose of this chapter is to examine mathematical proof and logical reasoning from the perspective of embodied cognition (Clark 1998; Gibbs 2005; Johnson 1987, 2012; Lakoff and Johnson 1980; Shapiro 2010, 2014; Varela et al. 1991). Although there are multiple theoretical and methodological approaches to embodiment, Varela (1999: 11–12) characterized its essential aspects as follows:

> Embodiment entails the following: (1) cognition dependent upon the kinds of experience that come from having a body with various sensorimotor capacities, and (2) individual sensorimotor capacities that are themselves embedded in a more encompassing biological and cultural context … sensory and motor process, perception and action, are fundamentally inseparable in lived cognition.

In recent decades, researchers have investigated how the body is implicated in tasks ranging from remembering personal experiences to group collaboration, and in contexts including language learning, science, music, and emotions (Shapiro 2014). This attention to the body has started to break down a long-standing paradigm that viewed cognition as amodal and abstract, based solely "in the head." The role of the body in mathematical thinking, learning, and teaching specifically has been addressed in a range of settings, from young children learning to count objects (Kiefer and Trumpp 2012) to adults teaching about differential equations (Rasmussen et al. 2004, see also Abrahamson and Lindgren 2014; Edwards 2009, 2010, 2011; Edwards et al. 2014; Hall and Nemirovsky 2012; Radford et al. 2009). This "turn to the corporeal" (Rotman 1993) has enriched and deepened our understanding of

L. D. Edwards (✉)
St. Mary's College of California, Moraga, CA, USA
e-mail: ledwards@stmarys-ca.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_6

mathematical thinking, revealing that the body is thoroughly involved in the learning and doing of mathematics. In addition, attention to embodiment has broadened the focus within mathematics education research beyond written symbols, images, and oral speech to include modalities such as gesture and other bodily movements, eye gaze, rhythm, and prosody (Edwards et al. 2014; Hall and Nemirovsky 2012).

Within an embodied cognition framework, mathematics is not seen as a transcendental, formal collection of rules and patterns, unrelated to everyday thinking and experience, but instead as a human intellectual product, one which develops both historically as a discipline over time and ontologically as it is constructed by an individual learner. It is socially constructed, but not in an arbitrary way, being both constrained and enabled by the biological capabilities and physical situatedness of human beings. Embodiment does not negate the influence of social interaction and culture; rather it grounds it in shared biological constants (Hall and Nemirovsky 2012; Núñez et al. 1999). As stated by Hall and Nemirovsky (2012: 212), "We think of concepts (in mathematics but also in other domains) as forms of modal engagement in which bodies incorporate and express culture."

In the current analysis, we will investigate the possible conceptual roots or cognitive precursors to mathematics and deduction, taking into account existing research about early cognitive development. In addition, we will examine multimodal interactions among emerging mathematical experts engaged in creating a new proof, with the goal of better understanding the relationship between the body and this essential area of mathematical practice.

Mathematics, or at least "advanced" mathematics such as proof, has at times been seen as very different from other kinds of thinking. For instance, Manin stated, "The gaping abyss between the habits of our everyday thinking and the norms of mathematical reflection must remain intact if we want mathematics to fulfill its function" (Manin et al. 2007: 36). A great deal has been made of the "abstract" and "formal" nature of advanced mathematics, with some scholars setting apart a separate "world" of formalism, distinct from simpler mathematical worlds connected to either embodiment or symbolism (Tall 2008). While mathematics, as a form of discourse and practice, certainly has its own unique characteristics, including the non-ostensive nature of its objects, one of the goals of this chapter is to highlight ways in which mathematical cognition and communication are closely connected to everyday, nonmathematical thinking and language.

One of the central principles of embodied cognition is that of continuity; under this principle, even complex thought such as mathematical proof is seen as arising from simpler conceptual mechanisms (Johnson 2012). Embodiment theory and the principle of cognitive continuity motivated the current work: to investigate how the specialized kind of thinking and communication involved in mathematical proof is related to other kinds of human thought and activity. Following scholars like Johnson (2012), we propose that deductive proof and logic are built from the same basic conceptual building blocks as are more mundane kinds of thought. An important analytical tool used to explore this proposal is found in cognitive linguistics, which has identified many such building blocks, such as mental spaces, image schemata, metaphors, and conceptual blends (Dancygier and Sweetser 2005; Evans and

Green 2006; Fauconnier and Turner 2002; Johnson 1987, 2012; Lakoff and Johnson 1980). A central tenet of cognitive linguistics is that language is not a collection of formal rules and productions, where form is unrelated to content, but is based on primary embodied experiences as well as collections of unconscious mental mappings linking together familiar experiences and ideas in order to create new ones. These groupings of familiar concepts or experiences are known as mental spaces or input spaces (Fauconnier 1994; Fauconnier and Turner 2002). An important mechanism within this framework is the blending or integration of input spaces. As defined by Fauconnier and Turner (2002: 89), "Conceptual integration … connects input spaces, projects selectively to a blended space, and develops emergent structure."

Conceptual integration can be seen as a general mechanism that encompasses more specific mappings such as conceptual metaphor (Fauconnier and Lakoff 2009). Whereas conceptual integration may link two or more input spaces to create blended space, conceptual metaphor projects the logical structure of a single input space (known as the source domain) onto a single target domain (Lakoff and Johnson 1980). Conceptual metaphors have been used in the analysis of mathematical ideas ranging from arithmetic to calculus (e.g., Bazzini 1991; Lakoff and Núñez 2000; Núñez et al. 1999). Another important element of cognitive linguistics is the notion of image schemata. An image schema "is a condensed redescription of perceptual experience for the purpose of mapping spatial structure onto conceptual structure" (Oakley 2007: 215). An example of an image schema that undergirds an important mathematical idea is that of a physical container. Containers have an inside that contains objects, an outside, and a boundary. It is unlikely that we could understand the idea of a mathematical set without having had many experiences with putting objects into, and taking them out of, physical containers, experiences that allowed us to build up this image schema (Lakoff and Núñez 2000; Mandler 2004).

This analysis will examine proof using an embodied cognition framework as well as the tools of cognitive linguistics. Following Johnson, the goal is to demonstrate that "we do not have two kinds of logic, one for spatial-bodily concepts and a wholly different one for abstract concepts. There is no disembodied logic at all. Instead, we recruit body-based, image-schematic logic to perform abstract reasoning" (Johnson 2012: 181).

This chapter utilizes existing research as well as newly analyzed data to sketch possible elements of a "body-based, image-schematic logic" underlying mathematical reasoning and proof. Framing questions include the following: In what ways is mathematical proof continuous with other kinds of thinking, both developmentally and cognitively? What are plausible conceptual underpinnings that help people understand the notion of mathematical proof? Can language and gesture provide evidence for continuity in the relationship between proof and the body?

I will argue that, indeed, there is continuity between the practice of mathematics and other kinds of socio-cognitive practices. In particular, I will argue for continuity at several levels. These include the following:

- Continuity between the ways that proof is described and discussed and basic embodied physical experiences, as indicated by unconscious conceptual mappings or metaphors. The evidence for these foundational mappings for proof will also include nonverbal modes of communication such as gesture.
- Continuity between the notions of physical causality constructed in early infancy and childhood and the way that we think about inductive and deductive logic.
- Continuity in the way we express epistemic or logical conditionals, verbally and gesturally, between mathematical and nonmathematical discourse settings.

The overall goal is to demonstrate that, far from being an arcane form of thinking unrelated to everyday concerns, mathematical proof is grounded in very basic human experiences. Although the practice of proving has certainly been tested, refined, and constrained over time by the mathematical community, proof is as embodied as any other powerful concept utilized by human beings as they explore patterns, create new structures, and test regularities.

Mathematical Proof and Logical Deduction

Mathematical proof has been characterized as an explanation accepted by a community of mathematicians (Balacheff, cited in Hanna 1990: 9); more formally, proof has been defined as:

> [A] finite sequence of sentences such that the first sentence is an axiom, each of the following sentences is either an axiom or has been derived from preceding sentences by applying rules of inference, and the last sentence is the one to be proved (Hanna 1990: 6).

There is a rich body of research within the field of mathematics education about proof (Lin et al. 2009), whether addressing how it is taught and learned (Balacheff 1991; Nardi and Knuth 2017), the impact of interactive technologies (Chazan 1993; Laborde 2000; Roy et al. 2017), or differences in proof schemes and the use of logic (Harel and Sowder 2007). This chapter is concerned not with the teaching or learning of proof, but with how people conceptualize the process of proving, and the product of that process, a proof. The analysis will take a top-down approach, first examining ideas about proof as a whole, and then addressing the fundamental elements of proof, namely, logical deductions and conditional statements. At each level, we hope to demonstrate continuity and connections between embodied human experience and what is often considered to be the most abstract of mathematical activities.

How Is Proof Conceptualized?

In order to examine how people think about proof, we will look at how they express themselves about this subject, drawing on both written and oral language, as well as examining physical gestures that occur in conjunction with speech. This analysis is based on foundational work in cognitive linguistics (Evans and Green 2006) which holds that language reflects deep conceptual structures and mechanisms, including image schemas (Talmy 1988), metaphors (Johnson 1987, 2012; Lakoff and Johnson 1980), and conceptual blends between mental spaces (Fauconnier 1994; Fauconnier and Turner 2002). By examining the particular images, language, and gestures used to describe mathematical proof, we can infer the nature of these underlying structures.

Two examples of descriptions of proof in written texts are given below, the first by a mathematics educator and the second by a "working" mathematician:

Example 1: A proof is a transparent argument, in which all the information used and all the rules of reasoning are clearly displayed and open to criticism. It is in the very nature of proof that the validity of the conclusion *flows from* the proof itself, not from any external authority (Hanna 1995: 46, emphasis added).

Example 2: I'd like to spell out more what I mean when I say I proved this theorem. It meant that I had a *clear and complete flow of ideas*, including details, that withstood a great deal of scrutiny by myself and by others (Thurston 1994: 175, emphasis added).

From the point of view of cognitive linguistics, the terms used to refer to the abstract process of proof are not accidental or arbitrary, but instead point to underlying conceptual structures. Both of the texts above utilize a specific term, "flow." Under the current analysis, the use of the term "flow" is evidence for an image schema based on a universal perceptual experience, the sight and feel of moving water. In the "flow" image schema, water moves in one direction, with a certain amount of force, in a continuous stream. This image schema is utilized by the writers above precisely because it contains elements that correspond to the way that they conceptualize mathematical proof: either as a unidirectional force that links the conclusion to the proof or as a "stream" of clear, complete, and convincing ideas, one following the next.

The image schema of "flow" is related to a conceptual metaphor that structures our understanding of thought in general. Sweetser identifies this metaphor as "THOUGHT (or REASONING) IS MOTION THROUGH SPACE," and gives as an example the phrase "We don't seem to be getting anywhere" (1998, paragraph 10). An example of this metaphor can be found in way the late mathematician Maryam Mirzakhani described what it is like to work on a difficult proof: "It is like being lost in a jungle and trying to use all the knowledge that you can gather to come up with some new tricks, and with some luck you might find a way out" (*New York Times*, July 18, 2017). In this example, she describes her reasoning process in finding a proof as motion through space, moving from a location in a confused, junglelike place to a place of more clarity outside the jungle. As we will see below, the metaphor of reasoning as motion through space can be seen expressed via speech, written language, and gesture.

Evidence from Emerging Mathematicians

In addition to looking at extant written definitions and other texts addressing mathematical proof, data were gathered from a group of successful mathematics learners that have not received a great deal of attention within mathematics education research. These are doctoral students in mathematics; these students can be seen as emerging experts in mathematics who have gained a base of foundational knowledge in the field and are in the process of learning to create original proofs of their own (Marghetis et al. 2014). As such, they are knowledgeable about the process of proving, as well as being familiar with different kinds of proof and even how to teach mathematical proof.

A total of 12 doctoral students in mathematics at a major research university in the United States participated in pairs in a qualitative study based on a 90-min clinical interview (Edwards 2010). The interview consisted of three parts: an initial set of questions about their experiences with mathematical learning, teaching, and proof; a second part in which they had 40 min to work together to prove an unfamiliar conjecture; and a final segment in which they were asked to evaluate a visual "proof."

During the second, proof-finding part, the students looked for a proof of the following conjecture:

> Let f be a strictly increasing function from [0, 1] to [0, 1].
> Prove that there exists a number a in the interval [0, 1] such that f(a) = a.

The sessions were videotaped (the interviewer left the room while the students were working together on the proof) and later transcribed and gestures annotated. As with this chapter, the goal was to investigate how the students conceptualized mathematical proof.

Metaphors for Proof

When directly asked what a proof was, the doctoral students gave descriptions which were similar to those cited above by Hanna. For example:

- AC: It's a set of logical reasoning that begins with a premise and leads to a conclusion.
- AW: I would say it's just, you know, a well-thought-out sequence of steps that nobody would refute …. In practice, it's just—it—a very, very solid argument in which each step proceeds logically from the last.
- AS: A rigorous proof would be based on the axioms of mathematics that we've set up …. Actually, following it step by step so that your conclusion always follows from some kind of logical steps.

These definitions include elements of the metaphor that conceptualizes reasoning as motion through space, as shown in terms like “leads to,” “sequence of steps,” and “following it step by step” (although it could be argued that the “steps” of a proof that the students describe are a “fossilized” or dead metaphor that no longer have the connotation of movement, but instead refer to written lines on a page; see Edwards 2010).

In fact, the students’ more casual discourse about proof provides even clearer evidence of a motion-based metaphor grounding their understanding, as shown in the following phrases:

> “the destination,” “the forward direction,” “walking it back,” “you want to end up over here,” “you get kind of bogged down,” “you get to a certain point,” “I don’t wanna go any further,” “we’ll try the other way,” “maybe I don’t know where I’m going,” “at some point, maybe I can, like you know, see the goal,” “there’s so many ways you could go,” “the better way to go”

These phrases reveal a more specific metaphor than the general REASONING IS MOTION THROUGH SPACE identified by Sweetser. The new metaphor incorporates an image schema called “source-path-goal” (Lakoff and Johnson 1980; Johnson 1987; Talmy 1988). This image schema is based on the universal embodied experience of moving oneself from a particular starting location to a specific destination, via a path. This metaphor for proof is called A PROOF IS A JOURNEY (Edwards 2010). In this metaphor, rather than an undefined motion through space, a proof has a specific source or starting point (a premise or set of givens), a goal (a conclusion or that which is to be proved), and a set of steps, each of which needs to be logically valid (and each of which should lead toward the goal). Thus, the defining elements of a mathematical proof are linked to elements with similar roles in a directed, physical journey. This mapping is illustrated in Fig. 6.1, which shows the source-path-goal schema as applied to mathematical proof, and Table 6.1, which spells out the conceptual mapping between the source and target domains in the PROOF IS A JOURNEY metaphor.

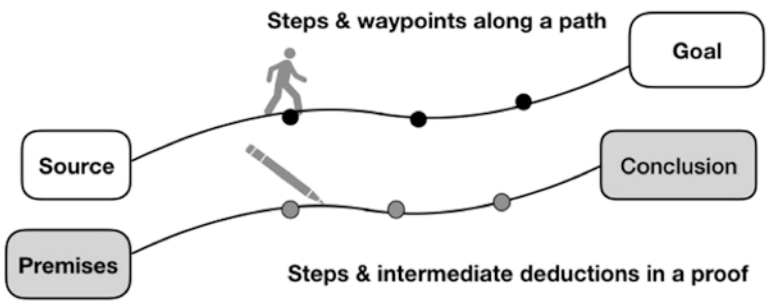


Fig. 6.1 The source-path-goal image schema applied to proof

Table 6.1 The PROOF IS A JOURNEY metaphor

A PROOF IS A JOURNEY	
Source domain: physical journey	Target domain: mathematical proof
• Starting point (source) • Destination (goal) • Steps • Possible sequences of steps (paths) • "Dead ends" or wrong paths that don't reach the destination • Obstacles to finishing the journey	• Premises (source) • Conclusion (goal) • Logical statements ("If A then B") • Possible sequences of logical statements (paths) • Sequences that don't result in the desired conclusion • Obstacles to completing the proof

Hanna (2000) quotes two mathematicians who use this JOURNEY metaphor explicitly when discussing the nature of proof:

> Rav suggests we think of proofs as "a network of roads in a public transportation system, and regard statements of theorems as bus stops." A similar metaphor is used by Manin (1998) when he says that "Axioms, definitions and theorems are spots in a mathscape, local attractions and crossroads. Proofs are the roads themselves, the paths and highways. Every itinerary has its own sightseeing qualities, which may be more important than the fact that it leads from A to B" (Hanna 2000: 7).

In using the metaphor of proof as a journey, Manin elaborates it to include the notion of "sightseeing qualities," that is, the idea that some proofs can explain and enlighten, rather than simply provide a technical justification of the given conjecture, a thesis promoted by Hanna among others (Hanna 1990, 2000).

Evidence from Gesture

In addition to the evidence from written and spoken language above, the use of co-verbal gesture by one of the doctoral students also revealed the metaphor of PROOF IS A JOURNEY, as shown in Fig. 6.2, a still from one of the videos in the doctoral student study.

In the video, the student WG says the following, in conjunction with the indicated gesture:

> 'cause you start figuring out, I'm starting at point A and ending up at point B. There's gonna be some road - where does it go through? And can I show that I can get through there?

WG begins the gesture sequence by touching a location near the top of his thigh ("point A"), and then points as he moves his right finger away from his body ("point B"). He then returns to his initial location and traces a fairly straight path outwardly through the air with his finger, pausing briefly after saying, "some road." He then makes a small horizontal circle with his outstretched hand, and finishes by retracing the path between the origin and end of the gesture a second time.

This gesture sequence clearly shows a physical motion through space as the student discusses proof, but the student also marks, verbally and spatially, specific

Fig. 6.2 Gesture illustrating PROOF IS A JOURNEY metaphor

locations corresponding to the premises/start of the journey ("point A"), and to the conclusion/end of the journey ("point B"). His spoken words also refer to "where the road goes." In the metaphor, the "road" is a sequence of logical statements, and the student also asks whether he can reach the desired conclusion ("get through there"). Thus, the gestures and speech both contribute to a multimodal expression of the "journey" metaphor for mathematical proof.

Conceptual Roots of Logical Deduction

Given that proofs are defined, in part, as sequences of logical statements, in order to understand proof more deeply from an embodied perspective, the next question we address is this: What are the conceptual roots of logical deduction? That is, how is it that people are able to understand and create logical statements of the form, "If P is true, then Q is also true," and use them in proofs?

There is a body of scholarship investigating the development and functioning of logical thinking, with notable early work by Piaget and Inhelder (1964) and including more recent work based on cognitive science (e.g., Best 2005; Johnson-Laird 1999; Kahneman e al. 1982). Although this research has identified changes in logical thinking within childhood and described errors in deductive reasoning, for the most part it was carried out with a view of reasoning as a purely mental facility, unconnected to the physical body.

From the point of view of embodied cognition, however, the concept of logical necessity, like other ideas, is ultimately founded on bodily/perceptual experiences (Mandler 2004; Talmy 1988). In addition, rather than seeing different kinds of reasoning as arising within age-demarcated phases, current research again identifies a continuity between the thinking of very young children and adults:

> Right from the beginning, or at least from a few months of age, babies function in ways that merge continuously into those of older children and adults. They form concepts, they have notions of different kinds, they generalize from their experience on the basis of concepts they have already formed … (Mandler 2004: 11)

It is this capacity for generalization and pattern noticing in young children that, we argue, forms the basis for the development of logic. From an evolutionary perspective, the ability to learn from experience by noticing (creating) patterns and drawing conclusions was essential to human survival; here, we focus on its development in children and later application to mathematics.

The argument is as follows: the development of an understanding of logic (where a premise necessarily "leads to" a conclusion) is based on the physical experience of causality (where one event causes the occurrence of a second). The concept of physical causality, constructed early in infancy, is itself based on the experience of contingency, where one event (usually) follows another. Thus, children's embodied and perceptual experiences with causality provide the conceptual template for the later construction of logical necessity.

Research within the past few decades has demonstrated that infants are able to perceive and respond to the contingency of physical events; that is, they can notice that if they cry, a parent is likely to appear, or, in an early experiment with 2-month-old infants, if they press their heads on a pillow, a mobile above their crib will turn (Watson, cited in Mandler 2004). As Mandler states: "Responsivity to the contingency of events is present at least from birth and is one of the most powerful factors governing perceptual learning and controlling attention" (2004: 96). Noticing and reacting to contingencies is a first step toward seeing two events as causally linked:

> An infant actively noticing that every time she drops something over the side of her high-chair her mother picks it up is conceptualizing a kind of "if-then" relation (suggesting that this is one of the image-schemas on which intuitive understanding of logic rests) (Mandler 2004: 98).

Further research has investigated infant understanding of physical causality, and evidence for awareness that one event can cause another is found even in very young infants (Gopnik and Schulz 2007; Leslie and Keeble 1987; Mandler 2004; Sperber et al. 1996). Furthermore, the concept of causality constructed by preschoolers does not seem to be based purely on statistical covariation between events, but on "a causal mechanism view of causality, in which causation is understood 'primarily in terms of generation transmission' of force and energy" (Shultz, cited in Gopnik and Shulz 2007: 9). That is, young children do not simply think that one event causes another because they occur together, but instead use early notions of transmitting force or energy between entities. This conceptualization of causation in terms of forces adds the characteristic of compulsion to the idea of contingency: not only does pushing the tower of blocks result in it falling over, but also the tower must fall because the push causes it to. As Mandler states, "It [the if-then notion] 'paints' force onto the object and kinetic information that the perceptual system provides, leading to the perception of causality" (2004: 99–100).

Thus, from a very early age, we are capable of building a schema for physical causality, such that if one event follows another in time, and physical contact or force exists between the entities involved, the first event is seen as causing the second. Another way of putting this is that if the first event happens, then the second event must also happen. This structure or schema for force has also been identified within cognitive linguistics as an embodied source for many abstract concepts, from personal relationships to political actions (Talmy 1988). This image schema for force includes the following properties (Evans and Green 2006: 187):

- Force schemas involve a force vector, i.e., a directionality.
- Force schemas have sources for the force and targets that are acted upon.
- Forces involve a chain of causality.

These properties map to analogous properties for logical deductions and proofs:

- Logical deductions have directionality ("If A then B" is not the same as "If B then A").
- Logical deductions have sources for the premises (previously proved propositions, postulates, and/or axioms) and the conclusions (the premise and its sources).
- Mathematical proofs involve a chain of logical deductions.

Thus, we propose that force/physical causality serves as a source domain for an unconscious conceptual metaphor underlying our understanding of logical deduction. In this metaphor, logical conclusion B "follows" premise A, with the conclusion having the same sense of necessity as physical causation: just as a physical effect "must" be a consequence of its cause, given valid reasoning, a logical conclusion "must" follow its premise. Table 6.2 spells out the metaphorical mapping from physical causation to logical deduction.

Cognitive Continuity in Conditionals

The logical deductions that comprise the building blocks of mathematical proof, typically expressed using if-then statements, belong to a linguistic category known as conditionals. Conditionals have been analyzed within a cognitive linguistics

Table 6.2 Metaphorical mapping between physical causation and logical deduction

LOGICAL DEDUCTION IS PHYSICAL CAUSATION	
Source domain: causation via physical forces: "This action causes that effect"	Target domain: logical deduction: "If A then B"
• Two entities • One is foregrounded or singled out (the "agonist" or effect) • The other is considered in terms of the effect it has on the agonist (the "antagonist" or cause) • Physical force • If the force of the antagonist is sufficiently strong, the result is motion of the agonist	• Two declarative statements • One is foregrounded as the "conclusion" • The other ("premise") is considered in terms of the implication that it has for the truth of the conclusion • Logical necessity • If the logical necessity connecting the premise to the conclusion is valid, then the truth of the conclusion is established

framework, seeing them as constructions within linked mental spaces (Dancygier and Sweetser 2005; Fauconnier 1994). Given the importance of if-then statements in mathematical reasoning, we will look at the relationship between conditionals used in everyday, nonmathematical settings, and those used in mathematics, to see whether this reveals another type of conceptual continuity.

There are several kinds of conditionals found in everyday discourse, typified by how the two clauses (premise and conclusion) are related (Dancygier and Sweetser 2005; Sweetser 1996). The most common type is the content (or predictive) conditional, in which the two clauses are semantically related, and in which the outcome or conclusion is contingent on the action in the premise taking place. An example would be the statement, "If you pet the cat, she will bite you." A conditional like this is not taken to mean that the cat will bite 100% of the time when petted, or that she might not bite even if she is not petted. But in general, when the cat is petted, she does bite.

The conditional of most interest in the current context is called an epistemic conditional, in which the speaker carries out a more formal logical reasoning process. Two examples are "If the car is in the driveway, he must be home" and "If x is even, then $x/2$ is an integer" (Dancygier and Sweetser 2005: 17). These kinds of conditionals go beyond contingency or a possible connection to what Dancygier and Sweetser call a "metaphoric 'compulsion' of the speaker's reasoning process" (p. 20) in which the speaker is "forced" to draw the given conclusion, based on either inductive reasoning ("the car is almost always in the driveway when he is home") or deductive logic (the mathematical definition of "even"). The conceptual roots of this metaphorical "compulsion" are spelled out in the previous section, where physical causality is seen as providing the source domain or conceptual template for logical deduction.

Logical deductions as utilized in mathematics can be seen as a type of epistemic conditional; however, in order to reduce ambiguity, the mathematical community has added constraints to such statements that are not found in everyday discourse. In everyday discourse, a content conditional is typically biconditional; that is, "If P then Q" also implies "If not P then not Q." Using the example from above, "If you pet the cat, she will bite you" would also imply that if you don't pet the cat, she won't bite you. This biconditionality (also called alternativity, Dancygier and Sweetser 2005) is explicitly rejected when using logical if-then statements in mathematics: "If P then Q" does not imply "If not P then not Q." This difference can cause difficulties for students when learning the more specialized discourse of mathematics, where students who first learn conditionals in everyday discourse may carry over the assumption of biconditionality into the more restricted setting of logical/mathematical deduction. When this happens, students are seen as making errors in logical thinking (Evans et al. 1993).

Within both everyday and mathematical discourse settings, the use of conditionals is often accompanied by physical gestures, and these gestures can serve as a source of data about the underlying conceptualization of conditional statements. If we find that the kinds of gestures utilized with conditionals within mathematics are similar to those used outside of mathematics, this would be additional evidence for cognitive continuity between everyday and mathematical thinking. By examining an existing corpus of participants in televised talk shows, Sweetser and Smith (2015) identified a characteristic gesture that often accompanies conditional statements uttered in nonmathematical discourse settings. In a study involving 402 video clips, the researchers found that conditionals used in this setting were generally accompanied by a particular hand motion: the speaker moved his or her hand along a transverse axis through space, starting on the speaker's left and moving toward the speaker's right, in parallel with the verbalization of an "if-then" statement. For example, this gesture was used by the author Michael Pollan on a talk show when saying, "If you're not hungry enough to eat an apple, you're not really hungry" (Sweetser and Smith 2015: 13).

We examine the use of epistemic conditionals, in the form of deductive statements, among the participants in the proof study to see whether their gestures reflect a similar conceptual structure as epistemic conditionals in everyday discourse. There were, of course, numerous examples of epistemic conditionals in the form of logical if-then statements throughout the interview sessions. These were particularly prominent during the section when the pairs of students were working together to find a proof for a conjecture. These epistemic conditionals included statements like the following:

1. "Like if you start above the line, then you stay above the line"
2. "If you compose *f* with itself a bunch, like, every time it's gonna keep going up"
3. "It's kind of stupid but you like, can't draw the picture if it, 'cause like, if you go right here, and like, you know …"

Although there are conditional statements in this corpus that did not begin with the word "if," in the current analysis, we looked only at those that did. Some statements were full conditionals, in that they included both premise and conclusion. Some of these conditional statements explicitly included the term "then" (Example 1), while others did not, but still expressed the premise and conclusion in full

(Example 2). In addition, as is typical in unrehearsed dialog, there were fragmentary or partial statements that utilized the word "if" but did not include a conclusion (Example 3, above).

Epistemic conditionals accompanied by gesture were most common when the participants were talking to the interviewer. During the proof-finding part of the session, when the students were working together and the interviewer was out of the room, the students were usually facing the blackboard, holding chalk, and writing or drawing. With their hands thus occupied, there were fewer gestures associated with conditionals than when explaining their work or answering the interviewer's questions.

Within the 12 interview sessions, there were a number of instances where the students, when uttering a conditional, demonstrated the same kind of transverse, left-to-right gestural movement found by Sweetser and Smith (2015). For example, in the first proof session, which involved two female graduate students, the students produced a total of 41 epistemic conditionals in their speech, including 11 that were incomplete. Of all the epistemic conditionals, four were accompanied by this typical transverse gesture (one sequence that included three instances of this gesture is analyzed below).

The similarities of these epistemic-accompanying gestures, which are found across mathematical and nonmathematical contexts, serve as evidence for a cognitive continuity, in that the same gesture that accompanies an epistemic conditional in everyday discourse is also found in the more specialized discourse involved in doing and discussing mathematical proof.

Table 6.3 illustrates an example of an epistemic conditional spoken by a student, accompanied by a series of gestures. Stills from the videos of the gestures are shown in the middle column of the table, labeled as Fig. 6.3a–i. In this example, doctoral student AB is discussing an episode in which she was helping her undergraduate students understand a new mathematical concept/procedure.

In this example, the epistemic conditional that the student is expressing can be summarized as: "If you have a scalar function and a vector function, then the rule for finding their product is the same as the rule for finding the product of two scalar functions." The sequence of gestures accompanying the student's speech is very rich, when characteristics such as hand shape and orientation, hand location, and movement of the hands through space are considered. We'll begin by looking at the direction of motion of the gestures.

Consistent with other epistemic conditionals, the sequence includes left-to-right motion along the transverse axis; in fact, this transverse motion occurs three different times, as shown in the pairs of figures above:

1. Figure 6.3a, b: A relatively small left-to-right transverse motion of the left hand, as AC begins by saying, "If you have *some scalar function* of T and *some vector function of T*." This sequence also includes a change in orientation of the left hand: when holding it on the left, AC uses an upward-opening (horizontal) C-shape as if "bracketing" or "holding" a scalar function. As she moves her hand to the right, she rotates her wrist so that when she says, "vector function," the

Table 6.3 Discourse segment by AC about scalar functions

AC: Well, I guess, so, the other day they were trying to prove that, um, if you have *some scalar function of T* Int: Uh huh AC: —and some vector function of T	 Figure 6.3a  Figure 6.3b	Left hand starts in horizontal C-shape (“bracket”) facing upward on left side of body Left-to-right motion with left hand along transverse axis, ending in middle of body, with C-shape turning vertical
Int: Uh huh AC: —that the *derivative of their product …*	 Figure 6.3c  Figure 6.3d	Left-to-right motion with left hand along transverse axis, with left hand open and facing outwards. Left hand begins on left side of body and ends in middle of body

(continued)

Table 6.3 (continued)

is the same …	 Figure 6.3e  Figure 6.3f	Rapid left-to-right motion with left hand along transverse axis. Left hand starts in loose horizontal C-shape (“bracket”) facing upward on left side of body and ends in pointing gesture to the right
AC: … product rule essentially that you know from just, you Int (talking over): Uh huh	 Figure 6.3g  Figure 6.3h	A complex motion in which the left hand begins by pointing downward, then is moved in a circle twice around the right hand while saying “you know,” ending up open and facing the speaker
AC: know from like *scalar functions*	 Figure 6.3i	Left hand moves to right and finishes in horizontal C-shape (“bracket”) on left side of body This is the same shape and location as when the phrase “scalar function” was initially uttered

Note: Italicized text indicates speech that is co-timed with gestures

C-shape is now vertical. She thus uses both hand shape and hand location to gesturally distinguish the two different kinds of functions.
2. Figure 6.3c, d: A wider left-to-right transverse motion of the left hand, as AC says, “the derivative of their product.” In this case, the hand shape stays the same throughout, open and facing outward.
3. Figure 6.3e, f: After saying “derivative of their product,” AC pauses briefly, and then makes a very rapid left-to-right motion of her left hand while saying, “is the same,” starting with a horizontal C-shape and ending with a right-facing point.

As can be seen above, in addition to an overall left-to-right transverse movement that occurs three times during the sequence, gestures are also used to mark or indicate specific mathematical objects, in a scheme that Calbris calls “two-entity opposition.” This happens when AC uses a horizontal “bracket” held to her left when saying “scalar functions” and then a vertical bracket held to her right when saying “vector functions.” As a second example, in Fig. 6.2c, the terms “derivative” and “product” have the same hand shape but are marked by left and right hand locations.

The discourse segment ends with AC discussing a “product rule” while using an iterative circular gesture during a pause in speech (possibly searching for her next words), and then verbally comparing it to the rule for scalar functions. Interestingly, the final gesture of the sequence, associated with the words “scalar function,” has an identical shape and location as the gesture used the first time these words were uttered. This is an example of using a specific hand shape and location in gesture space to “hold” a referent in discourse (Calbris 2008; McNeill 1992, 2005).

Overall, these gestures are consistent with prior research and theory related to gesture and if-then statements. Calbris (2008) has stated that in gesture space, the transverse axis can represent logico-temporal concepts, such as cause and effect, or before and after:

> A path in space or time is depicted by a left-to-right movement. But give that body symmetry allows this axis to account for splitting in two as well as two-entity oppositions, it can be used to oppose past and future, or precedence and successor, by locating the past on the left side and the future on the right side (Calbris 2008: 43).

The transverse axis of the body has been also called “the axis of reading and writing, pointing to the right in the Western world” (Calbris 2008: 28). In this case, the motion of AC’s gestures is consistent both with the placement of the “cause” (premise) on the left and the “effect” (conclusion) on the right, as well as the left-to-right order in which premise and conclusion are generally written in English.

Discussion and Conclusions

In this chapter, we have presented three examples, utilizing several kinds of evidence, that show how the processes and ideas associated with mathematical proof are embodied phenomena, rather than existing purely “in the head.” In addition, this evidence helps to demonstrate conceptual continuity between mathematical proof and nonmathematical thinking and discourse. Whether looking at how seasoned and emerging mathematicians talk about proof, at the gestures of doctoral students, or at research on infant cognition, we find that proof and its building blocks, statements of logical deduction, are not abstract elements of disembodied rationality. Instead, we argue, these sophisticated forms of discourse make use of basic image schemata related to force and motion, and are supported by conceptual metaphors grounded in physical experiences.

One such experience is the perception or action of physical causality: it is claimed that our earliest experiences as infants who are able to perceive/conceive that one event physically causes another provide a template for later being able to say and understand that "A implies B" in a logical sense. Without the notion of physical causality, we would not be able to build more abstract notions of social and logical causality. Similarly, the experience of physical motion through space, beginning at one location, proceeding along a trajectory or path to a final destination, whether perceived or enacted, gives rise to the source-path-goal schema. And this everyday schema in turn is recruited when we build the notion of a proof that starts with a premise, proceeds through a number of "steps," and finishes with a conclusion.

We acknowledge that mathematical proof is a specialized cultural product and a specific form of discourse, developed and formalized over centuries in order to become a powerful tool for both solving practical problems and exploring patterns and structures. However, we would claim that the form that this discourse takes is not arbitrary, but rather is grounded in particular kinds of embodied human experiences. This chapter has attempted to illustrate several ways in which proof-related discourse is grounded in the body. In order to assist students of mathematics to succeed in participating in its most characteristic practice, deductive proof, it is hoped that an understanding of how this grounding can both facilitate and, in some cases, hinder the learning of proof will be of value.

References

Abrahamson, D., and Lindgren, R. (2014). Embodiment and embodied design. *The Cambridge handbook of the learning sciences*, 2, 358–376.

Balacheff, N. (1991) Benefits and limits of social interaction: The case of teaching mathematical proof. In: A. Bishop, S. Mellin-Olsen and J. Van Dormolen J. (eds.) *Mathematical knowledge: Its growth through teaching*, pp. 175–192. Dordrecht, The Netherlands: Kluwer.

Bazzini, L. (1991). From grounding metaphors to technological devices: A call for legitimacy in school mathematics. *Educational Studies in Mathematics,* 47: 259–271.

Best, J. (2005). Recognition of proofs in conditional reasoning. *Thinking and Reasoning* 11: 326–348.

Calbris, G. (2008). From left to right…Coverbal gestures and their symbolic use of space. In: A. Cienki and C. Müller (eds.) *Metaphor and gesture,* pp. 27–54. Amsterdam, The Netherlands: John Benjamins Publishing.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics* 24(4): 359–387.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Dancygier, B., and Sweetser, E. (2005). *Mental spaces in grammar: Conditional constructions.* Cambridge, UK: Cambridge University Press.

Edwards, L. D. (2009). Gestures and conceptual integration in mathematical talk. *Educational Studies in Mathematics* 70(2): 127–141.

Edwards, L. D. (2010). Doctoral students, embodied discourse and proof. In: M. M. F. Pinto and T. F. Kawasaki (eds.), *Proceedings of the 34th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 2, pp. 329–336. Belo Horizonte, Brazil: PME.

Edwards, L. D. (2011). Embodied cognitive science and mathematics. In: B. Ubuz (ed.), *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 2, pp. 297–304. Ankara, Turkey: PME.

Edwards, L. D., Moore-Russo, D., and Ferrara, F. (eds.) (2014). *Emerging perspectives on gesture and embodiment in mathematics*. Charlotte, NC: Information Age Publishing.

Evans, V., and Green, M. (2006). *Cognitive linguistics: An introduction.* Mahwah, NJ: Lawrence Erlbaum.

Evans, J. S. B., Newstead, S. E., and Byrne, R. M. (1993). *Human reasoning: The psychology of deduction*. Hove, UK: Lawrence Erlbaum.

Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge: Cambridge University Press.

Fauconnier, G. and Lakoff, G. (2009). On metaphor and blending. *Cognitive Semiotics* 5, (1–2): 393–399.

Fauconnier, G. and Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities.* New York, NY: Basic Books.

Gibbs Jr., R. W. (2005). *Embodiment and cognitive science*. Cambridge, UK: Cambridge University Press.

Gopnik, A., and Schulz, L. (eds.) (2007). *Causal learning: Psychology, philosophy, and computation*. Oxford, UK: Oxford University Press.

Hall, R., and Nemirovsky, R. (2012). Introduction to the special issue: Modalities of body engagement in mathematical activity and learning. *Journal of the Learning Sciences* 21(2): 207–215.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange* 21(1): 6–13.

Hanna, G. (1995). Challenges to the importance of proof. *For the Learning of mathematics* 15(3): 42–49.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics* 44(1–2): 5–23.

Harel, G., and Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In: F. Lester (ed.), *Second handbook of research on mathematics teaching and learning*, pp. 805–842. Charlotte, NC: Information Age Publishing.

Johnson, M. (1987). *The body in the mind: The bodily basis of reason and imagination*. Chicago, IL: University of Chicago Press.

Johnson, M. (2012). *The meaning of the body: Aesthetics of human understanding***.** Chicago, IL: University of Chicago Press.

Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology* 50: 100–135.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kiefer, M., and Trumpp, N. M. (2012). Embodiment theory and education: The foundations of cognition in perception and action. *Trends in Neuroscience and Education* 1(1): 15–20.

Laborde, C. (2000). Dynamic geometry environments as a source of rich learning contexts for the complex activity of proving. *Educational Studies in Mathematics* 44(1–2): 151–161.

Lakoff, G., and Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lakoff, G., and Núñez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York, NY: Basic Books.

Leslie, A. M., and Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition* 25(3): 265–288.
Lin, F., Hsieh, F., Hanna, G., and de Villiers, M. (Eds.) (2009). *Proceedings of ICMI STUDY 19: Proof and proving in mathematics education.* Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University.
Mandler, J. M. (2004). *The foundations of mind: Origins of conceptual thought*. Oxford, UK: Oxford University Press.
Manin, Y. (1998). Truth, rigour, and common sense. In: H.G. Dales and G. Oliveri (eds.). *Truth in mathematics* (pp. 147–159). Oxford, UK: Oxford University Press.
Manin, I. I., Manin, J. I., and Manin, Y. I. (2007). *Mathematics as metaphor: Selected essays of Yuri I. Manin* (Vol. 20). Providence, RI: American Mathematical Society.
Marghetis, T., Edwards, L. D., and Núñez, R. (2014). More than mere handwaving. In: L. D. Edwards, D. Moore-Russo, and F. Ferrara (eds.), *Emerging perspectives on gesture and embodiment in mathematics*, pp. 227–246. Charlotte, NC: Information Age Publishing.
McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago, IL: Chicago University Press.
McNeill, D. (2005). *Gesture and thought.* Chicago, IL: Chicago University Press.
Nardi, E., and Knuth, E. (2017). Changing classroom culture, curricula, and instruction for proof and proving: How amenable to scaling up, practicable for curricular integration, and capable of producing long-lasting effects are current interventions? *Educational Studies in Mathematics* 96(2): 267–274.
New York Times (July 16, 2017). *Maryam Mirzakhani, Only Woman to Win a Fields Medal, Dies at 40.* Retrieved December 18, 2018 from: https://www.nytimes.com/2017/07/16/us/maryam-mirzakhani-dead.html
Núñez, R., Edwards, L. D., and Matos, J. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational Studies in Mathematics* 39(1–3): 45–65.
Oakley, T. (2007). Image schema. In: D. Geeraerts and H. Cuyckens (eds.), *Handbook of cognitive linguistics*, pp. 214–235. Oxford, UK: Oxford University Press.
Piaget, J. and Inhelder, B. (1964). *The early growth of logic in the child: Classification and seriation*. London, UK: Routledge.
Radford, L., Edwards, L., and Arzarello, F. (2009). Introduction: Beyond words (Special Issue: Gestures and Multimodality in the Construction of Mathematical Meaning). *Educational Studies in Mathematics* 70(2): 91–95.
Rasmussen, C., Stephan, M., and Allen, K. (2004). Classroom mathematical practices and gesturing. *The Journal of Mathematical Behavior* 23(3): 301–323.
Rotman, B. (1993). *Taking God out of mathematics and putting the body back in: An essay in corporeal semiotics*. Palo Alto, CA: Stanford University Press.
Roy, S., Inglis, M., and Alcock, L. (2017). Multimedia resources designed to support learning from written proofs: An eye-movement study. *Educational Studies in Mathematics* 96(2): 249–266.
Shapiro, L. (2010). *Embodied cognition*. New York, NY: Routledge.
Shapiro, L. (ed.). (2014). *The Routledge handbook of embodied cognition*. New York, NY: Routledge.
Sperber, D., Premack, A., and Premack, J. (eds.) (1996). *Causal cognition: A multidisciplinary debate.* Oxford, UK: Oxford University Press.
Sweetser, E. (1996). Mental spaces and the grammar of conditional constructions. In: G. Fauconnier and E. Sweetser, (eds.) *Spaces, worlds, and grammar*, pp. 318–333. Chicago, IL: University of Chicago Press.
Sweetser, E. (1998). Regular metaphoricity in gesture: Bodily-based models of speech interaction. In: *Actes du 16e Congrès International des Linguistes*. Amsterdam, The Netherlands: Elsevier (CD-ROM).
Sweetser, E., and Smith, I. (2015, July). *Conditionals, mental spaces and gesture.* Paper presented at the 13th International Cognitive Linguistic Conference, Newcastle-on-Tyne, England.

Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal* 20(2): 5–24.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science* 12: 49–100.

Thurston, W. P. (1994) On proof and progress in mathematics. *Bulletin of the American Mathematical Society 30*(2):161–177

Varela, F. J. (1999). *Ethical know-how: Action, wisdom, and cognition*. Palo Alto, CA: Stanford University Press.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Chapter 7
Math Puzzles as Learning Devices

Marcel Danesi

Introduction

As Polya (1957) cogently argued, the use of puzzles and games has always been part of doing and learning mathematics since its emergence as an autonomous discipline (see also Parker 1955; Gardner 1998). The reason for this long-standing pedagogical practice may be that puzzles stimulate the imagination more so than any other type of mental faculty and are thus likely to be highly effective devices at various stages of the learning process. The purpose of this chapter is to consider the cognitive reasons supporting this implicit pedagogical principle. The use of puzzles and games in math education can be called, for the sake of convenience, educational recreational mathematics (ERM).

A specific word for what we now call a *puzzle* did not exist in antiquity. Activities in mathematics that we would now label in this way were called *propositions*. From the beginnings of mathematical history, these have constituted not only teaching devices, but also explorations of ideas by mathematicians themselves, often leading to discoveries within the field (Danesi 2018). For instance, the *Ahmes Papyrus* (1650 BCE) contains puzzles that were likely intended for the schooling of Egyptian youth but which also exemplified emerging ideas within mathematics itself. Similar texts are found throughout the ancient world. In a phrase, the early mathematicians knew that puzzles were intrinsic to mathematics itself and to its learning in school.

Although puzzles are used as ancillary pedagogical devices in math classrooms, it is rare to find entire courses and textbooks, akin to the *Ahmes Papyrus*, that revolve

M. Danesi (✉)
University of Toronto, Toronto, ON, Canada

Cognitive Science Network, Fields Institute for Research in Mathematical Sciences,
Toronto, ON, Canada
e-mail: marcel.danesi@utoronto.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_7

around them. Given the abundance of classic puzzles that have constituted the foundation of various branches of mathematics—for example, the Königsberg Bridges Puzzle, devised by Leonhard Euler, laid the groundwork for graph theory and the Rabbit Puzzle, invented by Leonardo Fibonacci, became the basis over time for the theory of sequences—the purpose of this chapter is to argue for the integrated use of such puzzles and their variants as part of ERM. The underlying objective is to give students the opportunity to "relive" historically how a particular field of mathematics emerged through the puzzle that gave it birth. Math teachers are adept at producing materials, manuals, and textbooks that incorporate puzzles as part of pedagogy (see, for example, Caldwell-Landsittell 2005). But there are few curricula that are based on puzzle-solving itself, as was the *Ahmes Papyrus* or the late-ninth-century puzzle collection compiled by Alcuin, *Propositiones ad acuendos juvenes* (see Hadley and Singmaster 1992), both of which understood that puzzles formed the core of mathematical learning.

Puzzles, Problems, and Games

As mentioned, an equivalent word for *puzzle* did not exist in any of the languages of the ancient world, although the concept that it encodes was implicit in early math textbooks such as the *Ahmes Papyrus*. So, labeling some ancient math activity as a "puzzle" today is a retrospective form of reference. The term *proposition* was used within geometry to refer to a problem in which some shape or figure had to be constructed in a certain way according to specific principles. The word *problem* was coined a little later with this sense, namely as a question or task that pertains to some geometrical situation. From this, the word was extended to cover any mathematical question that required a specific kind of answer. The word *proposition* however remained as an alternative word for a challenging puzzle, as can be seen in the title of Alcuin's famous collection—*Propositiones ad acuendos juvenes*.

A distinction between *problem, puzzle*, and *game* is of primary importance for the purposes of ERM. Generally speaking, a problem presents information that can be used unambiguously to reach a solution directly; a puzzle, on the other hand, presents information that appears to be incomplete or else conceals a twist or a clever trap, thus making it much more difficult to reach a solution. Both problems and puzzles are Q&A (question and answer) structures. The difference between the two can thus be shown graphically as follows:

Problem
Q → A (the question leads directly to an answer)
Puzzle
Q → (A) (the answer to the question is not immediately obvious)

A simple problem might be the following: "Given sides of length 15 and 23 in a right-angled triangle what is the length of the hypotenuse?" This has a straightforward solution because all the information that is needed to solve it is given to us, if we have learned the Pythagorean theorem. Now, in contrast, consider the following puzzle from the pen of the Renaissance Venetian mathematician, Niccolò Tartaglia:

> A man dies, leaving 17 camels to be divided among his heirs, in the proportions 1/2, 1/3, 1/9. How can this be done?

Dividing up the camels in the manner decreed by the father would entail having to split up a camel. This would, of course, kill it. So, the wily Tartaglia suggested "borrowing an extra camel," for the sake of mathematical argument, not to mention for humane purposes. With 18 camels, we arrive at a solution: one heir was given 1/2 (of 18), or 9; another 1/3 (of 18), or 6; and the last one 1/9 (of 18), or 2. The 9 + 6 + 2 camels given out in this way add up to the original 17. The extra camel could then be returned to its owner. Clearly, Tartaglia devised his puzzle as a ludic play on fractions, not as a realistic solution to a practical problem. As Petkovic (2009: 24) observes, this simple puzzle also offers up generalization possibilities—the crux of mathematical method. Tartaglia himself did so by finding solutions to the n-camel version of the puzzle. If there are three brothers, a, b, and c, and the proportions are $1/a{:}1/b{:}1/c$, then solutions are produced by the following Diophantine equation:

$$n/(n+1) = 1/a + 1/b + 1/c$$

The solutions are shown below:

$$\begin{aligned}
&n = 7(a = 2, b = 4, c = 8)\\
&n = 11(a = 2, b = 4, c = 6)\\
&n = 11(a = 2, b = 3, c = 12)\\
&n = 17(a = 2, b = 3, c = 9)\\
&n = 19(a = 2, b = 4, c = 5)\\
&n = 23(a = 2, b = 3, c = 8)\\
&n = 41(a = 2, b = 3, c = 7)
\end{aligned}$$

The key feature of Tartaglia's puzzle is that, unlike the previous problem, its solution is not obvious. It is common to refer to the effect that puzzle solutions produce in us as the "Aha" effect, reflecting the unexpectedness of the answers. It is relevant to note that Aha thinking has been found to originate in the right hemisphere of the brain—a fact that is especially critical in developing a cognitive theory of puzzles, as will be discussed (Bowden et al. 2005).

The concept of *mathematical game* is also of relevance to ERM. A game presents an initial state (I), asking us to reach an end state (E) via a set of rules. An example is Sudoku, which presents a grid with some numbers in it (I) and a set of rules (how to place numbers in the grid) in order to achieve an end state (E)—the completion of the grid. A mathematical game has the following structure:

Math Game
I → R → E (the initial state must be modified via rules to reach the end state)

A classic chess-derived math game is the so-called knight's tour, which asks the following (Conrad et al. 1994):

> Place a knight on the chessboard so that it visits every square once and only once.

There are many solutions to the puzzle, with the earliest one dating back to the ninth century in the *Kavyalankara,* a Sanskrit work on the nature of poetry. These need not concern us here. Suffice it to say that games such as the knight's tour require an abundant use of imagination to solve, even if we already know what is expected of us in the end, unlike Tartaglia's puzzle where the answer was not initially obvious. So, since we know that there is a solution (E), the game produces a "Eureka" effect, rather than an Aha one. This results from working out a way (or ways) to reach the end state, which might seem intractable at first. It is an expression of satisfaction more than one of surprise.

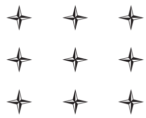
Needless to say, the line between a puzzle and a game is a blurry one, and this is perhaps why the two terms are used interchangeably, despite the fact that they refer to different psychological processes. Both are to be incorporated systematically within ERM, as will be discussed. As Trigg (1978: 18) has aptly observed, the term "recreational" should be taken at face value—creating math in an imaginative way. The objective of ERM is to allow students to explore mathematics through puzzles and games. Determining which ones are to be included under this rubric is a subjective act. As Trigg (1978: 21) remarks: "Recreational tastes are highly individualized, so no classification of particular mathematical topics as recreational or not is likely to gain universal acceptance." The only pedagogical principle involved is to make sure that the choice of puzzles and games is synchronized to the overall flow of the curriculum. This topic will be discussed more concretely below.

The Aha, Gotcha, and Eureka Effects

Aha thinking is *the* defining cognitive characteristic of puzzles. The classic example used to demonstrate this by psychologists is the so-called nine-dot puzzle (Fig. 7.1):

> Without your pencil leaving the paper, can you draw four straight lines through the following nine dots?

Fig. 7.1 Nine-dot puzzle



Many students attack this puzzle by joining up the dots as if they were located on the perimeter (boundary) of a square or flattened box. But this reading of the puzzle does not yield a solution, no matter how many times they try to draw four straight lines without lifting the pencil. A dot is always "left over." At this point, Aha thinking comes into play: “What would happen if the four lines were to be extended beyond the assumed imaginary box structure of the dots?” That hunch turns out to be the relevant Aha insight. One solution is as follows (Fig. 7.2):

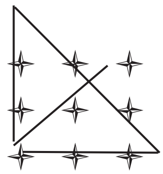


Fig. 7.2 A solution to the nine-dot puzzle

The first appearance of a dot-joining puzzle is in Sam Loyd’s *Cyclopedia of 5000 puzzles, tricks, and conundrums with answers* (1914). The puzzle is now used in psychology to study imaginative thinking (Kershaw and Ohlsson 2004). It requires solvers to literally look beyond the implied box figure of the puzzle. The term “lateral thinking” was proposed by De Bono (1970) to refer to the Aha thinking involved in solving the nine-dot puzzle, because a solver must think beyond the boxlike structure that the puzzle initially suggests. In his intriguing book *Aha! Insight!* Gardner (1979) exemplifies how this type of thinking unfolds through actual puzzles, allowing us to engage in such imaginative thinking directly.

The key pedagogical aspect of this kind of thinking is that it can be generalized after the initial Aha solution. The above puzzle is a 3 × 3 version of a dot-joining puzzle. By solving 16-dot, 25-dot, and various other puzzles, is it possible to uncover some general principles hidden within it? Is there a correlation between number of dots and number of connecting lines? This line of “post-solution thinking” is the essence of recreational mathematics. After solving a number of more complex dot-joining versions, a chart emerges, which suggests a relation between the number of dots and the lines required to solve it:

Dots	*Lines required*
3×3	$(3 + 1) = 4$
4×4	$(4 + 2) = 6$
5×5	$(5 + 3) = 8$
6×6	$(6 + 4) = 10$
…	…
$n \times n$	$n + (n - 2) = 2n - 2 = 2(n - 1)$

Needless to say, research on this type of puzzle has revealed more complexity. The point here is that it allows students to explore a pattern that may be packed into it. Unpacking that pattern is the role of logical reasoning; discovering the pattern in the first place is the role of Aha thinking.

Consider now the following classic puzzle that turned up for the first time in an arithmetic textbook written by Christoff Rudolf and published in Nuremberg in 1561 (De Grazia 1981):

> A snail is at the bottom of a 30-foot well. Each day it crawls up 3 feet and slips back 2 feet. At that rate, when will the snail be able to reach the top of the well?

This puzzle requires linking the counting process with the physical scenario to which it refers in an ingenious way. Since the snail crawls up 3 ft., but slips back 2 ft., its net distance gain at the end of every day is, of course, 1 foot up from the day before. To put it another way, the snail's climbing rate is 1 foot up per day. At the end of the first day, therefore, the snail will have gone up 1 foot from the bottom of the well, and will have 29 ft. left to go to the top (remembering that the well is 30 ft. in depth). If we conclude that the snail will get to the top of the well on the 29th day, as many students do (in my own teaching experience), we will have fallen into the puzzle's hidden trap. On the *second* day it starts at 1 from the bottom; on the *third* it starts at 2 from the bottom; so, on *28th* day it starts at 27 from the bottom. This means that the snail has 3 ft. to go to the top on that day. It goes up the 3 ft., reaches the top, and goes out, precluding its slippage back down. For the sake of historical accuracy, it should be mentioned that the original puzzle archetype is found in the third section of Fibonacci's *Liber Abaci* (1202):

> A lion trapped in a pit 50 feet deep tries to climb out of it. Each day he climbs up 1/7 of a foot: but each night slips back 1/9 of a foot. How many days will it take the lion to reach the top of the pit?

Students react to this puzzle with a sense of having being duped. So, rather than the Aha effect, it thus produces a "Gotcha" effect, as Gardner (1982b) aptly designated it. When we fall into the puzzle's trap, we really do not like it. However, pedagogically, the Gotcha effect is still very important—it warns us to read information carefully and extract from it the required interpretation. Moreover, in this case, the puzzle can be used to illustrate the meaning and value of the "number line," since the well can be envisioned as such a line and movements up and down that line as points on it.

As mentioned, rather than an Aha effect, math games produce a Eureka effect. "Eureka" means "I have found (it)" in Greek, and is famously connected to what Archimedes supposedly shouted when he envisioned a way to determine the purity of gold by applying the principle of specific gravity. Producing this effect is as important as producing the Aha and Gotcha effects—once a student reaches the point of exclaiming Eureka, then one can safely say that the student has grasped the hidden principle in a game and thus is well on the way to grasping the concept involved.

In sum, the overarching goal of ERM is to produce all three effects—Aha, Gotcha, and Eureka—through puzzles and games. The premise is that these stimulate the imagination and thus get learners to engage in the same kind of creative thinking that mathematicians have engaged in since the start of their discipline.

Objectives of ERM

The main pedagogical objective of ERM, as discussed above in the case of the nine-dot puzzle, is to get students to model solutions, searching for any general principle embedded in them. Consider the *eight queens puzzle*, which is really a game as defined here, whereby eight queens must be placed on an 8-by-8 chessboard in such a way that none of the queens is able to capture any other queen (with the normal rules of chess). A solution (or end state) requires that no two queens share the same row, column, or diagonal. One solution is given in Fig. 7.3.

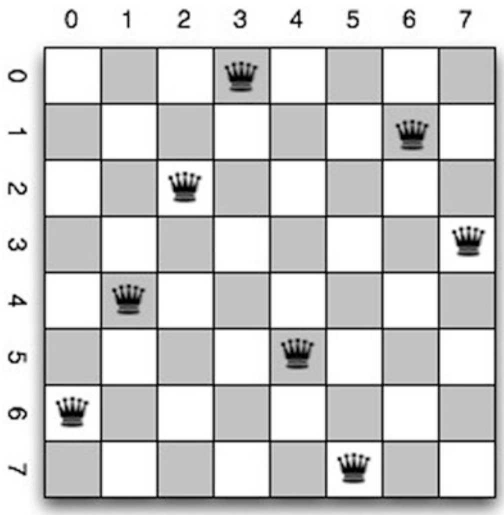


Fig. 7.3 A solution to the eight queens puzzle game

Now, students will invariably come up with different solutions. There are actually 92 distinct solutions; however if rotations and reflections of the board are taken into account, then it has 12 unique solutions. The puzzle game can now be used to get students to envision how to model an "*n*-by-*n*" solution. This need not concern us here. The point is that a simple game can lead to an engagement with fundamental mathematical method in the classroom and this is bound to enhance learning.

The puzzles and games that can be incorporated into ERM need to be, of course, tagged for their learning value. So, for example, cryptarithmetic puzzles can be identified as useful for the teaching of arithmetic. The topic of combinatorics can be presented and reinforced with puzzles and games involving sequential movement, such as the Towers of Hanoi Puzzle. The Monty Hall and Birthday Problems can be used profitably to impart basic notions of probability theory. Euler's Königsberg Bridges Puzzle is effective for introducing modern graph theory. Many tiling, packing, arrangement, and dissection puzzles, such as tangrams and Soma Cubes, lend themselves to the teaching of notions such as symmetry and impossibility. The list is actually endless and, as Trigg (1978) pointed out, the selection is really a subjective matter. The main thing to keep in mind is the didactic value of specific puzzles in a particular learning situation. So, the starting point for integrating puzzles and games into math education is developing an appropriate typology. Needless to say, there have been many proposals for classifying math puzzles. For example, the table of contents in Schuh's (1968) classic collection contains 267 distinct puzzle types. And his typology comes before the many new placement and mechanical puzzles that now flood the market. So, it is more practicable to concentrate on how specific puzzles and games can become entry points into mathematical ideas, rather than simply developing a generic taxonomy.

Virtually any classic puzzle or game can be incorporated into ERM. Consider the game of tic-tac-toe. As trivial as it might seem, it actually raises key questions related to probability and symmetry: What is the likelihood of winning if the X or the O is inserted in a particular location? What placement makes sense at the start? Can we reconstruct the winning moves and explain them? Answering these questions entails inferential analysis and hypothesis thinking. As Moscovich (2015: 15) has perceptively remarked: "Despite its apparent simplicity, tic-tac-toe requires detailed analysis to determine even a few elementary combinatorial facts, like the number of possible positions." Interestingly, all games should produce a draw; so, error and miscalculations, along with one opponent outwitting the other, are the only ways for winning to occur. Mathematics does not fail, but humans do, which is itself a significant pedagogical lesson to be imparted.

It is accurate to say that math educators, by and large, employ puzzles to illustrate or reinforce math notions or principles, complementing other materials. This constitutes an *ancillary* form of ERM. As far as can be told, an *integrative* use of puzzles and games in contemporary math education has been only occasionally contemplated. This topic will be broached below. At this point it is sufficient to outline the main pedagogical objectives of ERM in a general way:

1. Puzzles and games stimulate imaginative thinking. They are, in a sense, math thought experiments.
2. They challenge and motivate students to explore mathematical notions or principles creatively.
3. They allow students to explore how ideas in mathematics may have arisen.
4. Some math games, such as the eight queens puzzle, help develop spatial reasoning skills naturally.
5. Placement games, such as Sudoku, stimulate reasoning processes combined with spatial reasoning (what to place and where to place it).
6. Number puzzles, probability puzzles, graph puzzles, and many more span the whole domain of mathematical analysis in creative and ingenious ways.

Alexander (2012) has identified three dimensions of math cognition that are relevant to the present discussion—"pre-math," "math," and "mathematics." "Pre-math" is innate and intuitive, including a primitive sense of number and space. "Math" is what we learn as a set of formal skills, from elementary school to more advanced levels of education. It is what educators, policy makers, mathematicians, and many businesses want everyone to be competent in. "Mathematics" is the discipline itself, with its own professional culture, its research agendas and epistemologies, its own sense of correctness built around rigorous proofs, and so on. The boundaries among the dimensions are not clear-cut, and there are many cross-influences, but the distinctions are useful nonetheless. The goal of ERM is to transform "pre-math" into "math" and then to get the student to use the new knowledge to discover "mathematics" as a discipline.

Psychological and Pedagogical Aspects

To use an analogical construct from foreign-language learning theory, a distinction can be made between the *acquisition* of mathematical ideas and the *learning* of these ideas in a more formal fashion. This distinction was articulated first by linguist Krashen (1982, 1985). It is worthwhile reviewing it schematically here and considering its implications for ERM.

Noting how children develop their native languages naturally and how students struggle instead to master a foreign language at school, Krashen characterized the difference as one between *acquisition* and *learning*. The former characterizes spontaneous development in context, including early classroom input. Learning, on the other hand, is a conscious mode of analyzing the input that is activated when students know enough about the language to be able to reflect upon it formally. Krashen's distinction encapsulates something that math teachers have also felt intuitively—namely that students pick up certain things spontaneously but require conscious effort and focus to grasp other things. Acquisition is dominant during the processing of early input, when students pick up many new skills unconsciously. Learning, on the other hand, is dominant during later stages, when students attempt to understand "what is going on," so to speak.

Krashen derived his ideas from the work of the Russian psychologist Vygotsky (1961). For instance, his "$i + 1$" characterization of acquisition is derived from Vygotsky's notion of "zones of proximal development." This implies that children progress through zones of learning that are extended spontaneously as soon as they are able to understand new input by themselves—hence "i (input) + 1." Krashen claimed (1985: 1) that this "does not appear to be determined solely by formal simplicity," nor is it dependent on "the order in which rules are taught in language classes." It is triggered by the acquisition mode. The primary implication is that to get acquisition to unfold naturally involves providing input that contains "a bit" of information that is beyond the student's developing competence.

ERM aims to provide an "$i + 1$" learning environment throughout the course of the study. As Krashen argued, this type of environment will activate the modalities of the right hemisphere of the brain first, leading subsequently to the left hemisphere's ability to encode and formalize the ideas grasped naturally. In previous work, I have referred to this flow of learning as "bimodal" (Danesi 2003). The essence of bimodality can be seen in Gardner's (1982a: 74) statement: "Only when the brain's two hemispheres are working together can we appreciate the moral of a story, the meaning of a metaphor, words describing emotion, and the punch lines of jokes."

Bimodality theory is based on the observation that for the brain to grasp unfamiliar input it requires the experiential (probing) right-hemisphere functions to operate freely. These can be called R-Mode functions. When familiarity with the input becomes viable, then the analytical capacities of the left hemisphere, called L-Mode, come into play. Bimodality thus suggests a general principle of learning that can be called the modal flow principle:

> New notions and structures are learned more efficiently when the brain is allowed to process them in terms of an R-Mode (experiential) to L-Mode (analytical) flow.

So, during the initial acquisition stages, students need to assimilate input through creative R-Mode activities. This is where puzzles and games fit in—as R-Mode devices. But after this stage, students need to be exposed to formal explanations, practice drills, and other kinds of L-Mode techniques to reinforce what they have learned. The modal flow principle thus claims that: (1) experiential-creative forms of teaching belong to the initial acquisition stages; (2) teaching should become progressively more formal after these stages; and (3) the creative utilization of the new input belongs to the final reinforcement stages. Stage (1) is, as mentioned, an R-Mode stage, (2) an L-Mode stage, and (3) can be called simply a bimodal stage.

Needless to say, an advanced math student who is already in firm control of the required L-Mode skills through previous training will not have to spend as much time on the R-Mode phase as would a beginner. When students have mastered the L-Mode aspects of a concept, then they will be in a position to integrate them with the R-Mode ones as they are exposed to new mathematics. A consummate control of mathematics is, from a neuropsychological perspective, a bimodal feat, requiring the integrated contribution of both the R-Mode and the L-Mode to the understanding of a task.

The modal flow principle makes one fundamental demand on teachers—it requires them to identify a specific learning task as being novel or not. Generally speaking, something is novel when it involves structures or concepts that are either different or absent from the point at which the curriculum finds itself. Suffice it to say at that any task or input can be considered to be novel if the students demonstrate an inability to understand it or use it functionally. If that is so, instructional techniques that focus on analysis will be of little value, since the students have no preexisting L-Mode schemas for accommodating the new information directly. In order to make something accessible to the L-Mode, pedagogical experience dictates that the learner should be allowed to explore the new structures and concepts through R-Mode techniques. So, if graphs are to be discussed, an ideal R-Mode device to introduce them would be Alcuin's River Crossing Puzzles or Euler's Königsberg Bridges Puzzle. Once the initial R-Mode acquisition stage has been allowed to "run its course," the teacher can then "shift modes" and put the students in a frame of mind that allows them to reflect on the new structural patterns *in them-*

selves—namely the principles of critical path or graph theory. This implies the use of L-Mode techniques such as explanations, exercises, and the like. There is no way to predict when students will reach this stage, as Vygotsky maintained; but when they do it is obvious to teacher and learners alike.

From the foregoing discussion, the central question that applies to ERM is to determine whether puzzles should be used to maximum effect as ancillary or integral devices in a specific math classroom. The modal flow principle answers this question indirectly by suggesting that puzzles should be used during R-Mode stages, when they are integral to the acquisition process. During L-Mode stages the math concepts which they harbor should be taught formally.

To allow the R-Mode to process the novel input, students can be asked to provide variations to the puzzle given by extemporizing and adding their own versions to its basic format, no matter how many structural errors they might commit in the process. The actual time it takes for this stage to run its course will depend on the learners themselves (that is, on the kind of math know-how they bring to the classroom, on their previous familiarity with the concept in question, and so on). However, in order for the new structures and concepts to become part of long-term memory, this R-Mode stage must be followed up by techniques that allow the students to reflect and analyze the new content. This secondary stage in the flow cannot be circumvented. L-Mode knowledge will not necessarily emerge on its own in a *de facto* manner. So, ERM is really part of an integrated pedagogical system, amalgamating traditional mathematics education with puzzles in a cohesive way. Obviously, some types of learners will need fewer formal explanations and exercise reinforcement practice than others. Moreover, the type of instruction to be used will vary according to both learner cognitive style and type of input—some tasks are best taught through puzzle devices, and others through more formal ones. The way to determine what modal technique is appropriate is simply to try one out. Clearly, if some puzzle doesn't seem to work, then another one (or a different one within the same puzzle genre) must be considered. Similarly, the degree of utilization of the students' background knowledge will depend on the level of math competence reached by them—the more they are familiar with the concept the more the new material can be explained without ERM strategies. Commentaries on any pattern or feature related to, or derived from, the new content should be elaborated when required.

At various points in the learning process, individual students may manifest some persistent difficulty in utilizing a concept. At other points, they may show difficulty in using a certain concept in applications or in a mathematically appropriate fashion. At such points, it is obvious that they will need to *focus* on the concept itself. This aspect can be called the modal focusing principle:

> It is necessary from time to time in a course of study for the student to focus on some modal feature or pattern that is causing learning difficulty.

When a learner needs help in overcoming some error pattern that has become an obstacle to learning, then "L-Mode focusing" techniques should be used to allow that student an opportunity to relearn the feature in question. "R-Mode focusing" techniques, on the other hand, may be needed when a student shows the inability to apply a certain concept or structure; this is when puzzle techniques become highly useful. In essence, this principle claims that there will be points when students may need to stop and focus on certain aspects of the material introduced.

In an important work, Meyer et al. (2014) argue cogently for developing an overall approach to math education that they call puzzle-based learning (PBL). For these math educators, puzzles should be selected or designed to motivate students to think about framing and solving unstructured problems. The main goal, which is in line with bimodality theory, is to turn the implicit math concept inherent in certain puzzles into theoretical knowledge. The PBL approach is, in a phrase, a concrete proposal of how to teach mathematics bimodally. There is potentially an infinitude of others, allowing teachers to decide for themselves how to organize classroom input effectively.

Concluding Remarks

The historical imagination of mathematicians is imprinted in the classic puzzles and games that they have devised. Immersing oneself in this imagination implies grasping their puzzles. ERM is an attempt to project students into the imaginations of the greatest mathematicians of all time through the puzzles they created. To be used practically, it only requires an adjustment in the organization and timing of input within the normal course of pedagogy, at any level of education. How this can be done has been discussed schematically in this chapter.

Math educator Picciotto (2012) characterizes the use of puzzles in the classroom as a form of energy:

> Bottom line: the kind of energy the right puzzle brings into the classroom is phenomenal, with all types of kids. To get teachers to buy into this, the challenge is making the connection between puzzles and the core curriculum, rather than promote puzzles as something that just happens on the side.

It is that “energy” that ERM aims to harness. Without such energy, incidentally, it is unlikely that puzzles would have played such a key role in the history of mathematics itself. Puzzles are, as mentioned, explorations in the patterns that are inherent in some situation. As the Estonian biologist von Uexküll (1909) claimed, the brain is programmed to “model” such patterns in the form of cognitive artifacts. Puzzles are such artifacts, revealing that the brain is tuned into the structure of the world (since it rises from it). Through the puzzles that it generates, we are given glimpses of that structure a little at a time.

References

Alexander, J. (2012). On the cognitive and semiotic structure of mathematics. In: M. Bockarova, M. Danesi, and R. Núñez (eds.), *Semiotic and cognitive science essays on the nature of mathematics*, pp. 1-34. Munich: Lincom Europa.

Bowden, E. M., Jung-Beeman, M., Fleck, J. and Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences* 9: 322-328.

Caldwell-Landsittell, T. (2005). *Puzzled by math! Using puzzles to teach math skills.* Waco: Prufrock Press.

Conrad, A., Hindrichs, T., Morsy, H., and Wegener, I. (1994). Solution of the knight's Hamiltonian path problem on chessboards. *Discrete Applied Mathematics* 50: 125-134.

Danesi, M. (2003). *Second language teaching: A view from the right side of the brain.* Dordrecht: Kluwer.

Danesi, M. (2018). *Ahmes' legacy: Puzzles and the mathematical mind*. New York: Springer.

De Bono, E. 1970. *Lateral thinking: Creativity step-by-step.* New York: Harper & Row.

De Grazia, J. (1981). *Math tricks, brain twisters & puzzles.* New York: Bell Publishing Company.

Fibonacci, L. (1202/2002). *Liber abaci*, trans. by L. E. Sigler. New York: Springer.

Gardner, M. (1979). *Aha! Insight!* New York: Scientific American.

Gardner, H. (1982a). *Art, mind, and brain: A cognitive approach to creativity*. New York: Basic.

Gardner, M. (1982b). *Gotcha! Paradoxes to puzzle and delight*. San Francisco: Freeman.

Gardner, M. (1998). A quarter century of recreational mathematics. *Scientific American* 279: 68-75.

Hadley, J. and Singmaster, D. (1992). Problems to sharpen the young. *Mathematics Gazette* 76: 102-126.

Kershaw, T. and Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30: 3-13.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Krashen, S. D. (1985). *The input hypothesis*. London: Longman.

Meyer, E. F., Falkner, N., Sooriamurthi, R., and Michalewicz, Z. (2014). *Guide to teaching puzzle-based learning.* New York: Springer.

Moscovich, I. (2015). *The puzzle universe: A history of mathematics in 315 puzzles.* Buffalo: Firefly Books.

Parker, J. (1955). The use of puzzles in teaching mathematics. *The Mathematics Teacher* 48: 218-227.

Petkovic, M. S. 2009. *Famous puzzles of great mathematicians*. Providence, RI: American Mathematical Society.

Picciotto, H. (2012). Puzzles in math education. http://blog.mathedpage.org/2012/05/puzzles-in-math-education.html.

Polya, G. 1957. *How to solve it*. New York: Doubleday.

Schuh, F. (1968). *The master book of mathematical recreations*. New York: Dover.

Trigg, C. W. (1978). What Is Recreational Mathematics? *Mathematics Magazine* 51: 18-21.

von Uexküll J. (1909). *Umwelt und Innenwelt der Tierre*. Berlin: Springer.

Vygotsky, L. S. (1961). *Thought and language*. Cambridge, Mass.: MIT Press.

Chapter 8
Diagrams in Mathematics: On Visual Experience in Peirce

Vitaly Kiryushchenko

I think there could be a mathematical explanation of how bad your tie is (Russell Crowe as John Nash in Ron Howard's "The Beautiful Mind")

Just as we say that a body is in motion, and not that motion is in a body, we ought to say that we are in thought and not that thoughts are in us (Charles S. Peirce)

Introduction

Mathematicians use diagrams in their work all the time, whether they want to make use of Euclid's fifth postulate, to prove Fermat's principle, or to extract an algorithm that defines the seemingly chaotic movement of pigeons picking bread crumps from the ground. Using diagrams helps mathematicians identify patterns that solve particular mathematical problems by making the force of necessary reasoning *visually given*. A mathematical diagram, a paradigmatic use of which is exemplified in Euclid's *Elements*, is an *individual* image that instantiates *necessary* relations. As an observable entity, it allows a mathematician to experiment upon it and to visually demonstrate the necessity of a given conclusion. At the same time, it represents an abstract mathematical object. We do not use diagrams simply to facilitate our reasoning and then translate those diagrams into a formal calculus in order to make inferences. Diagrams *themselves* are immediate visualizations of the deductive process as such. The necessary character of deductive arguments is thus internal to the diagrams mathematicians construct (Sloman 2002).

There has been a surge in research on the use of diagrams in mathematics in the last two decades, with subjects ranging from the general role diagrams play in mathematical proofs (Barker-Plummer 1997; Kulpa 2009; Mumma 2010; Sherry 2009)

V. Kiryushchenko (✉)
York University, Toronto, ON, Canada
e-mail: kiryushc@yorku.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_8

to the role of analogical reasoning in scientific concept formation (Abrahamsen and Bechtel 2015; Nersessian 1992) and the use of diagrammatic tools in teaching mathematics (Bakker and Hoffmann 2005; Danesi 2016: 92–108; Hegarty and Kozhevnikov 1999; Legg 2017; Prusak 2012). The assumption shared by the majority of researchers is that studying diagrams may be seen as a way of bridging the gap between the purely platonic view of mathematics as a domain of abstract eternal forms, on the one hand, and theories that aspire to uncover the experiential basis of mathematical truths, on the other hand (Danesi 2016a: 15–18). This assumption is epitomized in the claim that the cognitive structure of mathematics presupposes a strong connection between mathematical abstractions and *metaphorical* cognition, and that our ideas of quantity and number systems are linked to our bodily experiences through what is known as "conceptual blending" (Lakoff 1999; Lakoff and Núñez 2000). To unpack the metaphors that instantiate such blending is to reveal the cognitive schematism which is deeply ingrained in human conceptual capacities and which shows that both surface figurative language and mathematics "are implanted in a form of cognition that involves associative connection between experience and abstraction" (Danesi 2016: 4). It seems that mathematical diagrams, which combine spatial and algebraic characteristics, and which allow continuous manipulations—as opposed to discrete symbolizations within formal algebraic systems—actually *embody* this associative connection in the best way possible.

One of the thinkers who pioneered the research of the role diagrams play in logical and mathematical reasoning was a logician and a mathematician Charles Sanders Peirce. In fact, Peirce went as far as to claim that there is no mathematical reasoning proper that is not diagrammatic (Peirce 1931–1958, Vol. 1, para 54; CP 2.216, CP 5.148). His system of diagrammatic logic, which he called "existential graphs," shows that diagrams are iconic signs that represent a link between visual experience, necessary reasoning, and imaginative experimentation. Thus conceived, the iconicity of diagrams, Peirce believed, tells us something important about the homological relationship between the grammar of visual language and the very machinery of thought, or thinking in actu (Peirce 1976 [NEM], p. 239; CP 4.6). In light of the above, the goal of this chapter is to delve into the question of what the iconicity of diagrams actually amounts to, or, in other words, what a mathematician actually *sees* when he or she is involved in reasoning by means of diagrams. From the Peircean perspective, to address this question is to define what diagrams are *as signs*, which characteristics they share with other signs, and which characteristics are unique to them.

Three Characteristics of Signs

According to Peirce, diagrams are iconic signs. By a "sign" Peirce broadly means anything (a thought, an emotion, a name, a mathematical formula, an existent physical object, a natural kind, a tool, the word "tool," a musical concerto, an action, a move in a chess game, a law of nature—anything at all) capable of standing *for* something else *in* some respect *to* someone who can interpret it (CP 2.228). Thus, "Santiago" in "Charles Santiago Sanders Peirce" stands for the name of a Biblical character to someone who can interpret it further as an expression of Peirce's gratitude to his friend William James. "Qf4" stands for Bobby Fischer's move in game six of the 1972 Spassky-Fischer World Chess Series, to Boris Spassky who, while pondering over his unfortunate situation, was preparing to resign. "Black Square" stands for a visual expression of the mathematical regularities that explain the basic relations between form and color in suprematism, to an art critic writing a book on Kazimir Malevich. "10" stands for *tetractys*, an equilateral triangle consisting of ten points arranged in four rows of 1, 2, 3, and 4, to a Pythagorean mathematician Eudoxus, who interpreted it as a mathematical expression of the universal cosmic harmony. The same number also stands for "impossible to be scratched by a piece of corundum" to a mineralogist as a description of diamond's hardness according to Mohs' scale, etc. (Peirce 1992–1998, Vol. 2, p. 326).

Three principal characteristics every sign possesses are important for understanding what Peirce actually means by a "diagram." First, Peirce insists that every sign, even if it is not actually interpreted, in order to be a sign, should be *capable* of being interpreted. The meaning of any and every sign thus consists in its future interpret-*ability*. From this, it follows that a sign's identity always lies in its reference to some other thing. Peirce was an idealist and, in his view, for anything at all to *be* is to stand for something else to a mind, that is—to be a sign. Given that everything, thus, is a sign, and that every sign is addressed to its possible future interpretations, an interpretation can result in nothing but the creation (or discovery) of a *new* sign. In Peirce's semiotics novelty, therefore, appears to be a necessary characteristic inherent in every act of interpretation. Second, although sign's being addressed to the possible future implies that no finite number of interpretations is exhaustive of its meaning, any and every interpretation has a short-term goal. Believing a proposition (say, "Gödel's Second Incompleteness Theorem shows that Peano arithmetic cannot prove its own consistency") is true amounts to being prepared to act habitually on this belief when the occasion presents itself (for instance, being prepared to employ a diagonal argument, or to clarify the idea of self-reference in Russell's paradox). Accordingly, the meaning of a sign depends on what habits of conduct it is going to bring about (Peirce 1992–1998, p. 432). Combining the idea of future interpretation and the idea of sign's being related to its object by means of possible conduct allows Peirce to define the meaning of a sign as the sum total of habits it would *ultimately* produce (Peirce 1992–1998, p. 346). This, in turn, brings about the definition of the ultimate aim of inquiry. In Peirce's own words, this aim consists of:

> the desire to get a settlement of opinion in some conclusion which shall be independent of all individual limitations, independent of caprice, of tyranny, of accidents of situation, of initial conditions, which does not confirm any belief but unsettles and then settles,—a conclusion to which every man would come who should pursue the same method and push it far enough (Peirce 1982, Vol. 3, p. 19).

What underlies the idea of the end of inquiry is *statistical reasoning*. As Peirce explains, "judging of the statistical composition of a whole lot from a sample is judging by a method which will be right on the average in the long run, and, by the reasoning of the doctrine of chances, will be nearly right oftener than it will be far from right" (CP 1.93). No matter where different members of a community of inquirers may begin, as long as they follow a certain method, the results of their research should eventually converge toward the same outcome. The method is formulated in Peirce's maxim of pragmatism: "Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object" (Peirce 1982, Vol. 3, p. 266). This formulation echoes Peirce's definition of sign as it suggests that meanings of our ideas depend on our capacity to predict outcomes of our experiments with the objects of those ideas, and to act on the outcomes. Because our capacity to interpret signs assumes our capacity to predict the outcomes of our possible future actions, signs bring habits about but, at the same time, they are catalysts that cause those habits to be reinforced or abandoned. Consequently, interpretation that follows the method, as it is described in Peirce's maxim, is self-corrective, i.e., characterized by self-controlled, habit-driven action.

Now what any language—whether natural, such as English, or mathematical, such as algebra—prima facie does is it brings individual objects under general concepts. Peircean signs refer to their objects through habitual action. According to the maxim quoted above, all our general idea of a thing amounts to is an account of our would-be responses to the changes resulting from our experiments with this thing. Signs, therefore, act like any language, with the only difference that they refer to their objects not through arbitrary convention, but through adaptive habitualized behavior. The relationship between the general and the particular through habits is the third characteristic of signs that is of importance for the idea of a mathematical diagram, to be discussed below.

Iconicity and Habitual Action

As iconic signs, diagrams are related to their objects by means of *likeness*. But what does it mean? What does likeness, revealed in vision (and, more broadly, in perception), amount to? Peirce himself recognized that there is a quandary here: "I myself happen, in common with a small but select circle, to be a pragmatist, or 'radical empiricist,' and as such, do not believe in anything that I do not (I think) perceive. … Only, the question arises, *What* do we perceive?" (CP 7.617–618; emphasis added). If we confine iconicity to the domain of vision only, we face a problem. For instance, in comparing what we see in a portrait of a mathematician Christiaan Huygens by Caspar Netscher, with what we would see if we lived in the seventeenth century and actually met Huygens, we juxtapose a two-dimensional image and a real person. In none of its parts the portrait is "like" Huygens himself. No matter how far we go into detail in trying to explain what particular feature serves as a ground for comparison, the "what" of similarity always escapes analysis (Eco 1992: 191–216). It should be realized, therefore, that when we are talking about similarity, we are talking about some sort of *perceptual schematism* rather than vision *simpliciter*. Peirce's principal suggestion is that what underlies this schematism is the isomorphism of not of substances, but of *relations* (Peirce 1992–1998, Vol. 2, p. 13; Stjernfelt 2007: 50–77; Paavola 2011); and he claims that maps, charts, geometric diagrams, and mathematical equations are primary examples of this isomorphism (NEM, p. xv; CP 4.530). Thus, for instance, the relation between points on a map of Toronto is isomorphic with the relation between the corresponding places on the Earth's surface, just as the relation between candles on a NASDAQ chart is isomorphic with the moves of the index's price over time. Similarly, the mathematical function "__is a square of__," or $f(x) = x^2$, is a mapping rule for a set of ordered pairs, in which one element is mapped onto the other, so that <2, 4> is followed by <3, 9>, etc. (Bradley 2004: 71–73). Imaging is no different. A feature, with respect to which a portrait is like its object, is always dynamic. It is a result of mapping of one set of relations onto another, revealing a character of the portrayed person based on the schematization of an imagined change. From this perspective, diagrams are iconic signs precisely because the similarity they convey reflects not the way their objects look, but the way they behave, their *modus operandi*: viewed relationally, diagrammatic iconicity is about how we *read* charts, *use* maps, *do* math, and *decode* a facial expression on a portrait. Diagrams copy the way their objects behave and, therefore, just like other signs, ultimately refer to habitual action.

A special case in this respect is represented by Peirce's existential graphs. The graphs replace formalized linear successions of syllogistic structures with a set of diagrammatic pictures in a state of constant transformation; they are conceived as schematic visual expressions of relations inherent in the action of *thinking itself*. The proper object of the graphs is thus the very machinery of ratiocination:

> It is requisite that the reader should fully understand the relation of thought in itself to *thinking*, on the one hand, and to graphs, on the other hand. Those relations being once magisterially grasped, it will be seen that the graphs break to pieces all the really serious barriers, not only to the logical analysis of thought, but also to the digestion of a different lesson, by rendering literally visible before one's very eyes the operation of thinking *in actu* (CP 4.6).

Graphic, nonlinear reading, as exhibited in the graphs, is important for a mathematician in several respects. First, Peirce's graphs are so designed as to show an immediate logical continuity of thinking in the form of a dialogue between two imaginary parties: the graphist and the interpreter. As Peirce puts it, "thinking always proceeds in the form of a dialogue—a dialogue between different phases of the ego" (ibid.). To this end, the graphs replace literal signification of functions, variables, and quantifiers with shapes mapped onto each other and composed of a variety of graphical conventions. The most basic conventions of the existential graphs are represented by the *sheet of assertion* (a blank sheet on which all graphs are being created); a *cut*, or negation (a linear separation that cuts whatever it encloses off from the sheet of assertion); a *line of identity*, asserting the existence of the individuals denoted by its extremities; and a *ligature* (a connection of two or more lines of identity). The empty sheet of assertion assumes that the truth of whatever is stated on it is a matter of agreement between the graphist and the interpreter. The empty sheet of assertion is, therefore, itself a graph. A cut that encloses an area of the sheet of assertions that is empty, i.e., contains no statement, is the pseudograph as it represents a statement that negates its own truth. In manipulating the graphs, a mathematician can actually *observe* a given argument visualized by a set of transformational rules as a number of continuously transforming pictures, and *experience* the meaning of the argument visually as a set of transitional states. It is for this reason that Peirce sometimes called his graphs the "moving pictures of thought" (CP 4.8), or "a portraiture of thought" (CP 4.11).

An important feature of the graphs is that the conventions and transformational rules, which constitute the grammar of the graphs, are devised as a surface structure that is not separated from what the graphs actually convey. In other words, the logical form of every graph appears to be an integral part of its overall message. Every graph thus conveys information and simultaneously provides a key to how this information is to be decoded. To use Marshall McLuhan's catch phrase, in this case, truly, "the medium is the message" (1994: 7–21). *Seeing* something and *understanding* how it works, or *what* is stated and *how* the statement is constructed, is a matter of one and the same act. Because every graph (including the sheet of assertions and the pseudograph) performs a visualization of the way its messages are encoded, using the graphs, or the moving pictures of thought, blurs the distinctions between the internal and the external, ratiocination and observation, and code and message. This feature of the graphs reflects a general characteristic Peirce ascribed to all iconic signs:

> Icons are so completely substituted for their objects as hardly to be distinguished from them. Such are the diagrams of geometry. A diagram, indeed, so far as it has a general signification, is not a pure icon; but in the middle part of our reasonings we forget that abstractness in great measure, and the diagram is for us the very thing. So in contemplating a painting, there is a moment when we lose the consciousness that it is not the thing, the distinction of the real and the copy disappears, and it is for the moment a pure dream,—not any particular existence, and yet not general. At that moment we are contemplating an *icon* (Peirce 1992–1998, Vol. 1, p. 226)

To rephrase, the likeness is neither in the sign, nor in the object, but in the way the two are brought together by an interpreting mind relative to some practical purpose it has (Parker 2017, p. 68). Given this, it might be claimed that the graphs are diagrams *par excellence*. The ground for the iconicity of a graph is the isomorphism between an experimental change in the relations between the parts of the graph and the corresponding transformation in its object—as is the case with diagrams in general. An important difference though is that the object of existential graphs is the thinking process, or the process of imaginative experimentation itself. The graphs do not simply show transitions from one thought to another so that we could further translate them into a formal language; they represent an identity between the action of thought and the continuity of movement in space exhibited in a graph (Paolucci 2017: 84–85).

Iconicity and Novelty

Peirce treated visual perceptions (and perceptions in general for that matter) as results of unconscious inferences (Hull 2017: 150). In his view, any percept is essentially a product of a long history of gradually habitualized, piecemeal adjustments to the ever-changing environment. If every visual experience is thus a readjustment, it is, as Peirce puts it, “constructed at the suggestion of previous sensations” that “are quite inadequate to forming an image or representation absolutely determinate” (Peirce 1982, Vol. 2, p. 235). Peirce’s overall conclusion is that “when we see, we are put in a condition in which we are able to get a very large and perhaps indefinitely great amount of knowledge of the visible qualities of objects” (Peirce 1982, p. 236). This being the fact, according to Peirce, “either we perceive some indeterminate properties or we perceive nothing at all” (Wilson 2017: 16).

Peirce’s view on the indeterminacy of perception is important in two respects. One will be discussed in the next section with regard to the relationship between the general and the particular in diagrammatic reasoning. Another is that the indeterminacy of perception implies that every act of visual experience—although what it delivers cannot be changed at will—presupposes interpretation and leaves space for errors, interpretive hypotheses, and imaginative musings about its object (Paavola 2011: 305; Vargas 2017). Naturally, if all perception involves interpretation, and if iconicity is an integral part of all reasoning, then diagrammatic aspects of reasoning, according to Peirce, are responsible for the creativity not only of mathematical cognition, but also of human cognition in general (Paavola 2011: 298). Peirce claims that, as far as knowledge is expressed in some language, “in the syntax of every language there are logical icons of the kind that are aided by conventional rules” (CP 2.280). Yet mathematical diagrams represent a unique case in this respect. On the one hand, they are deductions, and their primary goal is to represent patterns of deductive thinking, to express necessity. On the other hand, they are also capable of introducing new truths (CP 4.233). Peirce believed that this is explained by the fact that mathematical deductive reasoning, expressed diagrammatically, always involves *observation*:

> It has long been a puzzle how it could be that, on the one hand, mathematics is purely deductive in its nature, and draws its conclusions apodictically, while on the other hand, it presents as rich and apparently unending a series of surprising discoveries as any observational science. Various have been the attempts to solve the paradox by breaking down one or other of these assertions … The truth, however, appears to be that all deductive reasoning … involves an element of observation; namely, deduction consists in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of experimenting upon this image in the imagination, and of observing the result so as to discover unnoticed and hidden relations among the parts (CP 3.363).

Peirce even admits that the kind of experimentation a mathematician performs with diagrams is analogous to the kind of experimentation that is implemented in physical sciences:

> [R]easoning of much power has, as a historical fact, never been performed by means of words, or other sounds, nor even to any great extent by means of pure retinal sensations, but by means of muscular sensations and visual images which have in the imagination been put in motion, so that a sort of imaginary experiment is made; and the result has been observed inwardly, as that of a physical experiment is outwardly (NEM, p. 378; cf. CP 4.530).

This is true, Peirce says, not only in the case of geometry, but also in the case of syllogistic structures and algebraic equations. In fact, Peirce claims that, in *any* particular instance of mathematical reasoning, "there must be something amounting to a diagram before the mind's eye," and that "the act of inference consists in *observing* a relation between parts of that diagram that had not entered into the design of its construction" (NEM 4:353; cf. CP 2:279). For example, a particular case of "Barbara," written down correctly, represents a simple diagram that clearly shows the relationship between the three terms involved and actually *exhibits* the fact that the middle term occurs in both premises. Likewise, an algebraic equation is considered a rule that maps one relation between variables onto another in such a way that further manipulation could lead to the discovery of an unending series of new facts. A simple geometrical example would be Pythagoras' theorem. There are numerous proofs of this theorem, but the majority of them require that, in order to explain the relation among the three sides of a right triangle, a geometer should make a certain *rearrangement*; in the initial, Pythagoras's own version of the proof was the rearrangement of the four identical right triangles whose hypotenuses form a square. Based on these, as well as other, more complicated examples, Peirce shows that it is never the case that, in solving a mathematical problem, just thinking in general terms is enough. "It is necessary," he says, "that something should be *done*. In geometry, subsidiary lines are drawn. In algebra, permissible transformations are made. Thereupon, the faculty of observation is called into play. Some relation between the parts of the schema is remarked" (CP 4:233; Hull 2017: 149; Joswick 1988: 113).

Mathematics can discover new regularities due to the following two features that diagrams exhibit. First, because there is always an array of possible transformations that are implied by the very way a given diagram is constructed. Second, because we cannot predict in advance what particular transformations out of the array will be enacted, and what the ultimate result of those transformations will be (Stjernfelt 2007: 81–83). What these two features imply is that mathematics essentially is an *activity*, a habit-driven, and yet creative *practice* rather than a static deductive grammar that supplies rules for the contemplation of abstract mathematical forms (Campos 2009; Hull 2017). Within mathematical reasoning as a practice, imagination, in turn, has a threefold role to play. First, a mathematician forms a *skeletonized* iconic representation, a diagram, whether geometrical or algebraic, of the facts he or she is interested in considering. The principal purpose of the initial skeletonization of the problem for a mathematician, Peirce says, "is to strip the significant relations of all disguise" (CP 3.359). Second, a mathematician observes this diagrammatic picture until, at some point, "a hypothesis suggests itself that there is a certain relation between some of its parts." Third, he or she experiments upon the diagram in order to test his or her hypothesis, so that "it is *seen* that the conclusion is compelled to be true by the conditions of the construction of the diagram" (CP 2.278; cf. CP 3.560; Joswick 1988: 108–109).

To summarize, iconic signs in general, and diagrams in particular, show some relations that are constitutive of their objects and, at the same time, hide some others that may be discovered later. Consequently, on the one hand, a diagram connects us to inexhaustible possibilities of further interpretation. On the other hand, the way mathematicians construct their diagrams, together with the transformational rules implied by the construction, makes it the case that where the construction ultimately leads us is beyond our idiosyncrasies and individual whims. Diagrammatization, thus, may be understood as "a sort of self-controlled management of one's own thoughts, because the clarity and exactitude that is realized eventually in diagrams, in turn furthers the exactness of cognitive activities" (Hoffmann 2004: 133). More importantly, in the process of manipulation with diagrams, spatial imagination and abstract reasoning are presented not as two distinct mental faculties, but as two aspects of the same activity put to work together. The point is aptly summarized by Kathleen Hull: "Peirce's conception of a diagram is fundamentally and inseparably both conceptual *and* spatial insofar as reasoning by diagrams engages the continuum of spatial extension in the reasoning process" (op. cit., p. 147). Mathematics, in other words, is a practice that makes use of a set of particular cognitive mechanisms in order to creatively schematize together images and abstractions. This conclusion leads us to the third feature that diagrams share with other kinds of signs.

The General and the Particular

A diagram, unlike a word or a sound, represents a relation as a kind of connection between certain elements in immediate perception. Every instance of such representation is a skeleton idea of relations between certain things either in the mind or out there in the world (CP 7.426). On the one hand, a visually represented relation carries with it the full history of *particular* habitual responses that made it possible. On the other hand, the *kind* of connection it visually represents is an object of the general nature. Consequently, it might be said that, in some sense, generality is—in the case of diagrams, visually—perceived. Peirce is aware that, for an empiricist, this claim might seem not just improbable, but utterly meaningless. As he puts it, "Bishop Berkeley and a great many clear thinkers laugh at the idea of our being able to imagine a triangle that is neither equilateral, isosceles, nor scalene. They seem to think the object of imagination must be precisely determinate in every respect" (CP 5.371). But he still insists that some form of the realism about generality is indispensable.

Here is one way to describe the reasoning behind this belief. In any kind of observation, including mathematical, there is a uniformity in things we observe. This uniformity is due to the fact that any observation, in referring to its object, points at some general way to respond to its results every time certain conditions hold. But to be able to refer to an object in this way is to be a sign. Therefore, the uniformity implied by observation is due to the fact that we are capable of interpreting what we observe as a train of signs. As has been discussed above, any sign, as far as it is interpretable, is an *esse in futuro*. It *is* what it is *going to become* as interpreted in the future. Peirce's early account in "Some Consequences of Four Incapacities" offers the following definition of *thoughts* as signs:

> No present actual thought (which is a mere feeling) has any meaning, any intellectual value; for this lies not in what is actually thought, but in what this thought might be connected with in representation by subsequent thoughts … At no one instant in my state of mind is there cognition or representation, but in the *relation* of my states of different instants there is (Peirce 1992–1998, Vol. 1, p. 42).

There is nothing absolutely singular in thought, as what it is in the present moment amounts merely to a feeling, and anything of value lies in conditional expectations implied by its possible future interpretation. And, according to Peirce, the same is true of perception. What we are dealing with in perception amounts to "perceptual facts," which already contain an inferential element, rather than mere "percepts:"

> In place of the *percept,* which … is a construction with which my will has had nothing to do, and may, therefore, properly be called the "evidence of my senses," the only thing I carry away with me is the *perceptual facts,* or the intellect's description of the evidence of the senses, made by my endeavor. These perceptual facts are wholly unlike the percept, at best; and they may be downright untrue to the percept. But I have no means whatever of criticizing, correcting or recomparing them, except that I can collect new perceptual facts relating to new percepts, and on that basis may infer that there must have been some error in the former reports. … The perceptual facts are a very imperfect report of the percepts; but I cannot go behind that record. As for going back to the first impressions of sense, as some logicians recommend me to do, that would be the most chimerical of undertakings (CP 2.141).

There is nothing, then, that is absolutely raw and singular in perception at any given moment, as anything that is absolutely singular in perception at any given moment is only a brute existence, a purely denotative "this." The very immediacy of its presence does not allow us to say anything *about* it and, therefore, to treat it as a sign. Every visual perception proper contains something expected that is inseparable from what this visual perception, allegedly, simply *is* (CP 2.146). This means that generality, in the form of a potential or conditional future, is given to us in visual experience.

The kind of generality that is involved in mathematical observation is *necessity*. Expressed in diagrams, the deductive must of a conclusion becomes visually evident. And the evidence of such conclusion "consists in the fact that the truth of the conclusion is *perceived*, in all its generality, and in the generality the how and the why of the truth is perceived" (NEM, p. 317). What underlies this perspective is that, on the one hand, it does not seem right to reduce mathematical perception to a psychological process, i.e., to resort to some sort of intuitionist understanding of mathematics, according to which the truth of a relation amounts to a subjective claim. On the other hand, although there are objective, mind-independent relations out there that represent necessity, mathematical thought is always embodied in token signs. Just as it is impossible to get at the heart of an onion by peeling off all its skins, it is impossible to get at the heart of a necessary relation by stripping it off whatever particular symbols that happen to signify it (CP 4.6, 87). Peirce's overall point is that mathematical necessity cannot be reduced to a description simply attached to an arrangement of individual objects; it should be considered a real characteristic of mathematical signs *qua* signs—a characteristic that is available in an act of observation. Only, according to Peirce, what distinguishes diagrammatic signs is that they achieve the fusion of the general and the particular most effectively: by making this fusion visually available. As Catherine Legg (2012: 1) aptly puts it, in the case of diagrams, "[n]ecessary reasoning is in essence just a recognition that a certain structure has the particular structure that it in fact has."

Logic and Mathematics: The Perception of Totality

In "Some Amazing Mazes" (*The Monist*, 1908), Peirce writes:

> But mathematicians are rather seldom logicians or much interested in logic; for the two habits of mind are directly the reverse of each other; and consequently a mathematician does not care to go to the trouble (which would often be very considerable) of ascertaining whether the theoric step he proposes to himself to take is absolutely indispensable or not, so long as he clearly perceives that it will be exceedingly convenient; and the consequence is that many demonstrations introduce theoric steps which relieve the mind and obviate confusing complications without being logically necessary (CP 4:614).

Two years earlier, the same distinction was mentioned in Peirce's diary:

> The distinction between the two conflicting aims [of logic and mathematics] results from this, that the mathematical demonstrator seeks nothing but the solution of his problem; and, of course, desires to reach that goal in the smallest possible number of steps; while what the logician wishes to ascertain is what are the distinctly different elementary steps into which every necessary reasoning can be broken up …. In short, the mathematician wants a pair of seven-league boots, so as to get over the ground as expeditiously as possible. The logician has no purpose of getting over the ground: he regards an offered demonstration as a bridge over a canyon, and himself as the inspector who must narrowly examine every element of the truss because the whole is in danger unless every tie and every strut is not only correct in theory, but also flawless in execution (Fisch Chronological File (n.d.), Fragment on logician and mathematician, c. 1906).

Mathematicians are often reluctant to build bridges over the canyons filled with formal complications. What they want is not to examine every element of the logical truss, and not to dissect the process of reasoning into its simplest steps, but to discover the fastest and the most efficient way to prove (or discard) their assumptions. Seen from this perspective, the distinction Peirce aims at here is the one between the mathematical *practice* of making inferences and the logical *theory* that has those inferences as its objects of study. As noted by Kulpa (2009: 76):

> mathematicians usually produce informal proofs using much intuition and informal leaps of imagination, but still maintaining a certain discipline and rigour that convinces them that

> the result in principle *can* be formalized if need be. However, it is hard to hear a convincing answer to the question what exactly makes them so sure of that possibility.

Of course, mathematical intuition should not be perceived as being simply at odds with the capacity to produce long strings of formal proofs. Yet the distinction between the two is salient. For instance, when it comes to a computerized rewrite of a solution to a nontrivial mathematical problem, the absence of the intuitive guidance that initially paved the way to the solution leads to the exponential growth in the number of possible rewrites that are, at times, too much to handle (Kulpa 2009). Mathematicians thus tend to make justified shortcuts in their demonstrations—and this habit is nothing new. It was just as much in use among mathematicians back in Peirce's day. As Poincaré (2009: 178), for instance, once wrote, "If it requires 27 equations to establish that 1 is a number, how many will it require to demonstrate a real theorem?" Peirce himself was brought up in a family with two other mathematicians: his older brother James Mills and his father Benjamin, who is best known for the so-called "Peirce Criterion" used in statistics for the elimination of suspect experimental data. The incomprehensibility and hermetic character of Benjamin's lectures at Harvard were the subject of many legends and anecdotes. As a colleague of the elder Peirce once wrote:

> his intuition of the whole ground was so keen and comprehensive that he could not take cognizance of the slow and tentative process of mind by which an ordinary learner was compelled to make his step-by-step progress. In his explanations he would take giant strides; and his frequent "*you see*" indicated what he saw clearly, but that of which his pupils could get hardly a glimpse (Cajori 1890: 139; emphasis added).

What do, then, mathematicians *see* when they use diagrams in solving mathematical problems? They see relational schemata whose very construction, along with the transformational grammar implied by the construction, prompts certain changes that lead to the discovery of new relations among the parts of the schemata. Mathematicians approach problem situations by creating sets of skeletonized images in which not just particular moves, but certain general pattern dynamics is anticipated. But there is one more respect in which diagrammatic representation is of service for mathematical vision, and which may explain the likelihood of the aforementioned "giant strides" and "leaps of imagination" in mathematical demonstrations. What experimenting with diagrams allows mathematicians to see is not lengthy successions of discrete images following one another, but something that Gilles Deleuze, in trying to integrate Peirce's theory of signs into his interpretation of Henry Bergson's metaphysics of time, describes as *l'image-mouvement*.

A movement image, according to Deleuze, reflects not just a relation between different positions that the objects involved take at different moments, but a relation between the general patterns and the behavior those objects reveal and the corresponding qualitative changes of the image *as a whole* (Deleuze 1986: 56–70). In commenting on Bergson's *Matter and Memory* (1896), Deleuze (1986: 7) says that, "[when] one constructs a Whole, one assumes that 'all is given,' whilst movement only occurs if the whole is neither given nor giveable. As soon as the whole is given to one in the eternal order of forms or poses, or in the set of any-instant-whatevers, … there is no longer room for real movement." A movement image, according to Deleuze, is an image that cancels the dichotomy between the whole and changes in its parts because it is actually expressed as a whole by the very continuity of the transformation of its parts. What Deleuze's analysis implies is that a movement image *is* a movement, *conceived* or *anticipated* movement, enacted by the mathematical imagination. A diagram or a graph ("a moving picture of thought"), understood as such an image, unlike a string of discrete symbols, presents its meaning as accessible in its totality at any given *hic et nunc* during its transformation. We do not *read* it linearly, but *see* it in its entirety, with some stages already in the past, but still controlled in retention, and some other stages in the future, but already grasped in anticipatory protention. A mathematical graph is thus an image neither of a final state of affairs, nor of a passing instant. It is wholly determined by the Bergsonian *durée*, in which the past, the present, and the future are grasped together in the *dynamic totality* of a graph. The parts of it are in space, but the whole of it is in duration and change.

Conclusion

Due to a rich set of two-dimensional graphical properties, along with appropriate transformational rules and conventions, diagrams allow for creating structures that are more versatile and more efficient in solving mathematical problems than one-dimensional symbol strings. Moreover, diagrams exhibit a number of phenomena not occurring in the world of algebraic formalization. They combine deductive necessity and the capacity to generate new meanings, a priori ratiocination and visual perception, the visual grasp of the dynamic totality of a given argument, and perception of particular changes in its structure. The means by which we construct diagrams are our means, yet where this construction ultimately leads us is beyond our individual whims. Diagrammatic reasoning shows us that visual perceptions are, by nature, inferences, and that there is certain logic to their arrangement, such that, by following it, we can visualize the very continuity of thinking.

Peirce was one of the earliest authors on visual perception whose account included kinds, relations, modal qualities, and other kinds of general objects. He claimed that anything general is, by definition, relational, and that a relation is something that can always be represented visually. Overall, apart from all the specifics discussed above, Peirce's analysis shows that there is a dynamic element to perception, which serves as a bedrock for visual integration. Motion can tell us more than where an object is going, whether this object is a stock price, a point on a map, a mathematical function, or thought itself. It can also tell us what the object is. This is as true of mathematics as it is of the architecture of ordinary visual recognition: a pencil bouncing on a table, a butterfly in flight, or a closing door supports the recognition of these objects as such. An object and its stereotypical, habit-driven motion in some important sense make the object what it really is. Finally, what Peirce implies in the case of the graphs is that this dynamic element, along with the aforementioned features of visual perception, allows us to grasp the continuity of thinking itself. *Moving* pictures are needed in order to turn thought into a proper object. If there is a system of graphic conventions, there has to be a corresponding system of moves. Only then thinking can be caught in action. In what it *does*.

References

Abrahamsen, A. and Bechtel, W. (2015). Diagrams as tools for scientific reasoning *Review of Philosophy and Psychology* 6(1): 117–131.

Bakker, A. and Hoffmann, M. (2005). Diagrammatic reasoning as the basis for developing concepts: A semiotic analysis of students' learning about statistical distribution. *Educational Studies in Mathematics* 60: 333–358.

Barker-Plummer, D. (1997). The role of diagrams in mathematical proofs. *Machine Graphics and Vision* 6(1): 25–56.

Bradley, J. (2004). The generalization of the mathematical function: A speculative analysis. In: G. Debrock (ed.), *Process pragmatism. Essays of a quiet philosophical revolution*, pp. 71–86. Amsterdam, The Netherlands: Rodopi.

Cajori, F. (1890). *The teaching and history of mathematics in the United States*. Washington, D.C.: The U.S Government Printing Office.

Campos, D. (2009). Imagination, concentration, and generalization: Peirce on the reasoning abilities of the mathematician. *Transactions of the Charles S. Peirce Society* 45(2): 135–156

Danesi, M. (2016a). *Language and mathematics. An interdisciplinary guide*. Berlin: De Gruyter.

Danesi, M. (2016). *Learning and teaching mathematics in the global village. Math education in the digital age*. New York, NY: Springer.

Deleuze, G. (1986). *Cinema 1: The movement image* (translated by Hugh Tomlinson and Barbara Habberjam). Minneapolis, MN: University of Minnesota Press.

Eco, U. (1992). *A theory of semiotics*. Bloomington, IN: Indiana University Press.

Hegarty, M Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4): 684–689.

Fisch, M. H. (n.d.) Chronological file, Institute for American Thought, School of Liberal Arts, IUPUI, Indianapolis, IN.

Hoffmann, M. (2004). Peirce's "diagrammatic reasoning:" A solution of the learning paradox. In G. Debrock (ed.), *Process pragmatism. Essays of a quiet philosophical revolution*, pp. 121–144. Amsterdam, The Netherlands: Rodopi.

Hull, K. (2017). The iconic Peirce: Geometry, spatial intuition, and visual imagination. In: K. Hull and R. Atkins (eds.), *Peirce on Perception and Reasoning: From Icons to Logic* (147–173). New York, NY: Routledge.

Joswick, H. (1988). Peirce's mathematical model of interpretation. *Transactions of the Charles S. Peirce Society* 24(1): 107–121.

Kulpa, Z. (2009). Main problems of diagrammatic reasoning. Part I: The generalization problem. *Foundations of Science* 14(1): 75–96.

Lakoff, G. (1999). *Philosophy in flesh: The embodied mind and its challenge to Western thought*. New York, NY: Basic Books.

Lakoff, G. and Núñez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

Legg, C. (2012). The Hardness of the Iconic Must: Can Peirce's Existential Graphs Assist Modal Epistemology? *Philosophia Mathematica* 20, 1–24.
Legg, C. (2017). "Diagrammatic Teaching:" The role of iconic signs in meaningful pedagogy. In: I. Semetsky (ed.), *Edusemiotics: A handbook*, pp. 29–46. Singapore: Springer.
McLuhan, M. (1994). *Understanding media: The extensions of man*. Cambridge, MA: The MIT Press.
Mumma, J. (2010). Proofs, pictures, and Euclid. *Synthèse* 175(2): 255–287.
Nersessian, N. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In: R. Giere (ed.), *Cognitive models of science*, pp. 5–22. Minneapolis, MN: University of Minnesota Press.
Paavola, S. (2011). Diagrams, iconicity, and abductive discovery. *Semiotica* 186: 297–314.
Paolucci, C. (2017). Semiotics, schemata, diagrams, and graphs: A new form of diagrammatic Kantism by Peirce. In: K. Hull and R. Atkins (eds.), *Peirce on perception and reasoning: From icons to logic* pp. 74–85. New York, NY: Routledge.
Parker, K. (2017). Foundations for semeiotic aesthetics: Mimesis and iconicity. In: K. Hull and R. Atkins (eds.), *Peirce on perception and reasoning: From icons to logic,* pp. 61–73. New York, NY: Routledge.
Peirce, C. S. (1992–1998). *The essential Peirce*. *Selected philosophical writings.* Vols. 1-2. N. Houser and C. Kloesel (eds.). Bloomington, IN: Indiana University Press.
Peirce, C. S. (1982). *Writings of Charles S. Peirce. A chronological edition*. Vols. 1–6. M. H. Fisch, E. Moore, C. Kloesel, and Peirce Edition Project (eds.). Bloomington, IN: Indiana University Press.
Peirce, C. S. (1976). *The new elements of mathematics*. C. Eisele (ed.). Vol. 4. The Hague: Mouton.
Peirce, C. S. (1931–1958). *Collected papers of Charles Sanders Peirce*. Vols. 1–8. C. Hartshorne, P. Weiss, and A. Burks (eds.). Cambridge, MA: Harvard University Press.
Poincaré, H. (2009). *Science and method* (trans. by F. Maitland). New York: Cosimo Classics.
Prusak, N. (2012). From visual reasoning to logical necessity through argumentative design. *Educational Studies in Mathematics* 79(1): 19–40.
Sherry, D. (2009). The role of diagrams in mathematical arguments *Foundations of Science* 14(1–2): 59–74.
Sloman, A. (2002). Diagrams in the mind? In: M. Anderson, B. Meyer and P. Olivier (eds.). *Diagrammatic representation and reasoning*, pp. 7–28. London: Springer-Verlag.
Stjernfelt, F. (2007). *Diagrammatology: An investigation on the borderlines of phenomenology, ontology, and semiotics*. Dordrecht, The Netherlands: Springer.
Vargas, E. (2017). Perception as inference. In K. Hull and R. Atkins (eds.), *Peirce on perception and reasoning: From icons to logic.* New York, NY: Routledge.
Wilson, A. (2017). What do we perceive? How Peirce "expands our perception". In K. Hull and R. Atkins (eds.), *Peirce on perception and reasoning: From icons to logic***,** pp. 1–13. New York, NY: Routledge.

Chapter 9
Laws of Form, Peirce, and Cantor

Louis H. Kauffman

Introduction

The purpose of this chapter is to explore the idea of a sign, using G. Spencer-Brown's (1969) work "Laws of Form" is a pivot, a reference, and a place from which to make excursions into both simplicity and complexity. In order to handle the simplicity of the issues involved in thinking about distinction, Spencer-Brown's introduction of a language that has only one sign is an instrument of great delicacy.

The Spencer-Brown mark ⏋ is a sign that can represent any sign, and so begins semiotics in both universal and particular modes. The mark is seen to make a distinction in the space in which it is written, and so can be seen, through this distinction, to refer to itself. In the language of Charles Sanders Peirce, the mark is its own representamen and it is also its own interpretant. The sign that the mark produces for somebody is, in its form, the mark itself. By starting with the idea of distinction we find, in the mark, the first sign and the beginning of all possible signs.

> Right, that's how long it takes, not a day less,—Qfwfq said,—once, as I went past, I drew a sign at a point in space, just so I could find it again two hundred million years later, when we went by the next time around. What sort of sign? It's hard to explain because if I say sign to you, you immediately think of a something that can be distinguished from a something else, but nothing could be distinguished from anything there; you immediately think of a sign made with some implement or with your hands, and then when you take the implement or your hands away, the sign remains, but in those days there were no implements or even hands, or teeth, or noses, all things that came along afterwards, a long time afterwards. As to the form a sign should have, you say it's no problem because, whatever form it may be given, a sign only has to serve as a sign, that is, be different or else the same as other signs: here again it's easy for you young ones to talk, but in that period I didn't have any examples to follow, I couldn't say I'll make it the same or I'll make it different, there were no things

L. H. Kauffman (✉)
University of Illinois at Chicago, Chicago, IL, USA

Novosibirsk State University, Novosibirsk, Russia
e-mail: kauffman@uic.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_9

to copy, nobody knew what a line was, straight or curved, or even a dot, or a protuberance or a cavity. I conceived the idea of making a sign, that's true enough, or rather, I conceived the idea of considering a sign a something that I felt like making, so when, at that point in space and not in another, I made something, meaning to make a sign, it turned out that I really had made a sign, after all.

In other words, considering it was the first sign ever made in the universe, or at least in the circuit of the Milky Way, I must admit it came out very well. Visible? What a question! Who had eyes to see with in those days? Nothing had ever been seen by anything, the question never even arose. Recognizable, yes, beyond any possibility of error: because all the other points in space were the same, indistinguishable, and instead, this one had the sign on it.

I thought about it day and night; in fact, I couldn't think about anything else; actually, this was the first opportunity I had had to think something; or I should say: to think something had never been possible, first because there were no things to think about, and second because signs to think of them by were lacking, but from the moment there was that sign, it was possible for someone thinking to think of a sign, and therefore that one, in the sense that the sign was the thing you could think about and also the sign of the thing thought, namely, itself.

So the situation was this: the sign served to mark a place but at the same time it meant that in that place there was a sign (something far more important because there were plenty of places but there was only one sign) and also at the same time that sign was mine, the sign of me, because it was the only sign I had ever made and I was the only one who had ever made signs. It was like a name, the name of that point, and also my name that I had signed on that spot; in short, it was the only name available for everything that required a name.

In the universe now there was no longer a container and a thing contained, but only a general thickness of signs superimposed and coagulated, occupying the whole volume of space; it was constantly being dotted, minutely, a network of lines and scratches and reliefs and engravings; the universe was scrawled over on all sides, along all its dimensions. There was no longer any way to establish a point of reference: the Galaxy went on turning but I could no longer count the revolutions, any point could be the point of departure, any sign heaped up with the others could be mine, but discovering it would have served no purpose, because it was clear that, independent of signs, space didn't exist and perhaps had never existed. [A Sign in Space, *Cosmicomix* by Italo Calvino. Copyright © 1965 by Giulio *Einaudi Editore, S.p.A. English translation copyright © 1968 by Harcourt Brace & Company and Jonathan Cape Limited*] (Calvino(1965).

Finding Distinction

We begin by discussing (the idea of) distinction. If one looks for the definition of distinction in any dictionary the result is circular. Distinction is defined as a difference. Difference is defined as a form of distinction. The meaning of distinction as an indication of outstanding value is also an instance of special difference. Fields of study are founded in the use and examination of certain basic distinctions.

Mathematics is constructed set theoretically by using the concept of a collection. A collection is a distinction of membership. For example the set of prime numbers connotes the distinction between composite and prime among the positive integers. At the level of sets themselves, the empty set, denoted by brackets containing nothing { }, is a distinction between void and an empty container. The very sign for the empty set consists of two brackets (left and right) that together can be interpreted as

Fig. 9.1 Brackets

{ }

a container for something that is placed between them. In the case of the empty set, nothing is placed between the brackets. The brackets themselves are shaped as cusps (Fig. 9.1):

Each cusp can be seen as a process of bifurcation that gives rise to the distinction between the branches of the cusp. The two cusps (brackets) are mirror imaged with respect to one another, and it is this symmetry across an imaginary mirror between them that gives us the possibility to see them together as one container. The brackets are two and yet they are one (via the mirror symmetry).

At this point (in the encounter with the empty set) we reach a semantic divide between the mode of speaking of mathematicians trained in logical formalism and a wider analysis of language that I refer to as semiotic. In speaking of semiotics I am relying for its root meanings as expressed by Charles Sanders Peirce:

> [Semiotics is a] quasi-necessary, or formal doctrine of signs … which abstracts what must be the characters of all signs used by an intelligence capable of learning by experience, … and which is philosophical logic pursued in terms of signs and sign processes [Peirce, C. S., *Collected Papers of Charles Sanders Peirce*, vol. 2, paragraph 227].

> A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign. That sign which it creates I call the interpretant of the first sign. The sign stands for something, its object not in all respects, but in reference to a sort of idea which I have sometimes called the ground of the representation [Peirce –Vol. 2, p. 228].

Peirce goes on to say:

> The object of representation can be nothing but a representation of which the first representation is the interpretant. But an endless series of representations, each representing the one behind it, may be conceived to have an absolute object as its limit. The meaning of a representation can be nothing but a representation. In fact, it is nothing but the representation itself conceived as stripped of irrelevant clothing. But this clothing never can be completely stripped off; it is only changed for something more diaphanous. So there is an infinite regression here. Finally, the interpretant is nothing but another representation to which the torch of truth is handled along; and as representation, it has its interpretant again. Lo, another infinite series [Peirce – Vol. 1, p. 339] (Peirce (1931–1966)).

Peirce concentrates on the structure of signs and for him signs either are or stand for certain distinctions. To begin with signs is to begin with something apparently definite and yet, as soon as the discussion begins, we find that there are only signs (see above) "A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign." Thus what is in the mind of another person is also a sign, albeit a sign that is understood internally by that person. One can look and look for substance that may underlie the sign but the search always leads to more signs. In this expansion of signs related to signs, signs describing signs, signs, and interpretant signs, the self becomes yet another sign standing in relation to all the signs that work at the nexus that the person represents. The sign of the self becomes a limit of all the signs that are the life of that self. The distinction of a person is his or her sign of distinction, his or her sign of self.

Spencer-Brown (1969) in his book "Laws of Form" (here to be abbreviated LOF), makes a new semiotic start, beginning with the idea of distinction. Signs arise as we shall see, but Spencer-Brown begins with the pronouncement:

> We take as given the idea of distinction and the idea of indication, and that we cannot make an indication without drawing a distinction. We take, therefore, the form of distinction for the form (LOF, page 1).

It is as this point that Peirce and Spencer-Brown come into contact. For in Chap. 12 in the last sentence of Laws of Form, Spencer-Brown writes "We see now that the first distinction, the mark, and the observer are not only interchangeable, but, in the form, identical." Here the mark is the first made sign or indication of a first distinction. The observer can be identified with the interpretant in so much as the interpretant (see the quote from Peirce above) is an equivalent sign created in the mind of somebody, and must for its existence partake of the being of that somebody. At this nexus Spencer-Brown indicates the essential identity of sign, representamen, and interpretant. The three coalesce into the form that is the form of distinction.

The form of distinction becomes, in Spencer-Brown, a background for the entire play of signs that is the context of Peircean semiotics. We take the form of distinction for the form. And in this saying "the form" becomes a noun as elusive as it seems to be concrete, just as is the nature of the sign in Peirce. The form of a distinction drawn as a circle in the plane is geometrical form, the circle. But the form of distinction, the form of the idea of distinction, what is this form?

The echo from Peirce is clear as a bell. The form of distinction calls up a sign in the mind of some person. It is an amalgamation or superposition of all the signs for distinction in the history of that mind, that observer, or all observers. We come forth in the complexity of experience to a sharp idea of the distinct. We can give instructions for the performance of an act of distinction while simultaneously understanding that it is a creative act, not bound by any given set of rules or regulations.

The first lines of Laws of Form are worth reading, but we shall not quote them here. After some thought the reader may come to realize that Spencer-Brown's first paragraph is an amalgam of words that all stand for aspects of distinction: *definition, continence, boundary, separate, sides, point, draw a distinction, spaces, states, contents, side of the boundary, being distinct, indicated, motive, differ in value*. The paragraph is not a definition in the mathematical sense of definition: something in terms of previously defined things. There is no possibility to define distinction in terms of previous things that are not distinctions. The only possibility is to define distinction in terms of itself. We take the form of distinction for the form.

That first paragraph is nevertheless readable. How did this happen? How could readability arise from circularity? The answer is in the injunctive power of language. This same paragraph contains the phrases: "arranging a boundary so that a point on one side cannot reach the other side without crossing the boundary," "a circle draws a distinction," "a distinction is drawn," "spaces, states, or contents … can be indicated," and "contents are seen to differ in value." At once the paragraph is an amalgam of synonyms for distinction and it is a catalog of injunctions to arrange a boundary, to draw a circle, to indicate, and to see the difference in value. We are invited to take these steps and so enter into a contract of exploring the concept and practice of distinction.

Let us not forget that we have followed already the injunction of the first line: "We take as given the idea of distinction and the idea of indication and that we cannot make an indication without drawing a distinction." It is already given that there is a something called indication that entails the making of a distinction. And implicitly it is called up that a distinction could occur without any indication. We cannot make an indication without drawing a distinction. Can we have a distinction without making an indication? We are falling down the rabbit hole.

But here, we have to look and see. In most circumstances, to draw is to indicate.

A notion of privacy is another form of distinction. Can I hide distinctions within the boundary of my privacy distinction? Then I can pretend that there are distinctions that do not have indications. What a tangled web we weave in order to believe. In order to make a distinction without an indication, we are entangled in a web of new distinctions. The very act of drawing is a form of indication, and it must be concealed? Must we search for distinctions that are made without drawing a distinction? I sit before an emptiness. The emptiness is distinct for me. It is empty and I am empty before it. It is possible to have less action not more. And in the limit of acting gently in emptiness, or not at all, there seems to be the possibility of distinction without indication.

In the next few lines of Laws of Form, one finds the quote "If a content is of value, then a name can be taken to indicate this value." Already we have faced the multiplicity of names for a distinction, and making an indication is a special act that cannot happen without the making of a distinction. Nevertheless, it comes as a shock that suddenly a *name* can be called forth. A name can be taken to indicate a value. A distinction can be performed that allows the performance of a distinction. We begin to realize that in this condensed place where there is only creation of distinction, boundary or the crossing of the boundary, the only distinction is at first the distinction between nothing (the unmarked) and the act of creation, and then arises a distinction between name and act. If a state is of value then a name can be taken to indicate this value. If a distinction is a distinction then a distinction of distinctioning can be distinctioned to distinction this distinction. We are down the rabbit hole again. One side makes you smaller. One side makes you larger. Choose a door and pass through it. The act and the name are not different. The indication of a distinction, the crossing of the boundary of the distinction, and the distinction "itself" are in the form identical.

We come to the creation of a name and find that this is the same as the creation of a distinction. They are one and the same. And yet a name can be separated from the distinction to which it refers. The name can be taken to be a new distinction that refers to the first distinction. Indeed we can imagine that the original distinction (for example a circle drawn in the plane) is seen (in all quietness) to stand for, to indicate, itself. But in the act of recognizing this possibility that "it" could stand for "itself" we have made a distinction between "it" and "itself." We have allowed a condensation by making the possibility of a separation. The name and the sign are born in that process. The name, the sign, is Peirce's representamen, a sign residing in the mind of somebody. And we conclude that the sign, the name, and the original distinction all reside in the mind of somebody. At the point of condensation, the mind is the sign and the sign is the mind. No mind, no distinction. No distinction, no mind. We take the form as the form of distinction. Form is emptiness. Emptiness is form.

In his "A Note on the Mathematical Approach" Spencer-Brown writes "The act is itself already remembered … as our first attempt to distinguish different things in a world where, in the first place, the boundaries can be drawn anywhere we please. At this stage the universe cannot be distinguished from how we act upon it, and the world may seem like shifting sand beneath our feet." The act of naming is the key step toward a world of apparent distinctions. It is by naming a distinction that we call it into being. In the first page of the first chapter of Laws of Form one finds the "Law of Calling: The value of a call made again is the value of the call." It is enough to indicate the name once. For any name, to recall is to call.

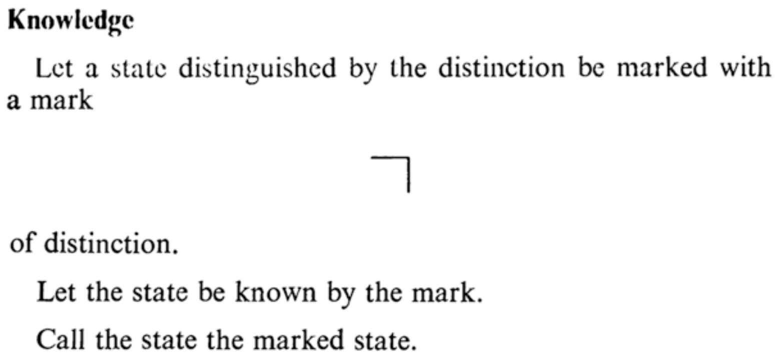
And then we find the "Law of Crossing: The value of a crossing made again is not the value of the crossing." At this point a distinction is made between crossing (the boundary of a distinction) and calling the name of a distinction. For "The value of a call made again is the value of the call." Crossing and calling appear to be given as terms in a similar level of speech, and yet they are declared to be different. We understand that the crossing of the boundary can be the act of naming the distinction. I cross into "riding" when I cross the boundary of balance and actually ride the bicycle. I name riding by actually engaging in the act of riding. If I cease to ride, then the value of riding ceases. The distinction of riding is no longer present. And yet it can still be named.

At this point, we have come to the end of Spencer-Brown's discussion of Laws of Form that makes no explicit use of a sign of distinction. The word "sign" has not yet occurred for Spencer-Brown. The first use of the word sign is in the next chapter of the book entitled "Forms Taken Out of the Form."

In this development, the injunctive mode has taken priority. The text tells its reader to "Draw a distinction." and to "Call it the first distinction." This should sweep away any notion that first distinction is an absolute concept.

The first distinction is the one that is under discussion. The form is the form of the first distinction. And so the form of the first chapter has shifted from the universal to the particular, and the form of distinction is the form of that first distinction. The form is inherent in any act of distinction. We find that at the point of intent "Let any mark, token, or sign be taken in any way with or with regard to the distinction as a signal. Call the use of any signal its intent." Here is the entry of the word sign into Spencer-Brown's consideration of distinction and form.

Now we listen again to Peirce. "A sign, or representamen, is something which stands to somebody for something in some respect or capacity." Indeed Spencer-Brown's mark, token, or sign is a sign in the Peircean sense. This sign is taken as a signal (in the condensation of Laws of Form, a signal is yet a sign) with regard to or



Knowledge

Let a state distinguished by the distinction be marked with a mark

⌉

of distinction.

Let the state be known by the mark.

Call the state the marked state.

Fig. 9.2 The Spencer-Brown mark

of the first distinction. Spencer-Brown does not say that this mark is the first sign, but sign it is and with it, it is possible to indicate the first distinction.

Finally there enters upon the stage of distinction the mark that will be the pivot for the formalism of Laws of Form. See Fig. 9.2.

Spencer-Brown writes, "Let a state distinguished by the distinction be marked with a mark of distinction." The mark is written upon one side of the first distinction. We shall take the liberty of illustrating this in Fig. 9.7. This mark is chosen to make and to indicate a distinction in its own form. The mark has (for the observer—our word for Peirce's somebody) an inside and an outside. Spencer-Brown says, quite explicitly, "Let each token of the mark be seen to cleave the space into which it is copied. That is, let each token be a distinction in its own form." And before this, he gives permissions:

"Call the space cloven by any distinction, together with the entire content of the space, the form of the distinction.

Call the form of the first distinction the form. … Let there be a form distinct from the form. Let the mark of distinction be copied out of the form into such another form.

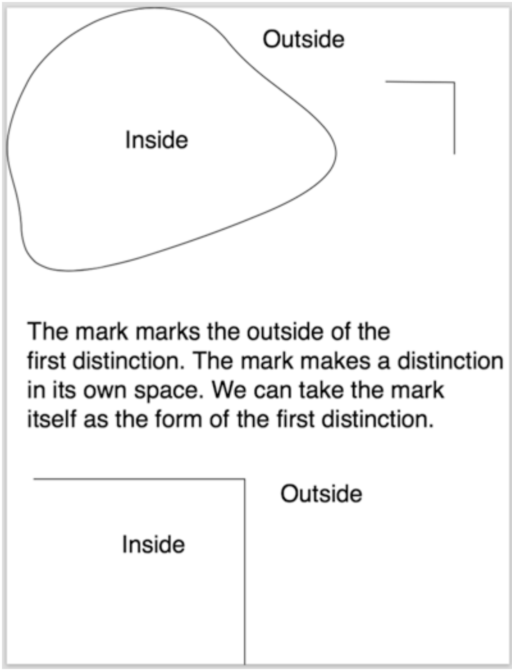
Call any such copy of the mark a token of the mark. Let any token of the mark be called as a name of the marked state. Let the name indicate the state."

Now the circle has again closed. Each token of the mark is a sign and a copy of the mark itself. Each token and indeed the mark itself is a distinction in its own form. There is now a plethora of signs, marks, and forms. They all indicate the marked side of the first distinction. Only one distinction is being discussed. As many marks as may be needed are available to signal this distinction. We embark upon not just form, but formalism and the inception of calculation.

Each mark in the expression on the left is a sign or name for the outside of the distinction made by the other mark in the expression. Each mark is the name of the other mark. The calling of a name made again may be identified with the calling of the name. And so we have the equation as indicated above, condensing the two marks to a single mark.

As we show in Fig. 9.3, the mark can indicate the outside of the first distinction, and we take the mark to make a distinction in the outer space of that first distinction.

Fig. 9.3 The mark indicates the outside of the first distinction



We see then that *we can take the mark itself as the first distinction.* This move brings the discussion directly in coincidence with C. S. Peirce's semiotics. If the mark or sign is the first distinction then it is a sign for itself. It is a sign that makes a distinction and it is a sign that stands for the outside space of that distinction. We are now in a position to summarize the semiotic development of the Spencer-Brown mark ⌉.

The mark is a sign that makes a distinction in the plane within which it is drawn. In that plane there is a distinction between the (bounded) inside and the (unbounded) outside of the mark. The mark is chosen to refer to the outside of the distinction that it makes in the plane. The mark can be seen to refer (via referring to the outside of the distinction that it makes) to itself as the boundary of that distinction. Thus we can write the law of calling in the form ⌉⌉ = ⌉.

Each mark in the expression on the left is a sign or name for the outside of the distinction made by the other mark in the expression. Each mark is the name of the other mark. The calling of a name made again may be identified with the calling of the name. And so we have the equation as indicated above, condensing the two marks to a single mark.

We take the mark to indicate a crossing from the state indicated on its inside:

$\overline{A}|$ denotes the state obtained by crossing from the state indicated by A.

Hence $\overline{\overline{\;\;}|}|$ indicates the state obtained by crossing from the marked state.

Hence $\overline{\overline{\;\;}|}|$ indicates the unmarked state.

In equations, we have the law of crossing: $\overline{\overline{\;\;}|}| =$

The value of a crossing made again is not the value of the crossing.

We have arrived at a self-referential nexus. The mark, first sign, refers to itself. The first sign is a name and it is identified with the action of crossing from the unmarked state (the state with no sign). We began with the idea of distinction.

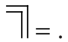
The sign of the first distinction acts as the transformation and the boundary between the unmarked state (the state with no sign) and the marked state. The sign of the first distinction is a signal of the emergence of articulated form. The sign of the first distinction is, in the form, identical with the first distinction.

Finding Primary Arithmetic

The formalism that we have arrived at is directly connected with mathematics. Let's recall where we are.

We have one sign ⌉ and two laws or rules about that sign:

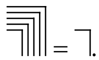
The law of crossing: ⌉⌉ = ⌉.

The law of calling: 

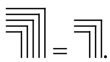
At this stage in the development of the sign, these laws are statements about naming and about the crossing of the boundary of an initial distinction. The initial distinction can be the distinction made by the sign itself. And yet there is another sign. It is the equals sign. And with that sign we enter mathematics. With the equals sign, we formalize condensations of reference and meaning.

It is implicit that we may write expressions such as 

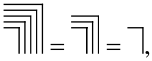
And we might wonder what this nest of signs can mean. And we find that we have already defined the meaning of this new sign as a fivefold act of crossing from the previous state, starting from the unmarked state. And being able to count, we know that this means that we will arrive at the marked state after such a process. Thus we have



An infinity of possible equalities of concatenations of signs has opened up before us and since we know how to count we can evaluate them all and find either the marked state or the unmarked state as an equivalent to each one. Do we need to know how to count to accomplish this task? We do not need to know how to count. We can apply the laws of calling and crossing where we find them. An empty cross with a cross over it can be regarded as an instance of the law of crossing:



The two innermost marks in the left-hand nest of marks are an instance of the law of crossing and we can erase them, forming the right-hand side with only three marks. Doing this once more, we find



and the marked value of the nest has been uncovered without the need for counting. Here is another way. Let *u* and *m* stand for the unmarked and marked states, respectively:

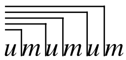
Agree that *u*⌉ has the marked state as its outside value and write

u⌉*m* to indicate this state of affairs.

Agree that *m*⌉ has the unmarked state as its outside value

and let *m*⌉*u* indicate this state of affairs.

Then we can evaluate the nest of marks by marking it with *u* and *m*.

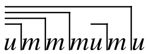


Similarly, ⌉⌉⌉⌉ = ⌉ by repeated application of the law of calling. And … = … = … = ….

Here we combine uses of the laws of calling and crossing when they are available. We see that there is an arithmetic of expressions written in the mark and the equals sign has taken on the crucial role of connecting expressions that indicate the same value.

Oh! You want to know the meaning of …!

It is a multiple action. Think of putting an unmarked signal u at the deepest spaces in the expression and marking it with *u* and *m* as we did before:



Note that in the space one crossing away from the outside there are two ms.

We take the rules that *mm* = *m*, *uu* = *u*, *mu* = *um* = *m*. Then any expression can be seen as indicating a multiple process of crossing and recrossing from the unmarked state of the first distinction. The signals interact with one another and produce the value of the expression as either marked or unmarked. The result is the same as that obtained by using the laws of calling and crossing on the expression. Here is the simplest arithmetic generated by a sign that makes a distinction. Spencer-Brown calls this the primary arithmetic or the calculus of indications.

Finding Logic

The primary arithmetic is a two-valued system. Every expression is either marked or unmarked. Remarkably, there is a translation to the two-valued logic of true (T) and false (F). Let *a* ∨ *b* denote "*a* or *b* "(inclusive or –*a* or *b* or both *a* and *b*), and

$a \wedge b$ denote "a and b." Let $\neg a$ denote "not a" and let $a \rightarrow b$ denote "a implies b." Recall that in symbolic two-valued logic one takes the equivalence $a \rightarrow b = (\neg a) \vee b$.

Now note that if we write algebraically about the primary arithmetic with the variables standing for either the marked or the unmarked states, *then ab is marked exactly when a is marked or b is marked.* This suggests that we take the interpretation T for the marked state and F for the unmarked state. Lets write

T = $\overline{}\rceil$ and F = .

Then $\overline{T}\rceil = F$ and $\overline{F}\rceil = T$. Thus we can interpret $\overline{a}\rceil$ as $\neg a$.

And then we have $a \rightarrow b = (\neg a) \vee b = \overline{a}\rceil b$

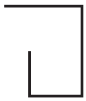
so that implication in logic becomes the operation $\overline{a}\rceil b$ in the algebra of the primary arithmetic. It is then easy to see that and is expressed by the formula $a \wedge b = \overline{\overline{a}\rceil \overline{b}\rceil}\rceil$ since the formula on the right is marked exactly when both a and b are marked.

In this way basic logic rests on the primary arithmetic and can be seen as a patterning of its operations and processes. I hope to have convinced the reader that this is a satisfactory entry into logic starting with the notions of sign and distinction. One can explore a great deal from this basis and I will stop here with only a hint of what may come.

One aspect of logic that comes forth at once is the role of paradox. Consider the liar paradox in the form $L = \neg L$. Rewriting into primary algebra, we find

$$L = \overline{L}\rceil.$$

Since the mark makes a distinction between its inside and its outside, this equation suggests that L must itself have a sign that indicates a form that reenters its own indicational space. L must have a sign as shown below:



In crossing from the state inside the reentering mark, we arrive again at the inside. The inside is the outside and the outside is the inside. The sign connotes a distinction that controverts itself and yet it is still a sign in the constellation of all signs and it still distinguishes itself in its own form. Nothing is left but the time of circulation in the oscillation of inside and outside, and beyond this state of time we have returned to void.

Finding Mathematics

Up to this point we have not actually ventured across a boundary into numerical mathematics. The construction of a sign that can stand for any sign and is self-referential involves no counting, no calculation, no algebra, and seemingly no arithmetic of any kind. It would appear that we have arrived at a pivot point where one could begin thinking about the growth of thought and language with no regard to the development of mathematics.

And yet mathematics has symbolic beginnings and is woven into the structure of language. What signs are the least signs needed for number? We might take on the sign | for 1, the sign || for 2, and generally the sign $n = |||\ldots|$ (with n vertical marks) for the integer n. In this mode we have $n + 1 = n|$. And $n + m = nm$, the juxtaposition of the marks for n and the marks for m:

$$1 = |,\ 2 = ||,\ 3 = |||,\ 4 = ||||,\ 5 = |||||\ \text{and so on.}$$
$$3 + 2 = ||| + || = ||||| = 5.$$

Arithmetic can grow from elemental signs and indeed we can use the Spencer-Brown mark to represent numbers with zero as the unmarked state:

$$0 =$$
$$1 = \overline{\;\;}|$$
$$2 = \overline{\;\;}|\overline{\;\;}|$$
$$3 = \overline{\;\;}|\overline{\;\;}|\overline{\;\;}|$$

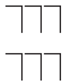
and so on.

Note that in order to represent numbers in this way, we must rescind the law of calling so that multiplicities of marks stand for different numbers. With the law of calling removed, we are no longer working with only one distinction. Each new number is a distinction in its own form. What about the law of crossing? It turns out that we can put it to service for defining multiplication as follows: We define $a \times b = \overline{a}|\#b$ where $\overline{a}|\#b$ means that we take each cross in b and insert a copy of $\overline{a}|$ underneath it. Then simplify the resulting expression using the law of crossing. Here is the example of 2 x 3:

$$2 \times 3 = \overline{2}|\#3 = \overline{\overline{\;\;}|\overline{\;\;}|}|\#\overline{\;\;}|\overline{\;\;}|\overline{\;\;}|$$
$$= \overline{\overline{\overline{\;\;}|\overline{\;\;}|}|}|\overline{\overline{\overline{\;\;}|\overline{\;\;}|}|}|\overline{\overline{\overline{\;\;}|\overline{\;\;}|}|}| = \overline{\;\;}|\overline{\;\;}|\overline{\;\;}|\overline{\;\;}|\overline{\;\;}|\overline{\;\;}|$$
$$= 6.$$

Here we have used the algebraic version of the law of crossing: $\overline{\overline{a|}|} = a$ for any a, and such an a can be one of our numerals, taking values beyond marked and unmarked.

This is the beginning of arithmetic, the gateway into the depths and beauties of mathematics. This foundation for the theory of numbers will clarify the deep quests of number theory. One can begin by wondering about the prime numbers. Six is not prime as we have just seen. It is a product of 2 and 3. The row of six marks is two rows of three marks and it is three rows of two marks. It seems that numbers want their own distinctions. After conversing with six we see that six prefers to be seen as



or as

but confinement to a single row is just not comfortable for a composite number. Let us find arithmetic anew by staying close to its origin in the origination of a sign.

Finding Ordinals

Remarkably, the structure of the Laws of Form expressions gives us a map of the transfinite ordinals. Let us explain. First recall that the transfinite ordinals of Cantor (1941) are an extension of the natural numbers. We begin with the natural numbers 1,2,3,… and then posit a new infinite number w that is greater than any natural number n. So now we have the ordered sequence 1,2,…; w. And we can continue with 1,23,…; w, $w + 1$,$w + 2$,…, $w + w = 2w$,…, $3w$,…, w^2,…, w^3,…, w^w, w^{w+1}… .

We shall translate these ordinals into Laws of Form expressions as shown below:

1 = 

2 = 

3 = 

…

To get higher we shall notate that generally, A + B = AB (juxtaposition) and $w^A =$ . Thus

$w =$ , $w^w =$ , $w^{w^w} =$ .

While $w^w + w + 1 =$  and $w^{w^w + w + 1} =$ .

It should be clear to the reader that the finite expressions in Laws of Form, taken only up to commutativity (AB = BA), each uniquely represents a treelike polynomial expression fragment of the transfinite ordinals!

Infinite expressions should be explored further.

For example, let $J =$ .

Then $J =$  $= J^w$ and $J = J^w$ are the important limit ordinals at the top of the hierarchy of the treelike transfinite ordinals that we have just indicated.

With this ordinal correspondence in place, we can translate a version of the Hercules and Hydra game of Kirby and Paris (1982) into Laws of Form expressions.

Take a finite expression such as E = .

Choose an empty mark in the expression.

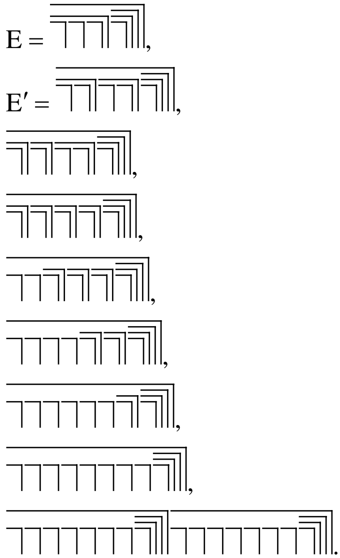
Determine the first mark that encloses this mark, making a sub-expression.

For example, E above has the sub-expression , with the leftmost mark the one we have chosen.

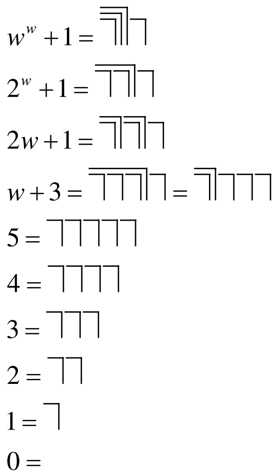
Now *remove the mark you have chosen and duplicate the resulting sub-expression to make a new expression E'*. Here the result is

E′ = .

(In the Kirby-Paris game one can make any finite number of duplicates. We shall restrict to one duplicate.) *The object of the game is to eventually reduce the expression to nothing.* Note that by the rules above, a single empty mark can be erased. Here are a few more moves in the game starting with E, above:



We are sure that the reader would like to finish this game! It will take quite a few moves, but not too many. Here is a shorter sequence. And in this shorter sequence we have labeled each expression with the corresponding ordinal. *The ordinals get smaller each time.* This always happens and it is the essence of the Kirby-Paris proof that *one can always win the Hydra game*. Any descending sequence of ordinals is finite, and so the game must end:



The most remarkable fact about the Hydra game is that, while we can prove that the game ends by using the transfinite ordinals, there is no proof of this fact in Peano arithmetic (Kirby and Paris 1982). The Laws of Form notation is a perfect expression of this fragment of the ordinals, and it is not just an entrance into Boolean algebra, ordinary arithmetic, and transfinite arithmetic. The Mark, via the Hydra Game, is an exemplar of Godelian incompleteness in what is surely its simplest form.

The Arctic Essay

This essay on the semiotics of Laws of Form was motivated by the author's discovery of the manuscript shown below. The manuscript was found at the bottom of an abandoned mine shaft in the frozen wastes of the Arctic Circle, abandoned surely in the 1800s. Written on crumbling paper and composed long before the conception of Spencer-Brown's book it is a mystery how the reference to the Spencer-Brown mark could have occurred in this manuscript. I have attempted in this essay to give sufficient background of a semiotic nature that the reader might be able to decipher the manuscript itself. The manuscript was entitled "A Sign in Space," but no author is indicated. I can only speculate that perhaps Spencer-Brown himself saw the manuscript, and yet that would not solve the puzzle of how it came to bear his name. Alas, the original document disintegrated into dust soon after it was found. This essay is all that is left. There is one clue. The document refers at a crucial point to C. S. Peirce. I suspect that this is a self-reference and that Peirce himself wrote the document. As for Spencer-Brown, Peirce time traveled into the future and took back these notes on Spencer-Brown's work. All that happened before the Russell singularity that made forward time travel impossible. There can be no other explanation.

A Sign in Space

Let $\rceil$ be the Spencer-Brown mark.

Let there be a distinction with inside denoted I, and outside denoted O.

Regard the mark as an operator that takes inside to outside and outside to inside.

Then

$$\overline{I}\rceil = O$$

$$\overline{O}\rceil = I.$$

Note that it follows that

$$\overline{\overline{I}|}| = \overline{O}| = I$$
$$\overline{\overline{O}|}| = \overline{I}| = O$$

For any state X we have $\overline{\overline{X}|}| = X$.
Introduce the unmarked state by letting the inside be unmarked.
Then I = .
And so

$$\overline{\ \ }| = O$$
$$\overline{O}| =$$

Therefore the value of the outside is identified with the mark and $\overline{\overline{\ \ }|}| =$.
The value of the outside is identified with the result of crossing from the unmarked inside:

$$\overline{\ \ }| = \overline{\ \ }|.$$

This equation can be read on the left as "cross from the inside "and it can be read on the right as "name of the outside." Once the inside is unmarked, then the mark itself can be seen to be the first distinction. The language of the mark is self-referential:

$\overline{\ \ }|\overline{\ \ }| = \overline{\ \ }|$ says that either mark names the other.

It is as though I were to wear a name tag that is a picture of myself. At the level of the form there is no difference between myself and a picture of myself. A sign can refer to another sign (cf. C. Peirce). The mark is seen as a sign and as a distinction between the inside of that sign and its outside. We take the mark, as sign, to refer to its own outside. In the form, the mark and the observer are identical. In the form, a thing is identical to what it is not.
Tat Vam Asi.

In this way one arrives at non-duality by abandoning form to void:

Abandon form to void.
Form is emptiness.
Emptiness is form.
The form we take to exist arises from framing nothing.
We take the form of distinction for the form.

Epilogue

In this chapter we have examined the use of and development of signs in relation to G. Spencer-Brown's Laws of Form and we have engaged in wordplay related to a real story "A Sign in Space" by Italo Calvino (giving an extensive quote from it in our introduction) and a fictitious document named "A Sign in Space" that seems to be a precursor to the work of both Charles Sanders Peirce and George Spencer-Brown. In fact, there is a long history of precursors to the semiotic signs at the base of mathematics, logic, and language. Alphabets are historical records and ongoing libraries of signs and the simplest of such forms such as the cuneiform and Sumerian signs. For example, consider the fragment in Fig. 9.3 from a Sumerian document, twenty-sixth century BC.

There in the document are a nest of left-shaped marks, and since they are nested the distinction they each make in the plane was clearly part of their use. In modern typography a relative of the Spencer-Brown mark is the square root sign, a connected sign that can be nested and arranged for mathematical purposes (Fig. 9.4):

$$\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{2}}}}}}$$

The language of Laws of Form was discovered, according to Spencer-Brown, in making a descent from Boolean algebra in which he found the notation of the mark, the role of the unmarked state, and the double-carry of mark as name and mark as transformation. In Boolean algebra and in symbolic logic the negation sign connotes transformation and it does not stand for a value (true or false). In the calculus of indications, viewed from the stance of symbolic logic, the mark is a coalescence of the value true and the sign of negation. This comes about because true is what is not false and the false is unmarked in Laws of Form. Hence $T = {\sim}F = {\sim}$ (using ~ as the sign for negation), but this cannot be said without confusion in symbolic logic since there is no inside to the sign of negation. In Laws of Form we can write $T = \overline{F}\rceil = \overline{\ \ }\rceil$ and the mark as container (as parenthesis) makes it possible for it to take on its double role of value and operator.

Wittgenstein (1922) says in the *Tractatus* (4.0621): "the sign '~' corresponds to nothing in reality." And he goes on to say (5.511): "How can all-embracing logic which mirrors the world use such special catches and manipulations? Only because all these are connected into an infinitely fine network, the great mirror." For Wittgenstein in the *Tractatus*, the negation sign is part of the mirror making it

Fig. 9.4 Sumerian document, twenty-sixth century BC.

possible for thought to reflect reality through combinations of signs. These remarks of Wittgenstein are part of his early picture theory of the relationship of formalism and the world. In our view, the world and the formalism we use to represent the world are not separate. The observer and the mark are (in the form) identical.

This theme of formalism and the world is given a curious twist by an observation that the mark and its laws of calling and crossing can be regarded as the pattern of interactions of the most elementary of possible quantum particles, the Majorana Fermion (Kauffman 2009, 2010, 2016). A Majorana Fermion is a hypothetical particle that is its own antiparticle. It can interact with itself to either produce itself or annihilate itself. In the mark we have these two modes of interaction as

$$\text{calling } \urcorner\urcorner = \urcorner \text{ and crossing } \overline{\urcorner}\!| = \;.$$

The curious nature of quantum mechanics is seen not in such simple interactions but in the logic of superposition (a kind of exclusive or) and measurement. Measurement of a quantum state demands the coming into actuality of exactly one of a myriad of possibilities. Thus we may write

$$\urcorner * \urcorner = \urcorner\urcorner + \overline{\urcorner}\!|$$

to indicate that the quantum state $\urcorner * \urcorner$ of a self-interacting Majorana Fermion $\urcorner$ is a superposition of marked and unmarked states.

Upon observation, one or the other (marked or unmarked) will be actual, but before observation the state is neither marked nor is it unmarked. We need a deeper step in semiotics to enter into quantum sensibility. The equation for this interaction can be written in ordinary algebra as $PP = P + 1$ where P stands for the Majorana particle and 1 stands for the neutral state of pure radiation. Then we recognize a famous quadratic equation $P^2 = P + 1$ with solution of the golden ratio $\left(1+\sqrt{5}\right)/2$ and multifold relationships with the Fibonacci numbers. Indeed this is the legacy of the Majorana Fermion as Fibonacci particle, fundamental entity in the most idealistic and yet soon to be practical searches for quantum computing and understanding of particles as well known as the electron. Each electron appears to be an amalgam of two Majorana Fermions. It is not the point here to start doing technical physics, but the moving boundary of sign and space is changed from the time of Wittgenstein and we should expect to see semiotic insight of a different kind from now on.

Charles Sanders Peirce came very close to inventing the mark $\urcorner$ in his "sign of illation" as shown in Fig. 9.5. [C. S. Peirce, "The New Elements of Mathematics," edited by Carolyn Eisele, Volume IV—*Mathematical Philosophy*, Chapter VI—The

Fig. 9.5 Peirce's sign of illation

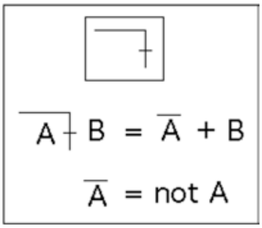


Fig. 9.6 The Peirce sign of illation

Logical Algebra of Boole, pp. 106–115. Mouton Publishers, The Hague—Paris and Humanities Press, Atlantic Highlands, N. J. (1976) (Peirce (1976).]

The Peirce sign of illation is used for logical implication and it is an amalgam of negation as the over-bar and logical or writing as a + sign on the left vertical part. See Fig. 9.6 for an illustration of this anatomy of the Peirce sign.

The mark ⌉ goes further since the unmarked state is allowed, and the operation of or is also unmarked and indicated by juxtaposition. Thus we still have the decomposition of $A\rceil B$ as "Not(A) or B" once the mark is understood to operate as negation. The largest difference is semiotic in that the mark can be taken as a universal sign and as a sign for itself. As such it has a conversational domain quite independent of Boolean logic. In this role, the mark can be seen as part of a wider context of distinction that informs and illuminates logic and mathematics.

Peirce spoke of a "Sign of Itself." Here is a key passage from his work.

> But in order that anything should be a Sign it must 'represent', as we say, something else called its Object, although the condition that a Sign must be other than its Object is perhaps arbitrary, since, if we insist upon it we must at least make an exception in the case of a Sign that is part of a Sign. Thus nothing prevents an actor who acts a character in a an historical drama from carrying as a theatrical 'property' the very relic that article is supposed merely to represent, such as the crucifix that Bulwer's Richelieu holds up with such an effort in his defiance. On a map of an island laid down upon the soil of that island there must, under all ordinary circumstances, be some position, some point, marked or not, that represents qua place on the map the very same point qua place on the island …
>
> If a Sign is other than its Object, there must exist, either in thought or in expression, some explanation or argument or other context, showing how – upon what system or for what reason the Sign represents the Object or set of Objects that it does. Now the Sign and the explanation make up another Sign, and since the explanation will be a Sign, it will probably require an additional explanation, which taken together with the already enlarged Sign will make up a still larger Sign; and proceeding in the same way we shall, or should ultimately reach a Sign of itself, containing its own explanation and those of all its significant parts; and according to this explanation each such part has some other part as its Object. [C. S. Peirce, *Collected Papers* – II, p. 2.230 – 2.231, edited by Charles Hartshorne and Paul Weiss, Harvard University Press, Cambridge (1933).]

There are extraordinary topological ideas in this passage. There is an implicit reference to the notion of a fixed point so that a map and its image must have a coincidence. There is the notion that sign and explanation will undergo recursion until ultimately the sign, the explanation, and the object become one. We have begun with a sign ⌉ that is a sign for itself in the sense that it represents the distinction that is made by the sign in its coincidence with an observer. And yet the

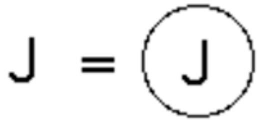


Fig. 9.7 Reentrant equation

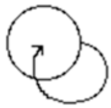


Fig. 9.8 A reentrant form

recursion is always possible. Consider the reentrant sign that was discussed in section "Finding Logic". The reentrant sign can be taken to be a solution to $J = \overline{J|}$ or, in a graphical mode, to be a solution to the re-embedding of *J* inside a circle as in Fig. 9.7.

"And yet, the equation $J = \overline{J|}$ asserts the reentry of *J* into its own indicational space, and it exhibits *J* as a 'part of itself.' The equation is the explanation of the nature of *J* as reentrant and can be taken as a description of the recursive process that generates an infinite nest of circles. It is only *J as an equation* that yields *J* as a Sign of itself. If we wish to embody the equation in the Sign itself then we need to allow the Sign to indicate its own reentry as we did in the last section with the symbol shown in Fig. 9.7. This symbol does 'contain its own explanation' in the sense that we interpret the arrow as an instruction to reenter the form inside the circle" (Kauffman 2001: 79–10) (Fig. 9.8).

In fact that reentry occurs ad infinitum as indicated in Fig. 9.9, from which we see that the equational reentry is recaptured from the self-standing form of Fig. 9.10.

We see that it is a matter of language that fuels the difference between a simple form that stands for itself such as the mark and those reentry forms that partake of infinite regress (as shown in Fig. 9.9) in order to attain self-reference. This infinite regress is a microcosm of the infinite regress of Peirce and allows us to solve for J as an unending nest of marks:

$$J = \overline{\overline{\overline{\overline{\overline{\ldots|}|}|}|}|} = \overline{J|}.$$

It must be mentioned that the work of Church and Curry on the Lambda Calculus (see Kauffman 1985, 1987, 1994, 2001, 2005, 2009, 2012; Buliga and Kauffman 2009) gives another approach to reentry:

Let $GX = \overline{XX|}$. Then $GG = \overline{GG|}$ and so we can take *J* = *GG* to obtain $J = \overline{J|}$ without any infinite regress!

How did this happen? At the level of sign and operation, *G* is a duplicating device. Given an *X*, *G* makes two copies of *X* and places them under a mark. The equals sign means that *GX is replaced by* $\overline{XX|}$, and the operation of *G* is defined by this replacement. When we replace *X* by *G* in the equation, we have put *G* in the

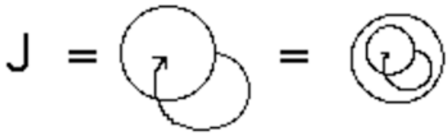


Fig. 9.9 Equation, indication, and reentry

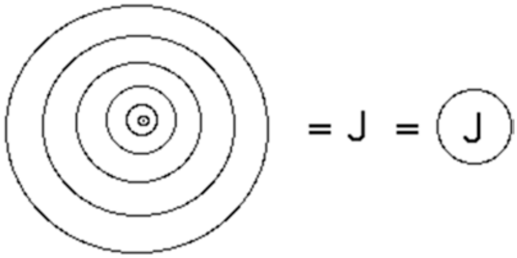


Fig. 9.10 Infinite regress and fixed point or eigenform

position to act on *G*. *G* does act and produces *GG* with a mark around it. But *GG* is now ready to act again and so *GG* moves into the temporal domain and instructs a recursion:

$$GG = \overline{GG} = \overline{\overline{GG}} = \overline{\overline{\overline{GG}}} = \overline{\overline{\overline{\overline{GG}}}} = \ldots$$

The infinite regress of *J* has been replaced by the inherent temporality of *GG*. The Church-Curry idea of recursion is in fact an outgrowth of the Russell paradox of the set of all sets that are not members of themselves. To see how this plays out in the realm of signs *let XY denote that Y is a member of X*. Taking this to heart, we define the Russell set by the equation $RX = {\sim}XX$.

As the reader sees immediately, *R* is now the duplicating Gremlin. We have shifted the interpretation of the mark to negation and we use ordered juxtaposition as membership. We find that if $RX = {\sim}XX$, then $RR = {\sim}RR$ and we now have the self-denial of the Russell set in regard to its self-membership. This could just as well have been written $RX = \overline{XX}$ and $RR = \overline{RR}$. We understand that this need not be a paradox. It is a reentry form and can be taken on its own cybernetic grounds. We have the option to view the Russell set temporally in the Church-Curry recursion. Then Russell oscillates in time between being a member of itself and not being a member of itself. The Russell pendulum avoids the Russell singularity.

In our fiction in this chapter, we referred to the Russell singularity as having made time travel into the future for the sake of mathematical and semiotic plagiarism impossible. In fact this was a reference to the weight of the ban (The Theory of Types) on temporal solutions to the paradox that was presented by Russell and Whitehead in their monumental work *Principia Mathematica*. With the recursive way out it may be that we have also released the demons of time travel once again upon an unsuspecting world.

When representation and explanation are insisted upon, then an infinite regress occurs due to the proliferation of signs that must indicate each stage of explanation. When this “noise” is reduced by the indicational power of an arrow, or the simple recognition of the presence of a distinction, then forms can stand alone and be recognized as being, in form, identical with their creators.

Along with the references quoted directly in the text, I have provided a selection of papers that I have written that are related to the themes of this essay. There is much to think about in this domain and we have only just begun.

Acknowledgement Kauffman's work in this chapter was supported by the Laboratory of Topology and Dynamics, Novosibirsk State University (contract no. 14.Y26.31.0025 with the Ministry of Education and Science of the Russian Federation). This chapter is dedicated to the memory of David Solzman, who introduced the author to many signs, including the work of Italo Calvino.

References

Buliga, M. and Kauffman, L. H. (2014). Chemlambda, universality and self-multiplication, In: H. Sayama, J. Rieffel, S. Risi, R. Doursat. and H. Lipson (eds.), Artificial life 14: Proceedings of the Fourteenth International Conference on the synthesis and simulation of living systems. Cambridge: MIT Press.

Calvino, Italo (1965). *Cosmicomix*. Torino: Einaudi.

Cantor, G. (1941). *Contributions to the founding of the theory of transfinite numbers.* New York: Open Court Publishing.

Kauffman, L. H. (1985) Sign and space. In: *Religious experience and scientific paradigms. Proceedings of the 1982 IASWR Conference, Stony Brook* (New York), Institute of Advanced Study of World Religions, pp. 118–164.

Kauffman, L. H. (1987). Self-reference and recursive forms. *Journal of Social and Biological Structures*: 53–72.

Kauffman, L. H. (ed.) (1994). *Knots and Applications.* New York: World Scientific Publisher.

Kauffman, L. H. (2001). The mathematics of Charles Sanders Peirce. *Cybernetics and Human Knowing* 8 (1–2): 79–110.

Kauffman, L. H. (2005). Eigenform. *Kybernetes* 34: 129–150.

Kauffman, L. H. (2009). Reflexivity and eigenform: The shape of process. *Kybernetes* 4(3).

Kauffman, L. H. (2010.) Reflexivity and foundations of physics. In: R. Amaroso, P. Rowlands, and S. Jeaers (eds.), *Search for fundamental theory" The VIIth International Symposium honoring French mathematical physicist Jean-Pierre Vigier, Imperial College, London, UK, 12–14 July 2010*, pp. 48–89. American Institute of Physics.

Kauffman, L. H. (2012). *Knots and Physics*, 4th ed. New York: World Scientific Publishers.

Kauffman, L. H. (2016). Knot logic and topological quantum computing with Majorana fermions. In: J. Chubb, J. Chubb, A. Eskandarian, and V. Harizanov (eds.), *Logic and algebraic structures in quantum computing and information*. Cambridge: Cambridge University Press.

Kirby, L. and Paris, J. (1982). Accessible independence results for Peano arithmetic. *Bulletin of the London Mathematical Society* 14: 285–293.

Peirce, C. S. (1931–1966). *Collected papers of Charles Sanders Peirce*, edited by Charles Hartshorne, Paul Weiss and Arthur W. Burks, 8 vol. Cambridge (Massachusetts): Belknap.

Peirce, C. S. (1976). *The new elements of mathematics*, edited by Carolyn Eisele, Volume IV – *Mathematical philosophy*, chapter VI – *The logical algebra of Boole*, pp. 106–115. Mouton Publishers, The Hague – Paris and Humanities Press, Atlantic Highlands, N. J. (1976).

Spencer-Brown, G. (1969). *Laws of form*. London: George Allen and Unwin.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. New York: Harcourt, Brace and Company.

Chapter 10
The Topology of Mathematics in the Mind and Its Interaction with Verbal and Written Language

Robert K. Logan and Izabella Pruska-Oldenhoff

Introduction

When we think of the human mind we most often think of its capacity for verbal language as we are the only living organism capable of speech. We are aware of the fact that the human mind is capable of mathematical thinking and think that mathematics was a later development of the human mind long after humankind had acquired language. In a book soon to be released in the Springer series *Mathematics in the Mind* edited by Marcel Danesi entitled *A Topology of Mind—Spiral Thought Patterns, the Hyperlinking of Text, Ideas and More,* we (Logan and Pruska-Oldenhof 2019) argue that human verbal language was as much a product of mathematical thinking as mathematics was a product of verbal thinking. We argue that the origin of verbal language, the origin of the mind, and the origin of mathematic thinking all happened at approximately the same time and that these three elements are basically interlinked. The human mind is a product of the brain and verbal language as was argued in *The Extended Mind: The Emergence of Language, the Human Mind and Culture* (Logan 2007), but verbal language as we have argued was dependent on the ability of humans to think in terms of sets employing a primitive form of set theory.

Before humans had verbal language, they lived in a world of percepts. Their communication was mimetic consisting of hand signals, facial gestures, body language, and nonverbal prosody or tones such as grunts and whines. They could only communicate about the here and now. Conceptual thinking only became possible with verbal language and our first concepts were our first words. These words acting as concepts linked to and represented all the percepts associated with those words. For example, the word water represents the concept of water and instantaneously triggers all of the mind's direct experiences and perceptions of water such as the water we drink, the water we cook with, the water we wash with,

R. K. Logan (✉) · I. Pruska-Oldenhoff
Ryerson University, Toronto, ON, Canada
e-mail: logan@physics.utoronto.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_10

the water that falls as rain or melts from snow, and the water that is found in rivers, ponds, lakes, and oceans.

The word “water” acting as a concept and an attractor not only brings to mind all “water” transactions but it also provides a name or a handle for the concept of water, which makes it easier to access memories of water and share them with others or make plans about the use of water. Words representing concepts speed up reaction time and, hence, confer a selection advantage for their users. And at the same time those languages and those words within a language, which most easily capture memories, enjoy a selection advantage over alternative languages and words, respectively.

The skill that made language possible and allowed a word acting as a concept to represent all of the percepts associated with that word was the mathematical ability to create sets, the set of all the percepts associated with that word. We suggest that the brain before verbal language was merely a percept processor and that afterwards it was able to conceptualize, i.e., operate with concepts. Each concept linked all the percepts associated with that concept. We conclude that the human mind naturally makes associations, creates categories or sets, and hence has the natural mathematical structure of set theory. We further suggest that verbal language emerges as a primitive form of set theory in that a set of percepts that are associated with each other or are similar are linked together with a word acting as a concept that unites all the members of that set. In a certain sense the primitive form of set theory we just described seems to be a precondition for the emergence of verbal language. It is not possible to determine the causal linkage between the primitive form of set theory and verbal language. We posit as is the case with a complex system where one cannot separate the top-down from the bottom-up causality that set theory and verbal language co-emerged. It is not that set theory caused verbal language to emerge or that language allowed set theory to emerge. Rather we would claim that mathematical set theory and verbal language self-organized into an emergent supervenient system.

The emergence of set theory according to this model of co-emergence would have preceded the emergence of enumeration as enumeration requires verbal language. There are two types of numbers, concrete numbers and abstract numbers. A pair of shoes, a yoke of oxen, or a brace of partridges are concrete numbers where the number is tied to the objects being enumerated. Concrete numbers have meaning only as units of the commodity they are designating and enumerating. The number “two” is abstract as it can apply to any set of two objects. We surmise that it arose in association with the perception of two people or two deer or two eggs or, even more relevantly, two fingers. Concrete numbers, such as “a brace of partridges” or “a yoke of oxen,” cannot be used to designate “two” as an abstract number and then be used to enumerate other objects. A brace of sandals is meaningless; instead one must refer to them as a “pair of sandals,” that is, as a concrete number or as “two sandals” where “two” operates as an abstract number. We would surmise that at some point in the evolution of language one particular concrete number came to represent an abstract number. We can only guess as to how this happened but certainly it is the case that numbers in the form of numerals like one, two, and three are basically concepts represented by words. It is no accident that the universal number system of most cultures is 10, the number of our fingers, or in some case 20 where counting included both the fingers and the toes. This explains why the term digit has two meanings, one meaning is a finger and the other meaning is a numeral.

The model that we have proposed of how verbal language and mathematical thinking co-emerged is an abduction or a just so story. It is a hypothesis but it cannot rise to the level of a scientific hypothesis because it cannot be falsified as the emergence of verbal language happened long before any scientific observations could be made. In fact, there could be no science before the emergence of verbal language as science requires conceptualization, which in turn requires verbal language.

Mathematics in the Mind Leads to Writing in Sumer and Writing Leads to the Further Development of Mathematical Thinking

We will argue that not only did mathematical thinking lead to verbal language but it also gave rise to written language through the development of mathematical notation. The very first notation for recording quantities was tally sticks in which the number of notches in the stick or antler corresponded to some quantity that the maker of the tally stick wanted to keep track of. The tally stick gave no indication of what was being tallied. The next step in the evolution of numerical notation was clay accounting tokens that archeologist Denise Schmandt-Besserat discovered in her digs in the Middle East especially in the Fertile Crescent between the Tigris and Euphrates rivers. These tokens had different shapes that corresponded to the things that they were enumerating which were agricultural commodities. The tokens were used as receipts for tributes paid by farmers to the priest-accountants as a form of taxation.

The agricultural commodities that the priest-accountants collected were redistributed to the workers that built and maintained the irrigation systems that made agriculture possible. The system of accounting tokens dates back to 10,000 years ago. They are similar to tally sticks except that the clay tokens have different shapes as each unique shape represented a different agricultural commodity. The token system remained basically unchanged for the first 5000 years of its use. Around 3200 BC the tokens were placed in spherical clay envelopes so the tokens would not become scattered and lost. After about a century of this, a bright priest-accountant suggested that they stamp the clay envelope while it was still wet with the tokens to be put inside so that they would not have to break open the envelope to see what tokens were inside the envelope. After a century or so of this practice another bright priest-accountant said why bother putting the tokens inside the envelope once the envelope was stamped and voila the clay tablet was born. The next innovation came about as the commerce in Sumer expanded and large numbers of agricultural commodities were being transacted. It became a nuisance to press the same token into a tablet multiple times. The solution to this problem was that the token for the large and small measure of wheat, the ban and the bariga, came to represent the numbers 10 and 1, respectively. To distinguish the ban and bariga used as the numbers 10 and 1 instead of the large and small measure of wheat it was decided that the numbers would be designated by pushing the token into the clay tablet and that the agricultural commodities including the large and small measure of wheat would be represented by etching the shape with a stylus on the clay tablet that the token representing that commodity would make if it were pushed into the wet clay tablet. As a result, a bifurcation occurred in which signs representing numbers were distinguished from signs representing words. And this is how the idea of writing was born as the result of the mathematical thinking of the Sumerian priest-accountants.

The idea of writing spread from Sumer throughout the Eastern Hemisphere. It is possible that the Chinese writing system was inspired by Western writing systems as trade existed between China and the Middle East before the appearance of writing during the late Shang dynasty circa 1200–1050 BC. The other independent invention of writing took place in the Western Hemisphere in Mesoamerica beginning with the Zapotec writing system that has not yet been fully deciphered. We therefore cannot find a link between math and writing for the Mesoamerican writing systems as we do not know how that system emerged.

The Mesoamerican number system of a bar and a dot however is similar to the Sumerian ban and bariga with both the dot and the bariga representing the number one. The one difference is that the bar represented 20 versus the ban which denoted 10. One of the chief uses of writing was to keep track of the Mesoamerican calendar providing a possible hint of a connection between math and writing, but this is hardly convincing evidence.

With a written notation for both words and mathematical notation not only was communication enhanced but also mathematical thinking became more sophisticated. De Cruz and De Smedt (2013) argue that

> 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. [Signalling] an intimate relationship between mathematical symbols and mathematical cognition.

Thus, mathematical thinking gave rise to mathematical notation and writing which in turn led to the further development of mathematical thinking. Here we encounter a spiral structure to human cognition which will be the subject of the next section of this chapter.

Cyclic and the Spiral Structure of Human Cognition

Mathematical structures of set theory and verbal language co-emerged. Verbal language led to enumeration and enumeration in the context of accounting led to mathematical notation and writing. This cycle of math in terms of a primitive form of set theory led to verbal language that led to number words and enumeration, which led to accounting tokens and then the notation of numerals and writing. This in turn led to an increase in the cognitive ability of mathematicians and to the further development of mathematics. We suggest that this development represents a spiral structure. A spiral is generated by circular motion in a two-dimensional conceptual plane to which is added an increase or a decrease in the third direction perpendicular to the plane of the circular motion. In our case the circular motion is the cycle from mathematical structures to notation and back again and the third dimension is the increase in the mathematical cognitive development of the human mind. The spiral structure of the topology of the human mind appears many times in the evolution of human thought, as we will show, replicating the spiral structures found in nature.

Spiral forms abound in nature in both the abiotic physical world and in the biosphere. Examples of abiotic spiral structures range from the spiral structures of galaxies to that of hurricanes, and tornados to whirlpools in oceans and rivers and even to the spirals in our bathtubs and sinks as water goes down the drain. The spiral structure in the biosphere occurs in the helical structures of DNA, RNA, and proteins and in the many structures of both plants and animals. We begin the gallery of spirals in various biological specimens by reviewing three examples: the floret of the sunflower, which has the pattern of the Fermat's spiral that allows the maximum number of seeds that can be packed into the flower; the logarithmic spiral of the nautilus mollusk shell; the fractal spiral structure of Romanesco broccoli; and the tails of the veiled chameleon and the seahorse (Fig. 10.1).

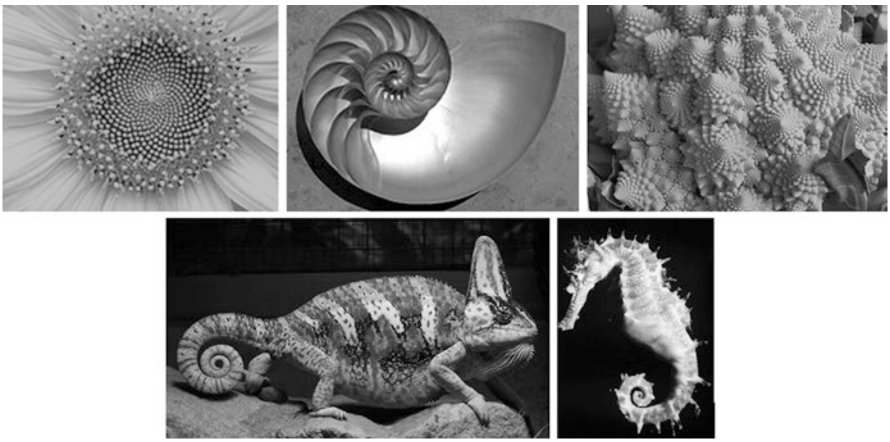


Fig. 10.1 Spiral structures

The Cyclic/Spiral Structure of Human Thought Patterns

The cyclic/spiral structure of human thought patterns is a universal characteristic of many cultures throughout human history. The idea of eternal return is the notion that events in the universe and in particular the events experienced by human kind repeat themselves so that they occur and reoccur forever and ever. This notion dates back to ancient Indian and ancient Egyptian philosophy. It was also an idea at the core of the beliefs of the Pythagoreans and the Stoics.

The notion in Indian philosophy and religion of the cyclic nature of time and existence can be dated as far back as 3300 BC and is characteristic of today's religion in India including Hinduism, Jainism, Buddhism, and Sikhism. The ancient Egyptians also believed in the notion of the eternal recurrence with their notion of *neheh*, roughly translated as endless recurrence.

Chinese historians as far back as 2070 BC formulated the notion of dynastic cycles in which a new dynasty emerges replacing an older one. At first the new dynasty is vital and dynamic but over time it degenerates until it is replaced by a new vital dynasty and the eternal cycle resumes once again.

There are two Greek myths that incorporate the notion of a recurring cycle, namely the myth of Sisyphus and the myth of Prometheus. According to legend Sisyphus was able to put death in chains so that mortals would not have to die. Death was able to escape, however, and together with the gods condemned Sisyphus to the endless cycle of pushing a heavy rock up a mountain only to have it roll back down the mountain and have Sisyphus return to the bottom of the mountain and once again push the rock to the top of the mountain and so on and so forth for eternity. The myth of Prometheus whose name literally means forethought is another story of the punishment of a mortal. Prometheus is credited with making mortals out of clay and then defying the gods he stole fire from them and gave it to humankind. Zeus was enraged by this act of theft and condemned Prometheus to be tied to a rock and have an eagle, a symbol of Zeus, come every morning and feed on his liver. Prometheus' liver would grow back overnight and the eagle would come again and feast on Prometheus' liver.

Kyklos is a traditional cyclic model of how political regimes evolve that was also described in Plato's *Republic*. The idea was later enlarged upon and elaborated by the Greek historian Polybius (200–118 BC). The cycle begins with society in anarchy or a lack of government from which a strong figure emerges as a monarch. At first this form of government works well but as a result of inheritance of monarchs lacking moral stature the monarchy degenerates into a form of tyranny. The tyrannical regime is overthrown by the prominent citizens or aristocrats of the state to form an oligarchy that rules effectively until it degenerates through the corruption of the inheritors of the oligarchy. This oligarchy is overthrown by the ordinary citizens of the state to form a democracy which in turn degenerates into rule by the mob or anarchy and the cycle begins all over again.

Another form of cyclic thinking was developed by Pythagoras and his followers. They believed that everything progressed in predictable cycles, which might have been motivated by their knowledge of the cyclic movement of the heavenly bodies including the annual movement of the sun, the monthly cyclic movement of the moon as well as the recurring cycles of the planets.

Porphyry, a philosopher from Tyre in the Roman Empire wrote in *The Life of Pythagoras* that:

> The following became universally known: first, that he [Pythagoras] maintains that the soul is immortal; second, that it changes into other kinds of living things; third, that events recur in certain cycles and that nothing is ever absolutely new; and fourth, that all living things should to be regarded as akin. Pythagoras seems to have been the first to bring these beliefs into Greece (the bolding is ours).

Muḥammad ibn Khaldūn al-Ḥaḍramī was a Tunisian/Arab historian who developed a cyclic theory of the rise and fall of empires in which an empire is formed, prospers, and after a period of decline is conquered by another regime which creates its empire but incorporates some of the cultural elements of the conquered empire. The conquering empire then suffers the same fate as the empire it conquered and so on and so forth in a never-ending cycle.

The tradition of cyclic accounts of history continued into modern times. Giambattista Vico's philosophy of history, circa 1725, involved the notion of an advancement, a *corso*, followed by a *ricorso* or return. According to Vico, a society evolves to a high point in its development, its *corso*, and then regresses back to or returns to a more primitive time in its history, which he refers to as a *ricorso*. After the *ricorso* the society once again embarks on a new *corso* and progresses to a more advanced level of development to once again experience a *ricorso* or return. The spiral structure of history of *corso* followed by *ricorso* followed by another cycle of *corso* and *ricorso* and so on and so forth is characteristic of Vico philosophy of history. For Vico history has a spiral structure of recurring cycles of development and collapse. The cycle consists of three ages. The first and most primitive is the age of gods. The second age is the age of heroes, in which there is constant conflict between the rulers and the governed. The third age is the age of the people in which democracy emerges but which eventually collapses because of corruption and returns to the age of gods once again. There is some parallel with the Greek notion of Kyklos if we take the age of gods as monarchy given some monarchs or emperors claimed divinity or claimed to have been chosen by God. The age of heroes would correspond to the regime of the aristocrats.

Vico's work was largely ignored by his contemporaries. His influence was not felt until the nineteenth century when his ideas influenced Marx, Goethe, Humboldt, Dilthey, Nietzsche, and Gedamer and into the twentieth century when his influence was felt by James Joyce, Marshall McLuhan, and Mircea Eliade. The notion of the "eternal return" was a central part of Friedrich Nietzsche philosophy. He first formulated this idea in Aphorism 341, entitled "The greatest weight," in Book IV at the very end of *Die Fröhliche Wissenschaft* (*The Gay Science)* where he wrote:

> What, if some day or night a demon were to steal after you into your loneliest loneliness and say to you: "This life as you now live it and have lived it, you will have to live once more and innumerable times more; and there will be nothing new in it, but every pain and every joy and every thought and sigh and everything unutterably small or great in your life will

> have to return to you, all in the same succession and sequence—even this spider and this moonlight between the trees, and even this moment and I myself. The eternal hourglass of existence is turned upside down again and again, and you with it, speck of dust!" Would you not throw yourself down and gnash your teeth and curse the demon who spoke thus.

The idea of eternal return was not for Nietzsche just a poetic or philosophical idea but something he believed actually and literally happens, is happening, and has happened. He even argued for the idea making use of physics and probability arguing that if there is a finite amount of matter in the universe and there is an infinite amount of time then every configuration of matter in the universe would have to repeat itself. He even considered studying physics to confirm his hunch of the existence of the eternal return.

Nietzsche returned to the "eternal return" in *Thus Spake Zarathustra*, the next book he wrote after the Gay Science. In Chapter LVII The Convalescent he explicitly describes his notion of the "eternal return":

> O Zarathustra, who you are and must become: behold, you are the teacher of the eternal return, that is now your fate! That you must be the first to teach this teaching - how could this great fate not be your greatest danger and infirmity! Behold, we know what you teach: that all things eternally return, and ourselves with them, and that we have already existed times without number, and all things with us.

The notion of the "eternal return" reappears in the twentieth century with the work of Mircea Eliade. He contends that all religious practitioners not only celebrate the sacred but actually participate in ceremonies that reenact and relive those sacred events that gave rise to the notion of the sacred. This idea is indicative of Eliade's spiral thought patterns. Given the universality of this pattern of "eternal return" among religious practitioners across the globe, Eliade's observation hints at the idea that spiral thought patterns are a universal characteristic of the human mind. Eliade suggests that the "eternal return" is not just restricted to religious practitioners but that secularist scientists also entertain a notion of the "eternal return" to the sacred as they try to understand the origin of the universe and the rules that govern it.

Eliade argued that within the oral tradition time is not a linear progression of events but rather a cyclic repetition of sacred events that are re-experienced through myths and the performance of rituals that give meaning to the lives of preliterate humans. For them there is no distinction between the sacred and the profane or the secular. All activities such as hunting, gathering, mating, storytelling, dancing, music, and socializing are sacred.

> To summarize, we might say that the archaic world knows nothing of "profane" activities: every act which has a definite meaning—hunting, fishing, agriculture; games, conflicts, sexuality—in some way participates in the sacred (Eliade 1964: 27–28).

There is no religion within the oral tradition because all activity is invested with the notion of the sacred and hence there is no need for religion. There are violations of the sacred but these are treated harshly by the community and can result in banishment. In oral culture one cannot live with a double standard as is the case in literate societies where people behave in one way on their day of prayer and another way for the rest of the week.

As the final example of a recent example of cyclic/spiral thought patters we consider the work of Marshall McLuhan with whom one of us (RKL) collaborated and the other (IP-O) has studied extensively. We believe that a deeper understanding of McLuhan's life work and philosophy emerges by looking at the role of spiral structures in his understanding of media and culture. We are not suggesting that the spiral was foremost in his thinking but we believe that the archetypal structure of the spiral provides a frame in which a new view of McLuhan's work emerges and one that encompasses his reversals of figure and ground and that of the reversal of cause and effect as well as the retrieval and flip in his Laws of Media (LOM).

McLuhan's intellectual roots can be found in a number of other scholars, philosophers, literary figures, and artists who embraced a spiral structure in their thinking and artistic productions including Giambattista Vico, Johann Fitche, Georg Hegel, Karl Marx, James Joyce, TS Eliot, Edgar Alan Poe, and members of the Vorticism movement including Wyndham Lewis and Ezra Pound, Sigmund Freud, and I. A. Richards, McLuhan's professor in Cambridge where he did his PhD studies.

The spiral structure of purely physical and biological phenomena is primarily played out in physical space, whereas the spiral structures of philosophy, culture, human thought, scholarship, and artistic expression involve the time dimension. The movement back and forth in these domains entails the transitions from the present back to the past or forward into the future and vice versa from the past and the future to the present. The spiral structure unites the past, the present, and the future. According to McLuhan, "We live in post-history in the sense that all pasts that ever were are now present to our consciousness and that all the futures that will be are here now." He also suggested that, "the future of the future is the present."

McLuhan felt that an understanding of history was essential for understanding the future and the impact of new technologies. He often used the metaphor of the rearview mirror, a device by which we are able to determine what is about to overtake us from our past. Furthermore, according to McLuhan, history is not to be regarded as a series of events but rather as a dynamic process with a discernible pattern, which repeats itself from culture to culture and from technology to technology.

An example of McLuhan's cyclic reversals is the way in which he sees the relationship between the users of technology and the technologies themselves. In Understanding Media: Extensions of Man McLuhan (1964) suggests that technologies are extensions of their users but then he introduces a flip in which he also suggests that the users of their technology become the servomechanisms of their own extensions, their own technologies. "To behold, use or perceive any extension of ourselves in technological forms is necessarily to embrace it. By continuously embracing technologies, we relate ourselves to them as servo-mechanisms (McLuhan 1964, 46)." At first, technology serves as an extension of humankind serving our immediate needs but unbeknownst to us our tools slowly transform our environment and we become their servants or servomechanisms. Consider how the automobile has transformed our landscape especially in North America to suit the need of the automobile rather than their drivers and pedestrians.

McLuhan's reversal of figure and ground is another spiral flip, which he expressed in the following excerpts from his writings:

> My writings baffle most people simply because I begin with ground and they begin with figure. I begin with *effects* and work round to the *causes*, whereas the conventional pattern is to start with a somewhat arbitrary selection of 'causes' and then try to match these with some of the effects. It is this haphazard matching process that leads to fragmentary superficiality. As for myself, I do not have a point of view, but simply work with the total situation as obvious *figures* against hidden *ground* (Molinaro et al. 1987: 478).

McLuhan (1964: 62) saw the creative process of both the inventor and the artist as working backwards from the effect they wanted to create to the cause that would lead to the desired effect.

> A. N. Whitehead … explained how the great discovery of the nineteenth century was the discovery of the technique of discovery. Namely, the technique of starting with the thing to be discovered and working back, step by step, as on an assembly line, to the point at which it is necessary to start in order to reach the desired object. In the arts this meant starting with the effect and then inventing a poem, painting, or building that would have just that effect and no other.

McLuhan explained how effects precede causes by showing how the effect of the telegraph was the cause of the telephone and the effect of the telegraph and the telephone was the cause of the phonograph.

The Spiral Structure of the Tetrad or Laws of Media

We will also encounter figure/ground thinking when we encounter McLuhan's (McLuhan 1975, 1977; McLuhan and McLuhan 1988) Laws of Media (LOM) also known as the tetrad. The LOM is sometimes formulated in terms of four questions and sometimes as four statements. We present both, first as four questions and then as four statements.

Four questions:

(a) What does a medium enhance?
(b) What does a medium obsolesce?
(c) What does a medium retrieve that had been obsolesced earlier?
(d) What does a medium flip into when pushed to the limits of its potential?

Four statements:

1. Every medium or technology enhances some human function.
2. In doing so, it obsolesces some former medium or technology, which was used to achieve the function earlier.
3. In achieving its function, the new medium or technology retrieves some older form from the past.
4. When pushed far enough, the new medium or technology reverses or flips into a complementary form.

In the LOM enhancement as "figure" is to obsolescence as "ground," just as retrieval as "figure" is to reversal as "ground." The LOM represents a model for the evolution of artifacts. According to the LOM, every artifact when pushed far enough flips into a new more advanced artifact. As an evolutionary model, it explains the continuous emergence of new artifacts in the ongoing cycle of the four laws of enhancement, obsolescence, retrieval, and flip. Each cycle of these four laws (the tetrad) is linked to the previous one and to the next cycle or tetrad and hence has a spiral structure.

We believe that the *Laws of Media* most vividly illustrate the spiral structure of McLuhan's thought processes that allowed him to pioneer the emergence of the new interdisciplinary and multidisciplinary field of study of media ecology, a form of systems theory in which causality operates simultaneously top down and bottom up. The LOM represents the culmination of McLuhan's lifelong project to understand media and their impact on all aspects of human life. The spiral structure of the LOM mirrors the spiral structure of the internal workings of McLuhan's thought processes and is an important part of his legacy.

1. Spiral thought patterns enhance seeing both the liminal and the subliminal.
2. They obsolesce reductionist thinking.
3. They retrieve general systems theory, cybernetics, and emergent dynamics.
4. And pushed far enough they will flip into the ultimate control of human's artifacts to serve human needs and the liberation of humankind's subservience to their technology as their servomechanisms.

Hyperlinking, Interlinking, and Cognitive Connections

Everything is interconnected and linked. No man is an island and by extension no text is an island. Nor does anything or any text stand by itself. Every figure has its ground and operates in some environment. And no environment stands by itself, but is connected to other environments. There is nothing in the literature that stands by itself. Every text digital or non-digital text is connected to other texts, some explicit as in hypertexting and some subliminally.

Hypertext arose with the Internet and the World Wide Web in which information from one document or Web site is linked to information in another document or Web site automatically when the user clicks on a hypertext link of text that is underlined in blue. We are more aware of linking in the digital age because hyperlinking is a ubiquitous feature of cyberspace. But as we will now show, hyperlinking or linking has always been a feature of the human mind. We have already claimed that creating the sets of percepts that led to verbal language was a form of hyperlinking, but let's examine other examples beginning with the oral tradition. If we define hyperlinking or hypertext more generally as the linking of one set of data or information with another set of data or information in another location, it is possible to identify pre-digital forms of hypertext in the very origin of language, in the recital or performance of epic poetry such as Homer and in written documents and books including both hand-written and printed manuscripts and books.

Oral Forms of Hyperlinking

We have claimed that the emergence of verbal language is connected to the ability of the human mind to create a linkage between percepts that share a common property and hence create a set of percepts that could be represented by a word acting as a concept. Linking is therefore an essential element of the topology of the mind.

Hypertexting or hyperlinking can also be traced back to the oral tradition before writing if we consider the way in which epic poetry was generated through the use of oral-formulaic composition as was described by Parry (1993) and his student Lord (1960), author of *The Singer of Tales*. According to Parry and Lord oral poets or the singer of tales, as in the Homeric tradition for example, composed their epic poems extemporaneously by combining elements from a store of formulae that they had memorized. Each formula had a certain metrical signature (six-colon Greek hexameter in the case of Homer) and represented a certain key idea. During the performance of a poem before a live audience the poet or singer of tales would combine or hyperlink these stored formulae to generate a unique story. The hyperlinking took place in the poet's memory that was facilitated by the meter and the rhyme of the stored formulae. The amount of information that could be stored in this manner could be quite extensive. Havelock (1963) in his book *Preface to Plato* describes Homer as a "tribal encyclopedia. The *Iliad* and the *Odyssey* are not just tales of a war and a journey home from a war, but they contain all the information a Greek needed to know to operate properly in their society. Homer's epic poems served as a 'tribal encyclopedia'" transcribed to memory through the devices of poetry into hyperlinked formulae. They are a compendium of the wisdom of a culture. It should be noted that Homer is most likely a mythical figure. The date of the composition of the opus attributed to Homer based on the references to geographic locations is well before the emergence of the Greek alphabet. So, what goes as the work of Homer is most likely the transcriptions of the epic poetry compiled into the two collections of the *Iliad* and the *Odyssey*.

In addition to the oral tradition of long ago there is another form of oral hyperlinking if one considers the way in which a conversation or dialogue transpires in the everyday affairs of all of us. As the individuals in a conversation or a dialogue take turns they often refer back to things said earlier in the conversation.

Written Forms of Hyperlinking

With the written word new forms of linkage emerged. The marginalia and illuminations of hand-written manuscripts are examples of pre-digital hyperlinking. With print new forms of hyperlinking emerged in the form of footnotes, annotations, and indices as pointed out by Ted Nelson, who coined the term hypertext.

The closest parallel to digital hypertext in print is the Hebrew Talmud. The Talmud in its printed form is a purely literary form that most closely resembles digital hypertext as each page consists of a central text that is surrounded by comments on the central text and often comments on the comments. In 1483 only 43 years after the first appearance of the Gutenberg press Joshua Solomon Soncino printed the first individual tractates of the Talmud. Soncino's innovation was not only the use of print, but also the way in which he formatted the pages of the Talmud with the original Talmudic text in the middle of the page and the commentaries of two rabbis, Rashi and Tosafot, surrounding the central text in the margins (see the accompanying figures) (Fig. 10.2).

Digital Hyperlinking

Hypertext is defined in many ways. In the Webster-Miriam dictionary, hypertext is defined as "an arrangement of the information in a computer database that allows a user to get information and to go from one document to another by clicking on highlighted words and pictures." The definition found in Wikipedia, the hypertext open-access encyclopedia, is:

> Hypertext is text displayed on a computer display or other electronic devices with references (hyper[text]links) to other text which the reader can immediately access, or where text can be revealed progressively at multiple levels of detail (Wikipedia).

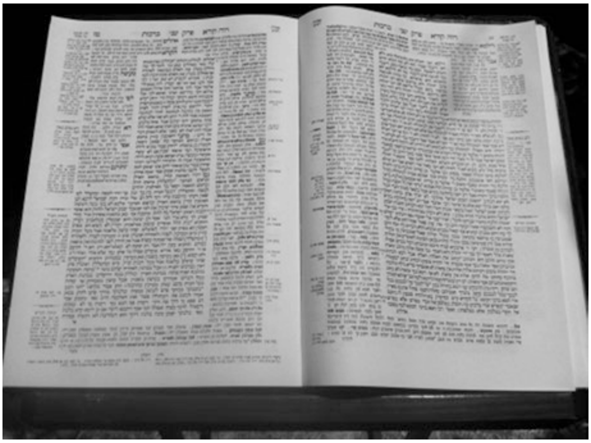


Fig. 10.2 The Talmud

These two definitions require that hypertext literature be made and experienced on a computer, and is thus medium specific. Some scholars take a different approach to defining hypertext, one that challenges the media-chauvinist argument. We find the definition of Hayles (2001), an American postmodern literary critic and electronic literature theorist, more useful because it pertains to both pre-digital and digital hypertext. "Hypertext has at least these three elements: multiple reading paths; text that is chunked in some way; some kind of linking mechanism that connects the chunks together so as to create multiple reading paths." Notice how this definition does not include the word "computer" or "database." Following Hayles' (2001) definition, hypertext is not limited to technology, content, or medium; rather it is an organizing structure for a genre of literature that is readily available on paper as well as on a computer. Hypertext is not inherently tied to electronic literature. This allows hypertext to be more than just a digital format and thus expands the concept of hyperlinking as we described in the section above.

When one hyperlinks data, one creates additional information through the network structures that are created that can have a treelike, rhizomic, or labyrinth-like structure. Hypertext is a special literary form of hyperlinking limited to written or spoken words. Examples of hypertext in digital formats include but are not limited to Wikipedia, computer networks, the Internet, the World Wide Web, individual Web sites, and Web pages to mention a few examples. Hypertext structure is a series of connections with no real beginning or end, just bunches of data to be stumbled upon. The order of stumbling becomes a path, which is rarely the same order twice (it is difficult to backtrack if you can't see where you've been like Alice in Wonderland, down the rabbit hole).

Hypertext utilizes modular theory to expand nodes of information and organize it to serve a higher function than pure input. These nodes of information diverge and expand into a networking on information, with no center and no margins. We will argue that hypertext is a method of organization that gives the reader/user more agency than non-hyperlinked text. It is by this agency that information is turned into knowledge.

Vannevar Bush and the Memex

The idea of hypertext was first conceived in a certain sense by Bush (1945) in an article he wrote for the Atlantic Monthly entitled "As We May Think" where he introduced the idea of the Memex (memory and index), a hypothetical device in which individuals would compress and store all of their books, records, and communications, "mechanized so that it may be consulted with exceeding speed and flexibility," and serve as a device that could hold humanities' collective memory. Bush did not think of this as a computer-based system but rather Bush described the Memex system as an electromechanical device enabling individuals to develop and read a large self-contained research library, create and follow associative trails of links and personal annotations, and recall these trails at any time to share them with other researchers. This device would closely mimic the associative processes of the human mind, but it would be gifted with permanent recollection. As Bush wrote, “Thus science may implement the ways in which man produces, stores, and consults the record of the race.”

Theodor Nelson, Project Xanadu

Theodor Nelson, perhaps inspired by the Memex model, began working on his own program that could link information together and mimic the associative thought patterns of its user. Nelson (1965) coined the terms hypertext to refer to linking bodies of text using electronic computers and hypermedia in his 1965 article *Complex Information Processing: A File Structure for the Complex, the Changing and the Indeterminate*. Those terms have gained common currency in the English language. He first formulated his ideas for hypertext when he was a graduate student at Harvard as early as 1960 when he began what he called Project Xanadu. Nelson saw a vision for a “digital repository scheme for world-wide electronic publishing.” Nelson formed a company to commercialize his ideas but Project Xanadu never bore fruit.

Tim Berners-Lee, The World Wide Web

Nelson’s idea of hypertext was successfully implemented with the emergence of the World Wide Web invented in 1989 by Tim Berners-Lee, who was working at the high-energy physics accelerator CERN in Switzerland. He wanted to solve the problem of the communication between physicists based in different parts of the world who had conducted experiments at CERN and wanted to share their data with each other and the general high energy physics community. He believed that a medium that allowed hypertexting or the associative linking of files that displayed both textual and visual information would serve the needs of the physicists that worked at and/or visited CERN. He also realized that this application running on the Internet would have thousands of applications outside of high-energy physics. Because the Web served the needs of the scientists at CERN he was able to convince the administrators there to support his project. In 1989 Berners-Lee developed the first version of the Hypertext Transfer Protocol (HTTP) for the distribution of hypertext and general hypermedia over the Internet. He also developed the first Web browser. Berners-Lee made the HTTP protocol available to the general public. He founded two organizations, the World Wide Web Consortium that maintains the standards for the operation of the Web and the World Wide Web Foundation that looks after ways to improve the Web as well as ways to access to it internationally.

World Wide Web brings a human dimension to IT, which facilitates collaboration. This is no accident because according to Berners-Lee (1999: 123) it was designed precisely to do this job: “The Web is more a social creation than a technical one. I designed it for a social effect—to help people to work together—and not as a technical toy.”

> The first web page went live on August 6, 1991. It was dedicated to information on the World Wide Web project and was made by Tim Berners-Lee. It ran on a NeXT computer at …. CERN. The first web page address was http://info.cern.ch/hypertext/WWW/TheProject.html. It outlined how to create Web pages and explained more about hypertext (https://www.businessinsider.com/flashback-this-is-what-the-first-website-ever-looked-like-2011-6 accessed on July 31, 2018).

The World Wide Web, Hyperlinking, and Cognition

The hypertext structure of the Internet/World Wide Web closely mimics the associative processes of the human mind. It utilizes modular theory to expand nodes of information and organize it to serve a higher function than pure input. These nodes of information diverge and expand into a networking on information, with its center everywhere and its margins nowhere. Hypertext is a method of organization that gives the reader/user more agency than non-hyperlinked text. It is by this agency that it turns information into knowledge. Just as spoken language allowed humans to conceptualize and written language allowed a more abstract level of thought that resulted in mathematics, science, social science, and philosophy so it is that hypertext promotes a systemic approach to organizing knowledge.

Conclusion

Hopefully the reader has seen the interlinking of mathematics, verbal language, mathematical notation, writing, digital hypertext, and spiral and hyperlinked structures of human cognition.

References

Berners-Lee, T. (1999). *Weaving the web*. San Francisco: Harper.
Bush, V. (1945). As We May Think. *The Atlantic* [Boston] July 1945: 101-08.
De Cruz, H. and De Smedt, J. (2013). Mathematical symbols as epistemic actions. Synthese 190 (1): 3-19.
Eliade, M. (1964). *Cosmos and history: The myth of the eternal return.* New York: Harper Torchbooks.
Havelock, E. (1963). *Preface to Plato*. Cambridge, MA: Howard University Press.
Hayles, N. K. (2001). *The transformation of narrative and the materiality of hypertext*. 1st ed. Vol. 09. Ohio State University Press, 21-39.
Logan, R. K. (2007). *The extended mind: The emergence of language, the human mind and culture.* Toronto: University of Toronto Press.
Logan, R. K. and Pruska-Oldenhof, I. (2019). *A topology of mind—Spiral thought patterns, the hyperlinking of text, ideas and more.* New York: Springer. (In preparation).
Lord, A. B. (1960). *The singer of tales.* Cambridge, MA: Harvard University Press.

McLuhan, M. (1964). *Understanding media: Extensions of man.* New York: McGraw Hill.

McLuhan, M. (1975). Communication: McLuhan's Laws of Media. *Technology and Culture* 16 (1): 74-78.

McLuhan, M. (1977). Laws of media. *English Journal* 67 (8): 92-94.

McLuhan, M, and McLuhan, E. (1988). *Laws of media: The new science*. Toronto: University of Toronto Press.

Molinaro, M., McLuhan, C. and Toye, W. (eds.) (1987). *Letters of Marshall McLuhan.* Toronto: Oxford University Press.

Nelson, T. (1965). Complex information processing: A file structure for the complex, the changing and the indeterminate. *Association for Computing Machinery (ACM) '65 Proceedings of the 1965 20th National Conference*.

Parry, M. (1993). *The making of Homeric verse: The collected papers of Milman Parry.* Oxford UK: Oxford University Press.

Chapter 11
Mathematical Fiction as an Interdisciplinary Source for Mathematics Courses: Resources and Recommendations

Frank Nuessel

Introduction

The purpose of this essay is to (1) define the notions of mathematics, literature, and mathematical fiction; (2) describe a resource-rich database for mathematical fiction; (3) discuss the rationale for using mathematical fiction in a mathematics course; (4) provide and discuss one selected exemplar of mathematical fiction; and (5) make recommendations for the use of mathematical fiction in a mathematics course. The following sections will provide details of each objective.

Definitions of Mathematics, Literature, and Mathematical Fiction

It is useful to introduce working definitions of certain key concepts in this essay before talking about the specifics. For this reason, descriptions of three important concepts "mathematics," "literature," and "mathematical fiction" are presented briefly.

The American Heritage Dictionary of the English Language (Morris 1979: 806) defines "mathematics" as "[t]he study of number, form, arrangement, and associated relationships, using rigorously defined linear, numerical and operational symbols." Devlin (2000: 7) poses the question "What is mathematics?" His simple response is that it is the "*science of patterns*" (Devlin 2000: 7). Devlin (2000: 8) goes on to state that:

> the patterns studied by the mathematician can be either real or imagined, visual, or mental, static or dynamic, qualitative or quantitative, utilitarian or recreational. They arise from the world around us, from the depths of space and time, and from the workings of the human mind. Different kinds of patterns give rise to different branches of mathematics. For example, number theory studies (and arithmetic uses) the patterns of number and counting;

F. Nuessel (✉)
University of Louisville, Louisville, KY, USA
e-mail: frank.nuessel@louisville.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_11

> geometry studies the patterns of shape; calculus allows us to handle patterns of motion; logic studies patterns of reasoning; probability theory deals with patterns of chance; topology studies patterns of closeness and position.

Likewise, literature may be defined as "an art form, or any single writing deemed to have artistic or intellectual value, often due to deploying language in ways that differ from ordinary usage" (Literature 2018). Moreover, there are distinct ways of categorizing literature (Literature 2018):

> Literature can be classified according to whether it is fiction or non-fiction, and whether it is poetry or prose. It can be further distinguished according to major forms such as novel, short story or drama; and works are often categorized according to the historical periods or their adherence to certain aesthetic features or expectations.

Mathematics and literature represent two distinct systems of communication with different representational modes. On the one hand, mathematics employs symbols to represent solutions to problems. In this regard, Danesi (2008: 37) observes that "math competence can be defined therefore as the ability to represent problems in an appropriate semiotic fashion, whereby the problem is converted into an appropriate sign-form." Danesi (2008: 119) further notes that "Story problems involve a straightforward translation of the language of the problem into the language of algebra."

Literature, on the other hand, is language based, and it makes use of a limited set of sounds (phonology), word-formation practices (morphology), and arrangement of words (lexicon) within a sentence to represent speech. This linguistic system may also be represented orthographically. Chomsky (1965: 8) points out that language can "make infinite use of finite means" by which he means that a simple set of grammatical rules is capable of generating an infinite number of grammatical sentences; hence language has an essentially creative component.

Freeman et al. (2016: 283–284) make the following observations about formal writing in mathematics and ordinary language:

> Formal writing in mathematics is a precise language that requires accuracy in its expression, especially at higher levels of mathematics study …, though it also constitutes a large part of K-12 education: in the classroom, in textbooks, and on assessments. The language of mathematics contains mathematical statements (hypotheses, conjectures, axioms, and theorems), linguistic forms and properties, grammar (connectors, combinators), and symbols. This language is often information-dense and abstract. It is also vastly different than language used in social conversation …, as is the vocabulary of mathematics with mathematical meanings being much more exact and nuanced than their ordinary definitions.

While there are distinctions between mathematical and linguistic communication, there are also areas of overlap. In some cases, mathematics and language converge in what Danesi (2008: 118) calls "story problems." Danesi (2008: 118) defines a story problem, or a word problem as "one that requires solvers to translate the language used by the problem into appropriate mathematical language." Danesi (2008: 117-153) devotes an entire chapter to this type of mathematical problem. It should be noted that Polozov et al. (2015) have proposed a useful personalized word problem generator, which allows the individualization of mathematical word problems.

O'Halloran (2000: 362) offers an insightful proposal on the evolution of mathematical problems from their linguistic representation to their contemporary visual and symbolic formats when she states that:

> A historical look at the evolution of the genres of mathematical texts suggests that the lexicogrammar of mathematical symbolism may have evolved from natural language because mathematical texts were initially written in the prose form of verbal "rhetorical algebra." These texts contained detailed verbal instructions about what was to be done for the solution of a problem. In later texts, there appeared abbreviations for recurring participants and operations in what is known "syncopated algebra." The use of variables and signs for participants and mathematical operations in the last 500 years resulted in "symbolic algebra" and the contemporary lexicogrammar of mathematics. Thus, we may conjecture that the grammar of modern mathematical symbolism grew directly out of the lexicogrammar of natural language and this may explain the high level of integration of symbolic and linguistic forms in mathematical texts.

It has been said that learning mathematics is akin to learning a foreign language. Mathematical story problems require the student to use ordinary language and translate/convert it to the language of vocabulary of mathematics. The latter has a precise and specific grammar of its own; it has a precise vocabulary with symbols that have specific meanings understood by all mathematicians. It also has its own grammar shared by mathematicians worldwide, e.g., the convention that formulas are written from left to right. Likewise, there are specific typographical conventions for mathematical statements. These formats contain highly condensed and abbreviated information that may be converted to ordinary language. With regard to these formulae, O'Halloran (2000: 361) notes that "[m]athematics is multisemiotic because the linguistic, visual and symbolic systems differentially contribute to the meaning of the text."

It is also important to have a working definition of the primary topic of this chapter, namely, mathematical fiction. Mathematical Fiction (2018) provides the following general description:

> Mathematical fiction is a genre of creative fictional work in which mathematics and mathematicians play important roles. The genre may include short stories, novels or plays; comic books; films, videos, audios.

It should be noted that comic books, films, videos, and audios all derive from scripted material, so their inclusion in this definition is appropriate. Shloming (2012: 33–34) further observes that:

> The term "mathematical fiction" is used to describe the genre of fictional works that contain mathematics. Media of mathematical fiction vary and include but are not limited to novels, short stories, movies, plays and poems. Much as mathematics is not limited to one topic, neither are the subjects considered in works of mathematical fiction. Numerous mathematical fields are represented in mathematical fiction, such as number theory, algebra, geometry, and analysis.
>
> Reading and learning mathematics through mathematical fiction may accomplish two important educational goals; namely, humanizing mathematics and exploring mathematical ideas through literature. References to mathematics in fiction as major themes of the story can reflect and shape how society perceives mathematics. Enjoyment from reading mathematical fiction can enhance motivation to learn more mathematics and may alleviate math anxiety ….

The Alex Kasman Mathematical Fiction Web Site

The representation of the science of mathematics in fictional works is considerable. Evidence for this claim comes from the web site entitled “Mathematical Fiction” created by Alex Kasman (2018) (see Mann 2010; Nuessel 2012), a professor of mathematics at the University of Charleston in South Carolina who has also written mathematical fiction (Kasman 2005). The site is a veritable cornucopia of information about mathematics in fiction. At this writing, it contains 1274 works. In the “About” section of his web page, Kasman observes that “[s]ince it is not especially significant to the purposes of this list, I am not differentiating between fiction which refers to actual mathematics and literature in which the mathematics itself is fictional.” This web site is frequently cited in studies on the use of mathematical fiction in mathematics courses. Because of its detailed categorization of mathematical fiction and its rich description of its current 1274 annotated bibliographic entries, it is considered the best and most accurate resource for mathematical fiction. It contains multiple parts, which will be presented in the following subsections.

Kasman includes an extremely useful “Search the Mathematical Fiction Database” in his web site, which contains the following components designed to reduce the user’s work:

1. Keywords in Title
2. Keywords in Author
3. Keywords in Summary
4. Medium
5. Genre
6. Topic
7. Motif
8. Math Content Rating
9. Literary Quality Rating
10. Order (Publication Date, Most Recently Added/Modified and Math Content Rating, Literary Quality Rating)

Kasman has categorized mathematical fiction by medium, motif, genre, and mathematical topic. Each grouping will be discussed briefly in the following four subsections of this chapter.

Medium in the Kasman Mathematical Fiction Site

Kasman categorizes the media type in the following formats:

1. Available Free Online (137 entries)
2. Collection (14 entries)
3. Comic Book (14 entries)
4. Films (109 entries)

5. Novels (591 entries)
6. Plays (57 entries)
7. Short Stories (489 entries)
8. Television Series or Episode (31 entries)

It must be remembered that all of these formats derive ultimately from scripted documents, so the inclusion of nonprint media is suitable for this inventory.

Motifs in the Kasman Mathematical Fiction Site

According to Kasman (2018), the following motifs appear in his voluminous list of mathematical fiction from a group of 1274 works, and that number will certainly increase with time. Several of these motifs are somewhat tangential to mathematics proper. Nevertheless, this resource is the best and the creator is meticulous in terms of content and accuracy. Kasman invites people to add comments to the web site, and many of the entries contain useful information:

1. Academia (232 entries)
2. Aliens (96 entries)
3. Anti-social Mathematicians (118 entries)
4. Autism (22 entries)
5. Cool/Heroic Mathematicians (46 entries)
6. Evil Mathematicians (49 entries)
7. Female Mathematicians (220 entries)
8. Future Prediction through Math (49 entries)
9. Gödel (42 entries)
10. Genius (55 entries)
11. Higher/Lower Dimensions (72 entries)
12. Insanity (80 entries)
13. Math as Beautiful/Exciting/Useful (76 entries)
14. Math as Cold/Dry/Useless (39 entries)
15. Math Education (131 entries)
16. Möbius Strip/Nonorientability (31 entries)
17. Music (22 entries)
18. Newton (18 entries)
19. Prodigies (82 entries)
20. Proving Theorems (118 entries)
21. Real Mathematicians (149 entries)
22. Religion (117 entries)
23. Romance 227 entries)
24. Sherlock Holmes (16 entries)
25. Time Travel (58 entries)
26. Turing (28 entries)
27. War (56 entries)

Genre in the Kasman Mathematical Fiction Site

Kasman provides his own categories for mathematical fiction as follows:

1. Adventure/Espionage (98 entries)
2. Children's Literature (94 entries)
3. Didactic (97 entries)
4. Fantasy (138 entries)
5. Historical Fiction (187 entries)
6. Horror (37 entries)
7. Humorous (244 entries)
8. Mystery (156 entries)
9. Romance (18 entries)
10. Science Fiction (478 entries)
11. Not Science-Fiction, Fantasy, or Horror (679 entries)

Topics in the Kasman Mathematical Fiction Site

Kasman has classified the 1274 examples of mathematical fiction by mathematical topic as follows:

1. Algebra/Arithmetic/Number Theory
2. Chaos/Fractals
3. Computers/Cryptography
4. Fictional Mathematics
5. Geometry/Typology/Trigonometry, Infinity
6. Logic/Set Theory
7. Mathematical Finance
8. Mathematical Physics
9. Probability/Statistics
10. Real Mathematics

Each detailed entry also features Kasman's personal assessment of the quality of the mathematical content and its literary quality on a scale of 1–5. Finally, the site provides search possibilities for keywords in title, keywords in author, keywords in summary, medium, genre, topic, motif, and math content rating and literary quality rating.

Rationale for Using Mathematical Fiction in Mathematics Classes

At this juncture, it is important to ask why mathematical fiction belongs in the curriculum of a mathematics course. There are several answers. First, it provides personal entertainment and enjoyment. Second, these works offer an alternative medium to illustrate mathematical principles taught at any level (K-16) since they offer an ancillary textual representation of mathematical principles and tenets. Mathematics is a representational code just as language is. Finally, numerous scholarly articles emphasize the value of mathematical fiction to teach the principles of mathematics because they allow the student to consider the specific mathematical topic from the perspective of another discipline and another code. These studies bridge all teaching levels (K-16). A few examples of this burgeoning literature include the following selected items (Hohn 1961; Kilman 1993; Kribs and Ruebel 2008; Padula 2005, 2006; Shloming 2012, and Zambo 2005).

Shloming (2012: 4) points out that the essential value in using fiction about mathematics is to enhance the experience of acquiring some basic principles of the discipline, namely,

> Connecting literature with mathematics can further an understanding of mathematical concepts (Bosse and Faulconer 2008; Whitin and Whitin 2004) that are taught formally. Many novels and short stories in the mathematical fiction genre are of high literary quality and mathematical exposition. These novels and short stories can educate and motivate as well as entertain the reader. Mathematical fiction can be used before and during formal learning from a textbook.
>
> Interest in mathematical learning through informal knowledge is accelerating (Asklaksen 2006). However, the matter of utilizing fictions with their mathematical content or how these fictions can influence teaching has thus far received only limited scholarly attention.

In this same vein, Padula (2006: 43) states that mathematical fiction can "… motivate students; introduce mathematical ideas in an informative context; elaborate on topics; supply imaginative applications; and help clarify mathematics."

The introduction of mathematical literature into a mathematics class allows the student to recognize and distinguish mathematical and linguistic registers. In this regard, Schleppegrell (2007: 140) observes that Halliday (1978) points out that there are two modes of discussing mathematics. First, there is the everyday language register involving counting and measuring, which often lacks precision. Second, there is a mathematical register that requires course-related teaching and learning so that the student can acquire the precision required by mathematical representation. Finally, Schleppegrell (2007: 140) cites Halliday (1978: 195–196), who states that discussing mathematics requires the student to learn new "styles of meaning and modes of argument … and of combining existing elements into new combinations." In this regard, Schleppegrell (2007: 141) states:

> In doing mathematics, it is not enough to be able to work with the language alone; mathematics draws on multiple semiotic (meaning-creating) systems to construct knowledge: symbols, oral language, written language, and visual representations such as graphs and diagrams. In addition, it uses features such as order, position, relative size, and orientation in meaningful ways (Pimm 1987) Because concepts that mathematics construct are often difficult to articulate in ordinary language, mathematics symbolism has developed to express meanings that go beyond what ordinary language can express. For example, mathematics symbolism can be used to describe relationships of parts to whole, and to construct trends and patterns of continuous covariation that cannot be presented as precisely in natural language. Visual displays, in the form of graphs and diagrams, can represent the information presented in the mathematics symbolism in ways that language cannot (O'Halloran 1999).

An essential part of understanding the basic tenets of mathematics is to be able to distinguish between mathematical and linguistic registers, and to use them both appropriately. The inclusion of mathematical fiction allows the student to make this differentiation through exercises that involve the translation of one register into another: (1) ordinary language register to mathematical register, and (2) mathematical register to ordinary language register. Initially, these tasks will challenge the student. However, with sufficient practice, the student will acquire mastery and competency of both systems.

The studies cited in this section on the rationale for using mathematical fiction in mathematics courses find theoretical support for their differentiation of linguistic and mathematical registers in the significant treatise by Sebeok and Danesi (2000) on modeling systems. In simple terms, human beings have an innate capacity to create models of the world through their perceptual sensory systems (sight, hearing, touch, taste, smell). This process is called semiosis, which Sebeok and Danesi (2000: 5) define as "[t]he ability to make models is, actually, a derivative of *semiosis*, defined simply as the capacity of a species to produce and comprehend the specific types of models it requires for processing and codifying perceptual input in its own way." This process consists of four phases: sensory perceptions > semiosis > modeling > representation (Sebeok and Danesi 2000: 6).

Sebeok and Danesi (2000: 32) further note that:

> A cohesive modeling system is known in traditional semiotic theory as a code, a system providing particular types of signifiers that can be used in various ways and for diverse representational purposes … A language code, for instance, provides a set of phonetic, grammatical, and lexical 'instructions' that the producers and interpreters of words and verbal texts can recognize and convert into messages.
>
> Generally speaking, for some particular representational need there is an optimal code or set of codes that can be deployed.

Sebeok and Danesi (2000: 34) also point out that "[t]he use of a code to make signs is called *encoding*, the reception or interpretation of signs or texts is called *decoding*."

In acquiring the ability to understand the mathematical register, the student must develop the ability to understand, and subsequently use, the mathematical code or register in a meaningful and appropriate fashion in specific contexts (classroom, problem-solving activities, and written papers). In the fourth chapter of their book on modeling, Sebeok and Danesi (2000: 120) discuss the notion of "tertiary modeling system," which they define as "the system that undergirds highly abstract, symbol-based modeling." In this discussion, they (Sebeok and Danesi 2000: 120–129) offer examples from mathematics (geometry, algebra) to illustrate their point.

Mathematical literature provides the student with the ability to comprehend the mathematical register via ordinary language. Translational exercises will allow the student to take the mathematical information in ordinary language and develop a competency in representing it in mathematical code through the process of "code-switching."

Selected Example of Mathematical Fiction

In her detailed discussion of 26 novels and short stories that feature geometric themes, Shloming (2012) discusses one work that has received frequent reference in the discussion of the use of mathematical fiction in mathematics courses, namely, *Flatland: A Romance of Many Dimensions* by Edwin Abbott Abbott (1838–1926), a British teacher and author, which was published in 1884 (Abbott 2002). This novel is a prototypical work of fiction that addresses plane geometry in a completely understandable fashion.

Abbott's (2002) *Flatland: A Romance of Many Dimensions* has received a great deal of academic attention as a good example of mathematical fiction that is usually an ancillary reading for mathematics classes at all levels (Danesi 2003: 75–76, Danesi 2018: 115–119, Dotson 2006, Mann 2010, Shloming 2012: 39, 50–52, *passim*, Padula 2005, 2006, Wallace et al. 2011, Sriraman 2003, 2004, Sriraman and Beckmann 2018).

Kasman's (2018) entry in his web site Mathematical Fiction gives this novel a rating of 3.76 based on the quality of its mathematical content and 2.5 on the basis of its literary quality. This work is widely cited in academic papers about mathematical fiction. It is worth reproducing Kasman's comments on this text to give the reader a sense of how a mathematician approaches mathematical literature. After an extensive review of the annotations, I found them to be consistently good and accurate.

> This is the *classic* example of mathematical fiction in which the author helps us to think about the meaning of "dimension" through fictional example: a visit to a world with only two spatial dimensions.
>
> One of the genres used in this Website is "didactic". I classify works of fiction as "didactic" if the intention of the author is to use the fiction to teach mathematics. For example, Enzenberger's "Der Zahlenteufel" is didactic because the story does not really matter at all; the purpose of that novel is to interest the reader in the real mathematics that it discusses. The idea is that many readers who have trouble with abstract mathematical thinking will understand it better if it is included in a story and given some sort of fictional "reality".
>
> Many people do have trouble conceiving of higher dimensional geometry, and a reference to *Flatland* is now commonly used by people who are trying to help others understand this difficult concept. It does seem to help people to imagine creatures living and thinking in a two-dimensional universe and to imagine how *they* would perceive the three-dimensional objects that are familiar to us. So, people certainly use Flatland as a didactic work of mathematical fiction.
>
> However, I do not think Edwin Abbott Abbott was using math that way. It was not his goal to make the math more understandable and believable by including it in a story. Quite the opposite, in fact. I think Abbott thought of math as something that people would already understand and wanted to use that math to discuss certain non-mathematical ideas that were important to him. Moreover, he hoped that by using mathematics (a topic most people agree upon), he would be able to generate some agreement in discussing something more controversial. In particular, I think that what he really wanted to write about was not mathematics but the relationships between people and the relationship between people and the supernatural.
>
> Consider this: the main character, "a square", of *Flatland* has had an experience with something from beyond his own universe, something he cannot see entirely but can only glimpse in pieces. This has changed his view of his reality, changed his view of his relationship with the other creatures of Flatland, and he wants to share that information.

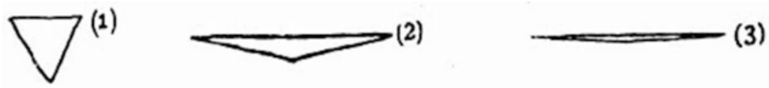
> Now, notice that because of his strangely repetitive name ("Abbott Abbott"), the author of Flatland could also describe himself as "A Squared" = "A^2." Abbott was a theologian. He presumably also believed that he could perceive God's existence, but not entirely, only in pieces. And although some of his ideas seem to reflect an old fashioned bias to modern readers (e.g. that the females in Flatland are line segments while the males are polygons) he was actually somewhat progressive for his day. His view of the relationships between people was also rather introspective for Victorian England.
>
> Consequently, I believe that the role of mathematics in Flatland was to provide Abbott with a language (the language of geometry) through which he could discuss non-mathematical ideas with the readers that he otherwise could not quite put into words.

Kasman invites readers of his Mathematical Fiction (2018) web site to send in their comments with the writer's name or anonymously. In the case of the *Flatland* (Abbott 2002) entry, there are 14 comments from different people. One of the comments points out that there are two film versions of *Flatland*, namely, *Flatland the Movie* directed by Johnson and Travis (2007), which features the voices of the well-known actors Martin Sheen, Michael York, Kristen Bell, and Joe Estevez in this animated short film. The second one (*Flatland: The Film*) is a full-length animated movie directed by Ehlinger Jr. (2007). It features the voices of actors Ashley Blackwell, Chris Carter, Megan Colleen, and Ladd Ehlinger, Jr.

Danesi (2018: 115–118) discusses *Flatland: A Romance of Many Dimensions* (Abbott 2002) in his *Basic Dictionary of Puzzles and Games*. The novel begins with an introduction to the two-dimensional world of "flatland" (Abbott 2002: 33–34):

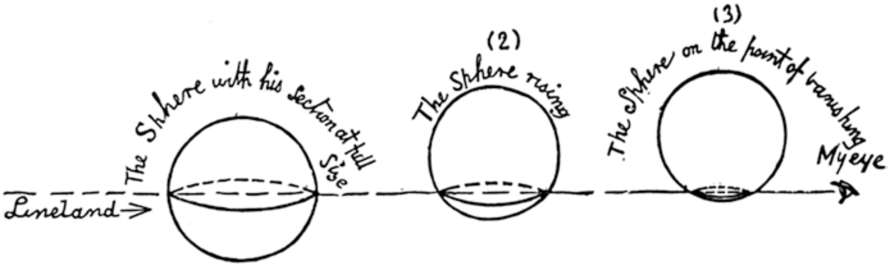
> I call our world Flatland, not because we call it so, but to make its nature clearer to you, my happy readers, who are privileged to live in Space.
>
> Imagine a vast sheet of paper on which straight Lines, Triangles, Squares, Pentagons, Hexagons, and other figures, instead of remaining fixed in their places, move freely about, on or in the surface, but without the power of rising above or sinking below it, very much like shadows—only hard and with luminous edges—and you will then have a pretty correct notion of my country and countrymen. Alas, a few years ago, I should have said "my universe": but now my mind has been opened to higher views of things.
>
> In such a country, you will perceive at once that it is impossible that there should be anything of what you call a "solid" kind; but I dare say you will suppose that we could at least distinguish by sight the Triangles, Squares, and other figures, moving about as I have described them. On the contrary, we could see nothing of the kind, not at least so as to distinguish one figure from another. Nothing was visible, nor could be visible, to us, except Straight Lines; and the necessity of this I will speedily demonstrate.
>
> Place a penny on the middle of one of your tables in Space; and leaning over it, look down upon it. It will appear a circle.
>
> But now, drawing back to the edge of the table, gradually lower your eye (thus bringing yourself more and more into the condition of the inhabitants of Flatland), and you will find the penny becoming more and more oval to your view, and at last when you have placed your eye exactly on the edge of the table (so that you are, as it were, actually a Flatlander) the penny will then have ceased to appear oval at all, and will have become, so far as you can see, a straight line.

The same thing would happen if you were to treat in the same way a Triangle, or Square, or any other figure cut out of pasteboard. As soon as you look at it with your eye on the edge on the table, you will find that it ceases to appear to you a figure, and that it becomes in appearance a straight line. Take for example an equilateral Triangle—who represents with us a Tradesman of the respectable class. Figure 1 represents the Tradesman as you would see him while you were bending over him from above; Figs. 2 and 3 represent the Tradesman, as you would see him if your eye were close to the level, or all but on the level of the table; and if your eye were quite on the level of the table (and that is how we see him in Flatland) you would see nothing but a straight line:



In a second passage in the novel (Abbott 2002:142–144), there is a description of what a person from "Spaceland" would perceive in Flatland reproduced here:

The diminished brightness of your eye indicates incredulity. But now prepare to receive proof positive of the truth of my assertions. You cannot indeed see more than one of my sections, or Circles, at a time; for you have no power to raise your eye out of the plane of Flatland; but you can at least see that, as I rise in Space, so my sections become smaller. See now, I will rise; and the effect upon your eye will be that my Circle will become smaller and smaller till it dwindles to a point and finally vanishes.



There was no "rising" that I could see; but he diminished and finally vanished. I winked once or twice to make sure that I was not dreaming. But it was no dream. For from the depths of nowhere came forth a hollow voice—close to my heart it seemed—"Am I quite gone? Are you convinced now? Well, now I will gradually return to Flatland and you shall see my section become larger and larger."

Every reader in Spaceland will easily understand that my mysterious Guest was speaking the language of truth and even of simplicity. But to me, proficient though I was in Flatland Mathematics, it was by no means a simple matter. The rough diagram given above will make it clear to any Spaceland child that the Sphere, ascending in the three positions indicated there, must needs have manifested himself to me, or to any Flatlander, as a Circle, at first of full size, then small, and at last very small indeed, approaching to a Point. But to me, although I saw the facts before me, the causes were as dark as ever. All that I could comprehend was, that the Circle had made himself smaller and vanished, and that he had now reappeared and was rapidly making himself larger.

Both passages from *Flatland: A Romance of Many Dimensions* (Abbott 2002) illustrate why this novel would be useful as collateral reading in a mathematics course on plane geometry. First, its plane geometric descriptions are accurate. Second, the inclusion of graphics that resemble plane geometric configurations provides the reader with a visual dimension that much fiction lacks with the exception of the graphic novel, which is a distinct hybrid visual and textual genre (Danesi 1983). Third, the novel allows the student to translate ordinary language descriptions of plane geometry into the precise language of this mathematical discipline. Finally, the second passage about "Spaceland" also provides the opportunity to discuss solid geometry or three-dimensional Euclidian space.

In addition to its important mathematical value, the novel also contains social criticism of the classism of Victorian England (1837–1901), e.g., the shapes of the characters correspond to their social position in society: (1) isosceles triangles (soldiers and workmen), (2) squares and pentagons (doctors, lawyers, and other professions), (3) hexagons (the lowest rank of nobility), and (4) circles (the priest class). Women, on the other hand, are only lines, which is a misogynistic element of the novel because they are perceived as a tiny dot, and therefore insignificant, in Flatland's two-dimensional world.

Recommendations

This section contains a set of recommendations for the use of mathematical fiction in a mathematics course.

1. Consult Alex Kasman's web site entitled "Mathematical Fiction." Its detailed annotations and commentary by Kasman himself as well as readers of his web site provide extremely useful observations and interpretations about each work. Especially useful are his categorizations of each work, and an assessment on a scale of 1–5 of the mathematical and literary quality of each work.
2. Select a work that is appropriate to the topic that you are teaching (see topics above).
3. Select a work whose mathematical value and accuracy have the highest rating (1–5).
4. Use mathematical fiction as a tool to show how the ordinary linguistic register is distinct from the mathematical one. A useful activity is to ask students to translate a passage from mathematical fiction to a mathematical register and vice versa. This allows the student to become fluent in both registers. A student can "translate" a mathematical code into ordinary language. Likewise, a student can translate the ordinary language of mathematical fiction into a mathematical code, thereby allowing the student to become fluent and proficient in both systems.

Concluding Remarks

Mathematical fiction is an excellent ancillary tool for inclusion in mathematics classes because it provides an ordinary language perspective of the distinct and unique mathematical register in mathematics textbooks. Its selective incorporation into a mathematics course allows students to differentiate the two codes or registers. Furthermore, mathematical fiction allows the student to translate the ordinary language of mathematical fiction into a suitable corresponding mathematical register. This type of encoding and decoding practice enhances the student's knowledge of the new mathematical code by seeing the distinctions between three semiotic systems (linguistic, symbolic, iconic). An example of mathematical fiction, Edwin Abbott's *Flatland: A Romance of Many Dimensions* (Abbott 2002), was used to exemplify the use of this type of fiction in a course on mathematics. Finally, a set of recommendations for the consequential introduction of mathematical fiction into a mathematics course was provided.

References

Abbott, E. A. (2002 [1884]). The annotated flatland: A romance of many dimensions. Introduction and notes by Ian Stewart. New York: Basic Books. [The complete original novel with graphics is available at the following Web Site: http://www.authorama.com/book/flatland.html. Web. 3 August 2018.

Asklaksen, H. (2006). Book review of *The Curious Incident of the Dog in the Night-Time* by M. Haddon. Notices of the American Mathematical Society 54(3): 343-345.

Bosse, M. J. and Faulconer, J. (2008). Learning and assessing mathematics through reading and writing. School Science and Mathematics 108(1): 8-18

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Danesi, M. (1983). Pedagogical graphics in second-language teaching. The Canadian Modern Language Review 40: 73-81.

Danesi, M. (2003). The puzzle instinct: The meaning of puzzles in everyday life. Bloomington, IN: Indiana University Press.

Danesi, M. (2008). Problem-solving in mathematics: A semiotic perspective for educators and teachers. New York: Peter Lang Publishing.

Danesi, M. (2018). Basic dictionary of puzzles and games. München: Lincom.

Devlin, K. (2000). The math gene. How mathematical thinking evolved and why numbers are like gossip. New York: Basic Books.

Dotson, Daniel. (2006). Portrayal of mathematicians in fictional works. CLCWEB: Comparative Literature and Culture https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1324&context=clcweb. Web. 9 August 2018. 8(4): Not paginated.

Ehlinger, L., Jr. (2007). Flatland. https://www.imdb.com/title/tt0972374/. Web. 3 August 2018.

Freeman, B., Higgins, K. and Horney, M. (2016). How students communicate mathematical ideas: An examination of multimodal writing using digital technologies. Contemporary Educational Terminology 7(4): 281-313.

Halliday M.A.K. (1978). Language as social semiotic. London: Edward Arnold.

Hohn, F. E. (1961). Teaching creativity in mathematics. The Arithmetic Teacher 8(3): 102-106.

Johnson, Dano and Travis, J. (2007). Flatland: The Movie. https://www.imdb.com/title/tt0814106/. Web. 3 August 2018.

Kasman, A. (2005). Reality conditions. Washington, DC: Mathematical Association of America.

Kasman, A. (2018). Mathematical fiction. http://kasmana.people.cofc.edu/MATHFICT/. Web. 31 July 2018.

Kilman, M. (1993). Integrating mathematics and literature in the elementary classroom. The Arithmetic Teacher 40(6): 318-321.

Kribs, C. M. and Ruebel, K. K. (2008). Exploring mathematical concepts in literature. Middle School Journal 40(1): 36-42.

Literature (2018). https://en.wikipedia.org/wiki/Literature. Web. 2 August 2018.

Mann, T. (2010). From Sylvia Plath's *The Bell Jar* to the Bad Sex Award. BSHM Bulletin 25: 58-66.

Mathematical Fiction. (2018). https://en.wikipedia.org/wiki/Mathematical_fiction. Web. 1 August 2018.

Morris, W. (ed.) (1979). The American Heritage dictionary of the English language. Boston: Houghton Mifflin.

Nuessel, F. (2012). The representation of mathematics in the media. In: M. Bockarova, M. Danesi, and R. Núñez (eds.), Semiotic and cognitive science essays on the nature of mathematics, pp. 165-208. München: Lincom.

O'Halloran, K. L. (1999). Towards a systemic functional analysis of multisemiotic mathematics texts. Semiotica 124 (1/2): 1-29.

O'Halloran, K. L. (2000). Classroom discourse in mathematics: A multisemiotic analysis. Linguistics and Education 10(3): 359-388.

Padula, J. (2005). Mathematical fiction: Its place in mathematics learning. Australian Mathematics Teacher 61(4): 6-13.

Padula, J. (2006). Mathematical Fiction for Senior Students and Undergraduates: Novels, Plays, and Film, Australian Senior Mathematics Journal 20(2): 36-44.

Pimm, D. (1987). Speaking mathematically: Communication in mathematics classrooms. London: Routledge and Kegan Paul.

Polozov, O., O'Rourke, E., Smith, A. M., Zettlemoyer, L., Gulwani, S., and Popović, Z. (2015). Personalized mathematical word problem generation. https://www.microsoft.com/en-us/research/uploads/prod/2018/01/ijcai2015-word-problems.pdf. 5 August 2018.

Schleppegrell, M. J. (2007), The linguistic challenges of mathematics teaching and learning: A research review. Reading & Writing Quarterly 23(2): 139-159.

Sebeok, T. A. and Danesi, M. (2000). The forms of meaning: Modeling systems theory and semiotic analysis. Berlin: Mouton de Gruyter.

Shloming, J. R. (2012). *Analysis of mathematical fiction with geometric themes*. Ph.D. Diss. Columbia University.

Sriraman, B. (2003). Mathematics and literature. Synonyms, antonyms or the perfect amalgam? Australian Mathematics Teacher 59(4): 26-31.

Sriraman, B. (2004). Mathematics and literature (the sequel): Imagination as a pathway to advanced mathematical ideas and philosophy. Australian Mathematics Teacher 60(1): 17-23.

Sriraman, B. and Beckmann, A. (2018). Mathematics and literature: Perspectives for interdisciplinary pedagogy. *Proceedings from the 9th International History, Philosophy and Science Teaching Group (IHPST)* 2007, Calgary, June 22–26, 2007. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.557.755&rep=rep1&type=pdf. 6 August 2018. No pagination.

Wallace, F. H., Evans, M. A., and Stein, M. (2011). Geometry sleuthing in literature. Mathematics Teaching in the Middle School. 17(3): 154-159.

Whitin, D. J. and P. Whitin. (2004). New visions for linking literature and mathematics. Reston, VA: National Council on English.

Zambo, R. (2005). The Power of two: Linking mathematics and literature. Mathematics Teaching in the Middle School 10(8): 394-398.

Chapter 12
Science, Magic, and the In-Between: Whence Logic

Inna Semetsky

In the course of modern history, science and magic have gradually become separated into a pair of binary opposites. While acknowledging what the "pure reason" of modernity considered to be a supernatural action, science nevertheless attempted to explain the latter in terms of a regular method of a direct cause-effect connection as a method in natural science, promptly arriving at a conclusion of either anomalous effect (as in magic) or anomalous cause (as in mantic). But can what is called magic still be considered a science—a science of hidden relations that are nevertheless, and in accord with Charles S. Peirce's pragmatic maxim, capable of producing real effects? Surely John Deely (2001) acknowledged Peirce's vision as rooted in science rather than mysticism. This chapter uses one of the Tarot cards called the Magician as an index of overcoming a schism between the dual opposites when positioned in the conceptual framework of semiotics that allows us to elucidate the meaning of this sign (Fig. 12.1).

Systems theorist Erich Jantsch (1980) defined consciousness as the degree of autonomy a system gains in the dynamic relation to its environment—thereby even the simplest chemical dissipative structure can possess "a primitive form of consciousness" (1980: 40). The image of the Magician represents such a trace of consciousness in the material universe, in agreement with Alfred North Whitehead's concept of protomentality. It was Ludwig von Bertalanffy, the founder of the general systems theory, who addressed the insufficiency of the analytical procedures of mechanistic science based on linear causality between two basic variables and attracted our attention to "new categories of interaction, transaction, teleology"

This essay is a modified and updated version of the earlier 2008 paper titled "The Transversal Communication, or: Reconciling Science and Magic" published in the journal *Cybernetics and Human Knowing,* Vol. 15, No. 2, pp. 33–48. See Semetsky (2008) in references, I acknowledge the original publication with gratitude.

I. Semetsky (✉)
Institute for Edusemiotic Studies, Melbourne, VIC, Australia
e-mail: irs5@columbia.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_12

Fig. 12.1 The Magician

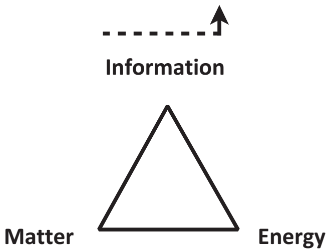


Fig. 12.2 The world as perfused with signs

(Bertalanffy 1972: xix): indeed, interactions between more than two objects create an unsolvable problem within the equations of classical mechanics. Importantly, the "interactions do not have to be *physical*; they can also be thought of as a transference of *information*" (Cilliers 1998: 3) along a semiotic bridge created as if by wave of the Magician's wand. According to MIT professor Seth Lloyd (2006), material elements such as "Earth, air, fire, and water … are all made of energy, but the different forms they take are determined by information. To do anything requires energy. To specify what is done requires information. Energy and information are by nature (no pun intended) intertwined" (2006: 44). Such intertwined relation parallels Peirce's triad. The world which, according to Peirce, is perfused with signs is thus intrinsically informational, and we can represent its dynamic structure by a diagram (Fig. 12.2).

Non-incidentally, the four tools on the Magician's table (aligned with the four suits in a Tarot deck: wand, pentacle, cup, and sword) correspond to the four elements available to the Magician in his alchemical laboratory: fire, earth, water, and air. Alternatively, they relate to the four Jungian functions comprising the Magician's semiotic reason: thinking, feeling, sensing, and intuiting. The Magician icon is a sign of mind being embodied in matter due to the evolutionary process of semiosis wherein Peircean Thirdness functions as a "mediation, whereby first [mind] and second [matter] are brought into relation" (Peirce CP 6. 7). The Magician's communicative action is a symbolic dialogue, an interaction, or relation as an ongoing event represented by means of two indices. While his right hand holding the wand points upwards, to the skies, the left hand points to the earth, thus enacting the Hermetic maxim in the ancient text *Emerald Tablet* that proclaims the formula of analogy: *That which is above is like to that which is below and that which is below is like to that which is above, to accomplish the miracles of one thing*. The number corresponding to the Magician in a deck is 1 as a symbol of the Whiteheadian *one* world without and within: *as above so below*. The action of signs crosses over the dualistic gap between mind and matter, science and magic, process and structure, subject and object, and partakes of a specific "communication mechanism which is capable of acting much faster than metabolic communication" (Jantsch 1980: 156). Such process may very well operate in qubits (Lloyd 2006) which are the swift bits of quantum information inaccessible to the usual sense perception.

Contemporary cosmology assigns to the natural world the status of a giant quantum computer that processes information in qubits. Hence follows the motto "it from bit" or rather "it from qubit" which means that the observable world arises out of information on the basis of which the universe computes its own dynamical evolution while actualizing potential reality as the computation proceeds. Ditto for the evolution of the human mind: mind and intelligence are evolving. However, the computational approach needs a bit (no pun intended) of qualification. At the cutting edge of cognitive science, computers are understood as dynamical systems that indeed manipulate bits, but these units of information are not strictly reducible to what in physics are called particles. They are moments in the flow represented, importantly, by analogue and not solely digital information. Lloyd (2006), stressing that the universal quantum computation proceeds in a dual (analogue-digital) mode, specifies the structure of the computational space in terms of a circuit diagram representing both logical gates (the places where qubits interact, thus exchanging/transforming information) together with unorthodox causal connections represented by the connecting "wires" or paths along which the information flows. Therefore these moments in the continuous flow of semiosis can be defined as discrete bits only within a certain context—that is, taken as already parts-of-the-whole (cf. Rockwell 2007). The flow of information enabled by semiotic communication establishes different and new relations so that the system's boundaries are crossed and traversed, and new boundary conditions of the system, or its external structure, are being established meanwhile sustaining the integrity of its internal structure in the manner of what Gilles Deleuze aptly called the fold as "the inside *of* the outside" (Deleuze 1988: 96).

As semiotician Floyd Merrell remarks, "the fascination of children with … Winnie the Pooh, and … Alice's adventures—also a favorite pastime of logicians, mathematicians, and physicists—attests to their import of 'primitive' perceptual and conceptual modes, keenly picked up by philosopher Gilles Deleuze" (Merrell 1996: 141). Is Tarot also such a "primitive" mode (Semetsky 2011)? Deleuze's philosophy employs Riemann's innovative geometry of surfaces as well as Lautman's notion of transcendence-immanence of ideas in mathematics. From Leibniz, Deleuze borrows the notion of "esoteric" infinitesimal calculus of ideas which are obscure problematic instances that, instead of being a priori direct, clear and distinct representations in consciousness (as Descartes proclaimed them to be) express the vague, fuzzy, and sub-representative "presentation of the unconscious" (Deleuze 1994: 192). Such calculus partakes of *mathesis universalis,* the hypothetical universal mathematics applicable to all branches of science while also unifying science with art, spirituality, and magic. Mathesis—also translated as the science of learning—represents "an alphabet of what it means to think" (Deleuze 1994: 182). Leibniz included pictures and "various graphic geometrical figures" (Nöth 1995: 274) as a possible medium of its characters; still his project remained unfinished. Ultimately, the "characters [of mathesis] were to be isomorphic with the concepts designated by them; [and] the universal signs were to be isomorphic with the facts of nature" (1995: 274). As for Deleuze, he was adamant that "to believe that mathesis is merely a mystical lore inaccessible and superhuman, would be a complete mistake" (Deleuze 2007: 143). Is the Magician a practitioner of mathesis? It sure appears so because it is mathesis that "transforms knowledge itself into a sensible object [and] insists upon the correspondences between material and spiritual creation" (2007: 151) in the tradition of Hermeticism.

The Magician is a symbolic problem-solver. Problems—not solely mathematical but existential as well—belong to the level of the virtual and unconscious, while solutions—to the level of actual, conscious experience. The realm of the actual "contains" identities (as subject to the logical copula) but the virtual realm is the domain of differences which are characterized by differential relations and corresponding singular points. Difference is a central concept in Deleuze's philosophy: "Everything which happens and everything which appears is correlated with orders of differences: differences of level, temperature, pressure, tension, potential, *difference of intensity*" (Deleuze 1994: 222). The unconscious ideas comprise "differential flashes which leap and metamorphose" (1994: 146) and amount to the new image of thought grounded in differences.

The Magician is just an idea, a virtual tendency. Yet, while seemingly musing *in potentia,* it still possesses a peculiar "feeling of the direction and end of various lines of behavior [as]… the feeling of habits working below direct consciousness" (Dewey 1922/1988: 26). The Magician's "transversal communications" (Deleuze and Guattari 1987: 11) between different levels bring life and vitality into the world of supposedly inert, unanimated, matter: matter becomes mindful! Physicist Henry Stapp (2007: 10) points out that John von Neumann, in his mathematical formulation of quantum mechanics, specifically coined *intervention* as a term describing the effects of free choices upon the physical word; yet, these free choices are themselves dependent on reasons, values, and unconscious motivations. And it is an act of intervention as the prerogative of the Magician that enables the functioning of this sign in the manner of an autocatalytic element representing "kinetics effective in this moment at each spacial point" (Jantsch 1980: 34). The relation between different levels or different terms can be described by the derivative of a function in the form **dy/dx**, where the values of the terms **x** and **y** do not have to be determined. What is important is that they exist absolutely and only in their relation to each other. Relations are prior (or external) to their respective terms! When we encounter a problem, its unknown and unidentified terms and conditions (of which we thus remain unconscious) are similar to the yet undetermined values of **x** and **y** that however are capable of determination precisely via their differential relation. The conscious mind then "incorporates all the power of a differential unconscious, an unconscious of pure thought which internalizes the difference … and injects into thought as such something unthought" (Deleuze 1994: 174). Understanding and meaning—solving a problem, becoming aware of the unconscious, creating a novel concept—derive "from the mathematical function of differen*t*iation and the biological function of differen*c*iation" (1994: xvi). Differentiation (with a "*t*") is the operation of difference, of inequality, and it is "in difference that movement is produced as an 'effect', that phenomena flash their meaning like signs" (1994: 57), thereby engendering the process of semiosis. But there is also differenciation (with a "*c*"), an in-itself or the second part of difference producing multiple "local integrations, as mathematicians say" (1994: 98). Such a double process of different/ciation, as the Magician's communicative action, appears to border on a magical act indeed when this sign intervenes between the different levels: the Magician lifts up the magic wand and makes "events turn into objects, things with meaning" (Dewey 1925/1958: 166) while actualizing the virtual reality of signs and bringing the unconscious to the level of consciousness.

Addressing the "social consequences of the misrepresentations of contemporary scientific knowledge" (Stapp 2007: viii), Stapp posits a mindful universe that consists of psychophysical (not just physical or material) building blocks and in which the transition from potentiality to actuality is indeed possible. "Idea-like qualities" (2007: 97) are therefore signs as part of parcel of semiotic hence "non-anthropocentric ontology" (2007: 97). Stapp contends that:

> the *physically described world* is built…out of objective *tendencies*—potentialities—for certain discrete, whole *actual events* to occur. Each such event has both a psychologically described aspect, which is essentially an increment in knowledge, and also a physically described aspect, which is an action that *abruptly changes* the mathematically described set of potentialities to one that is concordant with the increase in knowledge (2007: 9).

The actualization of potentialities hiding at the level of the unconscious is taking place due to the subjective, bottom-up, "intervention of the mind" (Shimony 1993/Vol. II: 319) into the chain of semiosis. Yet this very intervention may be considered objective in the sense of itself being implemented by a choice of a global, top-down, character analogous to the semiotic functioning of the relation between immanence and transcendence embedded in one inseparable process of semiosis. The choice of this kind may be accounted for by what philosopher of science Abner Shimony, addressing "the status of mentality in nature" (Shimony in Penrose 1997: 144), dubbed the hypothetical super-selection rule in nature that enables the very "transition between consciousness and unconsciousness … not … as a change of ontological status, but as a change of state" (1997: 150). The magic wand in the Magician icon thus is a symbol of "another kind of causation" (Peirce CP 6. 59) or a possible self-cause disregarded by classical science. The dynamics of self-organization (Jantsch 1980) proceeds in an autopoietic (cf. Varela 1979) manner along environmental perturbations and compensations effectuated by means of a semiotic "bridge, a transversality" (Guattari 1995: 23) between different, heterogeneous planes. Says Deleuze: "I undo the folds of consciousness that pass through every one of my thresholds, 'the twenty-two folds' that surround me and separate me from the deep" (1993: 93). This number corresponds to the 22 major cards in a Tarot deck. The Magician establishes coordination (Peirce's category of Thirdness) between the noumenal and phenomenal realms despite—or rather, due to—the original difference between the two so that the former becomes potentially knowable (counter to Kant) even if not presently known. The relation between *ens reale* and *ens rationis* does not mean their identity: the latter can never be completely preserved "in any advance to novelty" (Whitehead 1966: 107). However the Magician as a sign "that flashes across the system, bringing about the communication between disparate series" (Deleuze 1994: 222) of events creates a link between the physical world of facts and the world of objective meanings or values: for Whitehead, facts are creative or valuative due to the principle of creativity as a precondition for novelty.

Whitehead's philosophy of the organism posits actual occasions as spatiotemporal events endowed with experience that, albeit dim and not fully conscious, nevertheless defies the sharp bifurcation of nature into mindless matter and conscious mind. In contemporary physics event is defined as an actualized possibility of this event's objective tendency, or its *potentia*, to occur. In general relativity, events exert a causal influence on the very structure of events: structures are thereby evolving, that is, they are process~structures that defy the strictly linear causality of classical mechanics. The notation "~" (tilde or squiggle) is an unusual punctuation as a sign of what the cutting-edge empirical science of coordination dynamics indicates in terms of a "reconciliation of complementary pairs" (Kelso and Engstrøm 2006: 63) versus the otherwise disconnected opposites. Coordination dynamics exhibits "reciprocal causality" (2006: 115) which operates two-directionally: "from the bottom up (projection) and then from the top down (reinjection)" (Griffin 1986: 129)—just like as per symbolism of the Magician. The feature of double codification (cf. Hoffmeyer and Emmeche 1991) pertinent to the action of the Magician (analogue and digital, virtual and actual) relates to a specific problem in philosophy of science specified as the one that "for both Whiteheadian process and quantum process is the emergence of the discrete from the continuous" (Stapp 2007: 88). The operation of projection is significant in both mathematical and psychological terms. Stapp posits the hypothetical mechanism of a spontaneous quantum reduction event associated with "a certain mathematical 'projection' operator" (2007: 94) whose action seems to be direct (via projection) but which also causes indirect changes, producing faster-than-light effects, thus demonstrating in practice what has long been considered a "spooky" action at a distance.

The functioning of Tarot presupposes the projection of signs onto a surface, which always involves an ontological loss of dimension expressed, as Deleuze would say, in $(n - 1)$ dimensions. In contemporary cosmology, the so-called weak holographic principle (Smolin 2001) posits the world as consisting of processes or events, which can only be perceived through representations. Theoretically, representations—or, in semiotic terms, signs that by definition conform to the medieval *aliquid pro aliquo* formula—are all there is: they represent Whitehead's one kind of entity. These dynamical entities are "representations by which one set of events in the history of the universe receives information about other parts of the world" (Smolin 2001: 177). Because they occur on a scale unavailable to the ordinary sense perception, they sure enough can be seen only in their projected format: a loss in dimensions is thus implied. In cosmology, the reduction in dimensions is called *compactification.* We do not know, in general, the total number of hidden dimensions that may have been compactified (cf. Lloyd 2006). For example, a cinema projection on the screen compactifies our regular three-dimensional reality into only two dimensions. The screen metaphor is potent: it accords with the layout of Tarot pictures spread on a flat surface, making a surface literally a locus of meanings (cf. Deleuze 1990). The fact is that:

> the area of a screen—indeed, the area of any surface in space—is really nothing but the capacity of that surface as a channel for information. So, according to the weak holographic principle space is nothing but a way of talking about all the different channels of communication that allow information to pass from observer to observer … In short, the holographic principle is the ultimate realization of the notion that the world is a network of relationships. These relationships are revealed by this new principle to involve nothing but information (Smolin 2001: 177–178).

The layout of Tarot cards—functioning as a screen or projection—thus presents a spatiotemporal organization of informational bits and pieces (pun intended) in the form of signs embodied in pictorial representations. As regards the psychology of perception, "space-time ceases to be a pure given in order to become … the nexus of differential relations in the subject, and the object itself ceases to be an empirical given in order to become the product of these relations" (Deleuze 1993: 89) when unfolded and brought to consciousness, that is, actualized. The structure of the psyche does not contradict Lee Smolin's (2001) quantum account of the structure of space and time. It only makes us question whether we should continue positing psyche, in a Cartesian fashion, as a-dimensional and non-extended. Respectively, the quantum theory in its ontological interpretation (Bohm and Hiley 1993) posits the indivisible unity of the world, which is capable of being fully realized not as substantial but precisely as a relational or interactional system that continuously undergoes transformations between its various forms of manifestation.

When projected onto a pictorial spread, the virtual reality of signs undergoes transformations leading to a loss in dimensions at the level of our actual experience that "convey the projection, on external space, of internal spaces defined by 'hidden parameters' and variables or singularities of potential" (Deleuze 1993: 16). Hidden variables thus become exposed in practice: what was hiding in the depth of the psyche in the form of enfolded "ambiguous signs" (1993: 15) is literally brought to the surface and made available to consciousness nevertheless remaining deeply profound both conceptually and with respect to its informational content. It is because of the action of signs that "from virtuals we descend to actual states of affairs, and from states of affairs we ascend to virtuals, without being able to isolate one from the other" (Deleuze and Guattari 1994: 160). Top-down and bottom-up, again and again. It is "what is unseen [that] decides what happens in the seen" (Dewey 1998: 229): due to projection, we become able to actually *see* the compactified "scope of space and time that [becomes] accessible to observation" (Jantsch 1980: 4). The Magician performs the role of a hypothetical (in the framework of science) operator of projection when it enters into the "surface organization which assures the resonance of two series" (Deleuze 1990: 104), and the meaning created in practice is, paradoxically, even more "profound since it occurs at the surface" (1990: 10) in the form of projection of deep structures of the psyche. Mind embodied in matter extends itself both spatially and temporarily: consciousness in which the unconscious has been integrated "runs ahead and foresees outcomes, and thereby avoids having to await the instructions of actual failure and disaster" (Dewey 1922/1988: 133). This foreknowledge of the outcomes is the prerogative of Magicians, indeed!

The Magician is a symbol of *tertium quid* as the essence of semiosis, of the evolutionary dynamics of signs due to which they grow in meaning: "Essence is … the third term [which] complicates the sign and the meaning; it measures in each case their relation … the degree of their unity" (Deleuze 2000: 90), and it is the very "essence of the virtual to be actualized" (Deleuze 2003: 28). The Magician contains the conditions for unity, symbolized by the number 1, within itself. The recursive communicative feedback loops comprise the network of mutual interactions that establish a link between *res extensa* and *res cogitans*. As an unconscious idea implicit in the protomental nature, the Magician is virtually "extensive and enduring" (Dewey 1925/1958: 279), thus strongly defying the Cartesian postulate of mind as a non-extended substance. Creating a momentous "negentropy as semiotic information" (Spinks 1991: 71), the Magician is capable of trans-coding the analogue continuum of the universal *One* into the digital organization of *Many* particulars demonstrating as such the semiotic code-duality and hinting onto the solution to the continuous versus discrete problematic. Entropy as the invisible information is also the measure of ignorance (Lloyd 2006), and the boundary line separating the unseen invisible information from the visible depends, from the subjective viewpoint, on our own ignorance versus knowledge! The Magician performs what Jung, in relation to the psychology of the unconscious, called the transcendent function that he, in turn, derived from its definition in mathematics as grounded in complex (real and imaginary) numbers. It is the important function of "consciousness not only to recognize and assimilate the external world through the gateway of the senses, but to translate into visible reality the world within us" (Jung CW 8. 342). The relation between without and within:

> generates a tension charged with energy and creates a living, third thing—not a logical stillbirth in accordance with the principle *tertium non datur* but a movement out of the suspension between opposites, a living birth that leads to a new level of being, a new situation. The transcendent function manifests itself as a quality of conjoined opposites (Jung, CW 8. 189).

Tertium non datur is the excluded third, but depth psychology aims to achieve the reconciliation between the rigid opposites by means of the inclusion of "a third thing in which the opposites can unite … In nature the resolution of opposites is always an energic process: she acts *symbolically* in the truest sense of the word, doing something that expresses both sides" (Jung CW 14. 705). *Tertium non datur* thus becomes *tertium quid*—the included third: even if seemingly logically unclassifiable, it establishes a relation that connects or unites the perceived opposites. Such included middle "is not an average; it is fast motion, it is the absolute speed of movement. [It] is neither one nor two; … it is the in-between, the border or line of flight or descent running perpendicular to both" (Deleuze and Guattari 1987: 293). The Magician reconstructs the Neoplatonic Oneness by taking One out from the virtual realm (which is habitually considered supernatural, hence outside science) and bringing it down to earth and into the midst of the Many actual, flesh-and-blood, human experiences. Hence follows what Deleuze and Guattari (1987) called their mystical and magical formula expressed as One=Many.

Because the Magician's wand "reaches down into nature … it has breadth … to an indefinitely elastic extent. It stretches" (Dewey 1925/1958: 1). This stretch, as the new "magnitude of thirdness" (Deely 1990: 102), expands the event-horizon of knowledge because it "constitutes inference" (Dewey 1925/1958: 1) and contributes to the genesis of the fully fledged semiotic reason. A novel concept created by means of such a stretch—and effectuated as if by the "magic" wand of the Magician—has no reference outside itself. It is self-referential, just as genuine triadic signs, the logic of which appears paradoxical, if not totally magical, in the framework of dyadic, strictly two-valued, logic. It is "infinity [that] is self-referential" (Kauffman 1996: 293), and it is indeed a symbol of infinity crowning the Magician (Fig. 12.1) that indicates such self-referential, as though "magical," action.

While analytic reason denounces self-reference (dubbing such logic circular, hence begging the question), the action of signs is still "fundamentally linked to a logic: a logic of multiplicities" (Deleuze and Parnet 1987: viii) as tri-relative entities. In parallel to Peirce, Deleuze points out that "there are two in the second, to the point where there is a firstness in the secondness, and there are three in the third" (Deleuze 1989: 30). Taking two abstract terms A and B, Deleuze inserts the conjunction AND in between. Multiplicity contains an a-signifying rupture as difference—a pure relation, a gap—in which the conjunction AND intervenes in the mode of the included third: not in the opposition of A to B but "in their complementarity" (Deleuze and Parnet 1987: 131). The relational logic (semiotics) is not subordinate "to the verb to be. … Substitute the AND for IS. A *and* B. The AND is … the path of all relations, which makes relations shoot outside their terms" (1987: 57). The Magician's role is to be a Peircean interpretant that traverses the series symbolized by two "disparates," A and B. We therefore can construct another visual diagram displaying multiplicity as a genuine triadic sign where the *divergent*, heterogeneous series A and B *converge* on a paradoxical element symbolized by the AND that infinitely repeats itself in the evolutionary process of semiosis. The intensive, even if apparently "non-localizable" (Deleuze 1994: 83), conjunction AND forms a semiotic triangle (Fig. 12.3), and the Magician demonstrates the paradox of "the One-Whole of the Platonists" (Deleuze 1991: 93).

Notably, the conjunction AND foregrounding multiplicities as signs cannot be reduced to numerical addition. The process is of *summation* that, while suggesting a simple adding of information, in fact intensifies it by means of forming a logical product akin to multiplication and forming a power series. An intensified perception and understanding would have vitally increased in power, almost literally: there is an exponential growth there invoked yet by Peirce, but the Magician's transversal communication carries an exponent towards its limit as if graphically crossing the otherwise asymptotic line, thus reaching a threshold provided that the situation meets the conditions for actualization of the virtual. Meanings created in actuality are "effects that are not a mere dependence upon causes, but the occupation of a domain, the operation of a system of signs" (Deleuze and Guattari 1983: 86). The act of semiotic, transversal, communication confirms what Whitehead called the paradox of the connectedness of things (Whitehead 1966: 228). A semiotic triangle (as per Figs. 12.2 and 12.3) both *closes* on itself in the ternary structure and also *opens* itself to its becoming-other-than-itself because of novel meanings due to the inclusion of interpretants. Such is the paradox of self-reference (cf. Kauffman 2010; Kelso and Engstrøm 2006; Semetsky 2001b, 2001c) elicited by the logic of included middle peculiar to semiotics. A sequence of signs "adds up" to one enduring object. It is "when you invoke something transcendent [that] you arrest movement" (Deleuze 1995: 147), thereby demonstrating that a Tarot layout, conceptually, is like "any given multiplicity [that occupies] one area on the plane" (1995: 147) as the result of integration. In mathematics integration is represented by symbol $\int$ that represents the operation of summation symbolized by $\sum$.

The Magician creates the conditions for structural couplings defined as "a chain of interlocked … *communicative interactions*" (Varela 1979: 48f) embedded in the silent discourse of images. While the Arcanum that precedes the Magician and called the Fool (Fig. 12.4) conveys the image of facing the chaotic abyss with its unlimited potential, it is the Magician that brings order into the semiotic process because chaos as a source of potentially significant meanings is "seen as Creative" (Hoffmeyer and Emmeche 1991: 162).

In the Tarot deck the Fool's corresponding numeral is zero that appears to signify nothing (cf. Rotman 1987)—but not quite so. In fact, the presence of the Fool in each of the subsequently numbered Arcanum is a truism: 1 and 0 is still 1, 2 and 0 is still 2, and so forth. Partaking of the Deleuzian difference, imperceptible by itself, the Fool exemplifies zero-point energy, a quantum fluctuation (cf. Prigogine in Laszlo 1991) or pure information bordering on becoming active. Like an empty set $\varnothing$, an abstract entity of mathematical analysis that apparently signifies nothing, the Fool organizes meaning into what is intrinsically meaningless when it itself enters into relations following its symbolic leap into the abyss. Each whole number that indexes every one of the 22 major cards describes the property that contains zero in itself as an empty set. Each subsequent number can be marked off (or signed) by basic marks or braces .{}. This is the mathematical process of iteration during which the braces are repeated and "the empty set, {} … correspond[s] with zero; then 1 [becomes] the name of the property belonging to all sets *containing* the empty set, {}" (Noddings and Shore 1984: 51). The Fool plays the role, symbolically, of what Deleuze (1990) called an empty square; yet this emptiness or nothingness is what elicits the production of series, therefore becoming a precursor for putting them into

Fig. 12.3 Semiotic triangle

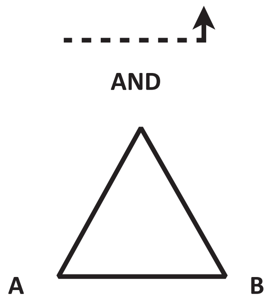


Fig. 12.4 The Fool

relation to each other "by virtue of its own power" (Deleuze 1994: 119). Such is the Fool's paradoxical semiotic significance even if signifying nothing!

According to Spencer-Brown's (1979) *Laws of Form*, logic is constructed in a repeating or replicating series on the basis of an initial act of making a distinction. In opposition to Russell, Spencer-Brown arithmeticized logic demonstrating its construction "from the basic intuitive act of making a distinction and two fundamental arithmetical acts: (1) making a mark to signify the distinction, and (2) repeating the mark" (Noddings and Shore 1984: 51) in agreement with Deleuze's (1994) fundamental philosophical conception of difference and repetition. Indeed it is "recurrence [that] makes novelty possible" (Dewey 1925/1958: 47). The unnumbered, and at first sight insignificant, Fool precedes the Magician that marks the distinction, thus becoming able to, quantum-mechanically, "create information out of nothing" (Lloyd 2006: 118), ex nihilo, the zero sign, the Fool. By virtue of its leap into the abyss, the Fool performs a quantum of action and thus initiates the evolutionary semiotic process. Surely, there "is nothing wrong with beginning from nothing. For example, the positive numbers begin from zero (the 'empty thing')" (2006: 45). Following the Fool's initial leap (Semetsky 2001a, 2005, 2013), it is the Magician that actively constructs logic as represented by multiple bracketing **{… {…}…}**, that is, making a difference and creating novel knowledge precisely due to repetition in the manner of the infinite series (Fig. 12.5).

There is "not merely 1, 2, 3, but 1, 2 in 2 and 1,2,3 in 3" (Deleuze 1986: 198). Each conjunction AND is the in-between relation that acts as a distributed marker

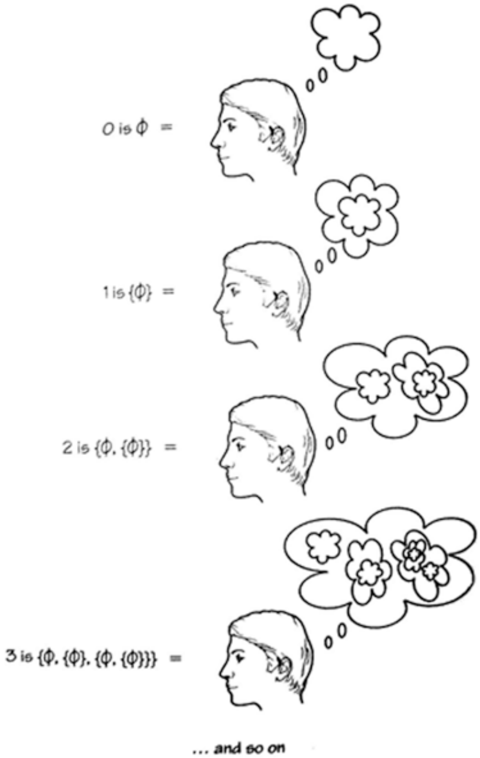


Fig. 12.5 Infinite series (Barrow 2000: 160; cf. Rucker 1982: 40)

of "a new threshold, a new direction of the zigzagging line, a new course for the border" (Deleuze 1995: 45). The Fool's zero then presents itself as "the germinal nothing … boundless possibility [and] boundless freedom" (Peirce CP 6. 217); this sense of freedom and infinite potential (cf. Peat 1997) becoming available for the Fool in its nonmetric world of topological space where void coincides with plenum and the dual opposites are in fact "inextricably connected to each other" (Kelso and Engstrøm 2006: 186). The mind is *in* the world, not outside of it. The action of signs is the "informationally meaningful, self-organizing coordination dynamics, a web~weaver" (2006: 253) where the tilde "~" is an index of a connective link weaved by the "magic" wand. The transition from the unconscious to consciousness (psychologically) or from virtual to actual (ontologically) indicates a "dynamic instability [that] provides a universal decision-making mechanism for switching between and selection of polarized states" (2006: 10), the latter functioning as "'attractors' of an underlying dynamical system" (2006: 10) of signs as patterns of coordinated activity. The Magician is immanent in matter in its capacity of a "virtual governor" (Juarrero 1999: 125), the function of which is nonlocal but distributed in the semiotic field in accordance with what Whitehead dubbed the fallacy of simple location. A newly created meaning is a new direction taken by means of the autocatalytic web built by the Magician's wand, a specific decision made because every actuality is "the decision amid 'potentiality' … The real internal constitution of an actual entity constitutes a decision conditioning the creativity which transcends the actuality" (Whitehead 1978: 93). In making a decision, the Magician may very well employ abduction that "comes to us as a flash. It is an act of insight" (Peirce CP 5. 181), or intuition, or imagination functioning analogous to a certain "automatism [as] the psychic mechanism of perception" (Deleuze 1993: 90).

Coordination dynamics exhibits tendencies as "preferences and dispositions" (Kelso and Engstrøm 2006: 10) between actual stable states, while in the metastable regime there are no states but relations. Mind-in-the world is a semiotic system that parallels the Deleuzian *fold* as a holistic structure held together by the tendency to couple or bind together as indeed exhibited by the Magician. The dynamic unity can be diagrammed as an oscillation between the opposites when the movement is projected onto a line, and the dynamics of semiosis is precisely what unites the opposite poles. Semiosis is not only the action but also transformation of signs, and self-reference is ultimately self-transcendence as "the creative overcoming of the status quo" (Jantsch 1980: 91). Such fecund empiricism is always already transcendental, and it is signs and "symbols [that] act as transformers" (Jung CW 5. 344). Mindful nature is by default complementary.

The founders of the science of coordination dynamics Kelso and Engstrøm (2006) reflect on their "fascination with what seemed at first a somewhat esoteric connection between philosophy and the science of coordination" (Kelso and Engstrøm 2006: xiii). Indeed! They notice that despite nature being described by quantum laws that sure allow complementarity between two seemingly mutually exclusive descriptions, our everyday practical experience habitually chooses between one true or right description and another false or wrong, hence ignoring the "shades of grey" (2006: xi) between them. The ubiquitous form of coordination dynamics (2006: 156) is presented as a symbolic equation cv-dot = f (cv, cp, F), where cv-dot is the rate of change of the coordination variable expressing the evolution of dynamic patterns (see also Kelso 1995), **cp** stands for the control parameter(s), and **F** is a chance fluctuation (perhaps, the very action of the Fool?). Thus the Magician's transversal link across "the brain~mind and brain~behavior barriers" (Kelso and Engstrøm 2006: 9) is not mystical but perfectly natural and it seems that surely, "mathematically speaking, [it] contains three different kinds of parameters" (2006: 157): the strength, or intensity, of coupling; the presence of intrinsic differences (not unlike in Deleuze's theoretical vision); and the always already present fluctuations or "noise." For Leibniz and Deleuze, such background noise would be composed of the unconscious or *little* perceptions as infinitesimal differentials.

At the symposium on developmental science in Stockholm in 1998, Kelso coined the principle of the *in-between* as the new scientific (rather than purely speculative) concept. The repudiation of the either-or mentality brought:

> a novel scientific grounding to age-old questions that all of us ask: Which is more fundamental, nature or nurture, body or mind, whole or part, individual or collective? … a great deal of the core essence of such dichotomized aspects seems to be located … in what Aristotle called the "excluded middle" … *The Complementary Nature* introduces a new

> meaning and application of the tilde or "squiggle" character ~, as in yin~yang, body~mind … Unlike the hyphen, the squiggle does not represent a simple concatenation of words, but … indicates the inextricable complementarity relation between them (Kelso and Engstrøm 2006: xiv–xv).

The tilde character "~" is a symbol for the abovementioned oscillation—or the Magician's very nature. The Magician's action can be expressed not just in *signa data* but in *signa naturalia* manifesting itself as the universal principle of the in-between. Such action is, however, implicit: it may be a hidden variable waiting to be discovered so as to take its place among the natural laws described in the language of mathematical physics meanwhile presenting itself, in psychological terms, as the unconscious "noise" striving to enter cognition. Kelso and Engstrøm (2006) point to some important nuances: while the laws of coordination, like physical laws in general, are matter independent, they are nonetheless function and context dependent; they govern and therefore make relatively predictable "*the flow of functional information*" (2006: 100). Information, albeit preserved, is being reorganized and redistributed. It becomes meaningful, functional, or active (cf. Bohm 1980), that is, capable of producing real effects in accord with Peirce's pragmatic maxim. This means that by practically stepping into the flow of semiosis—of which we, theoretically, as signs among signs are a constituent part anyway—we, by virtue of reading and interpreting signs represented by Tarot icons, become able to exercise a degree of predictability within each specific context. In agreement with the so-called triangle argument constructed on the basis of Einstein's relativity theory (Fig. 12.6), "me-now" can become simultaneous with "me-tomorrow" in practice, at the level of empirical reality:

The dotted lines indicate simultaneity; simultaneity implies coexistence; and the coexistence relation is indicated by the two-headed arrow, not unlike the double-directedness inscribed in the imagery of the Magician. The ancient law of analogies as applied to space—*as above so below*—has its temporal correlate in the *Emerald*

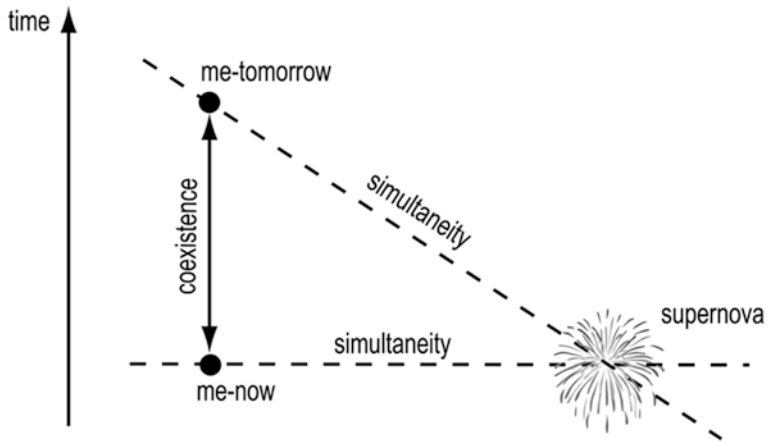


Fig. 12.6 The triangle argument (Kennedy 2003: 63)

Tablet expressed as follows: *That which was is as that which will be, and that which will be is as that which was.* Once again, the Hermetic philosophy which is even today considered mystical and magical manifests its uncanny affinity with the developments in science. The infinitely distant "supernova" may be considered conceptually equivalent to the vanishing point—a zero at infinity—in a perspectival composition. The Magician immersed in semiosis enables a specific, pointed onto by Whitehead, organization of thought that makes precognition possible. The Magician's creative wand establishes directedness, that is, "a vector [that] already indicates in which direction the new structure may be expected" (Jantsch 1980: 46)[1], and such action "terminates in a modification of the objective order, in the institution of a new object … It involves a dissolution of old objects and a forming of new ones in a medium … beyond the old object and not yet in a new one" (Dewey 1925/1958: 220), but within Leibniz's zone of indiscernibility between the two.

Novelty, as a change in a system's behavior, is described in nonlinear mathematics as a phase transition. Such ubiquitous state is far from equilibrium, yet it provides an initial impulse to the unfolding dynamics of signs. Semiosis does presuppose an initial condition of "uneasy or unstable equilibrium" (Dewey 1925/1958: 253) symbolized by the Fool tip-toeing at the edge of the abyss. It is when "frozen in their locations in space and time" (Kennedy 2003: 53) that past, present, and future events symbolized by the Tarot pictures in a specific layout demonstrate their coexistence quite in accord with the block-universe view of relativity theory. In the layout of pictures, the signs' diachronic dimension becomes compactified into a single synchronic slice when the dynamical process of semiosis is projected, that is, momentarily *frozen in its location in space-time* because of the quality of relatedness functioning in accord with the rules of projective geometry. Synchronization is but an example of self-organized coordination (Kelso 1995), and the Magician, by exercising the coordination dynamics, exhibits the semiotic value of the ultimate growth in knowledge and understanding.

To conclude, let us address one more "mystery" presented by the relation between the three worlds, namely physical, mental, and Platonic (see Penrose 1997, 2004). Because the Platonic world is inhabited by mathematical truths, but also due to the

---

[1] In the earlier publication titled "Interpreting Peirce's abduction through the lens of mathematics" (Semetsky 2015) I suggested a vectorial diagram on the complex (Gaussian) plane as a model for knowledge structure incorporating abduction as an unconscious inference. Peirce called such a mode of thought instinctive reason. The Magician's semiotic reason can be modeled by means of geometry on the complex plane using imaginary numbers—dubbed *magical* by physicist and mathematician Sir Roger Penrose (2004)—coupled with real and together forming complex numbers. The imaginary number $\boldsymbol{i}$ as a square root of minus 1 does "appear to play a fundamental role in the working of the universe" (Penrose 2004: 67) including, as implied by the Whiteheadian *one* world without and within, the working of the human mind. Leibniz called them amphibian: in-between being and nothingness. As Lou Kauffman points out, it is "remarkable that domains imaginary with respect to arithmetic are virtually real with respect to geometry" (1996: 293). Raising a complex number to the $\boldsymbol{n}$-th power multiplies its angle by $\boldsymbol{n}$. It was Riemann who merged projective geometry with the idea of complex numbers. On the Riemann's "number sphere" zero and infinity are but two opposite poles. In quantum mechanics, zero (vacuum) is a source of infinite energy.

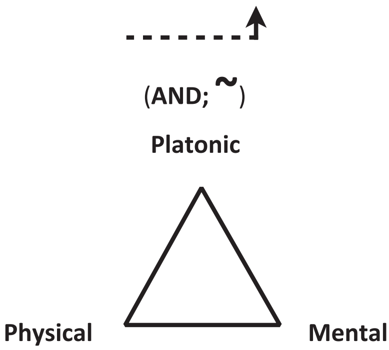


Fig. 12.7 Another semiotic triangle

"common feeling that these mathematical constructions are products of our mentality" (Penrose 1997: 96), the dependence of the natural world on strict mathematical laws appears mysterious. But it seems that we can take away the flavor of mysticism pertaining to this relation if we consider it properly semiotic and construct yet another semiotic triangle, thus confirming, once again, the presence of the Magician functioning in the manner of a "squiggle" or Deleuze's conjunction AND (Fig. 12.7).

In the language of coordination dynamics, physical and mental worlds form a complementary "body~mind" pair in accordance with the logic of included middle, thus confirming an assertion that there exists a "part of the Platonic world which encompasses our physical world" (Penrose 1997: 97). As such, it is when projected onto the physical level, that is, compactified, that the Platonic ideas can become "accessible by our mentality" (1997: 97) in the manner of another "the unconscious~consciousness" complementary pair. The rules of projective geometry establish mapping as the one-to-one correspondence—like in a perspectival composition towards a vanishing point—thus implying isomorphism between the archetypal ideas of the Platonic world and the *coupled together* mental and physical worlds. We can conceptualize a semiotic triangle in terms of such a composition, however with a shifting frame of reference or point of view. If and when a vanishing point shifts into the mental world, this leads to isomorphism between a mental representation and the other two worlds: the world of ideas coupled with the physical world of our actions and behaviors. In fact the very quality of this point being _"vanishing" makes such composition somewhat a-perspectival (cf. Gebser 1991), especially from the viewpoint of the Magician per se.

There is another nuance here. The Magician's unusual, or virtual, logic (cf. Kauffman 1996, 2010) "energizes reason [and] provides the real possibility and the means for opening of communication across boundaries long thought to be impenetrable" (Kauffman 1996: 293). Such semiotic reason transcends narrow rationality and reaches "into a world of beauty, communication and possibility" (1996: 293) while going beyond given facts into a world of interpretable signs, meanings, and values. What inhabits the Platonic world is not only the True but also the Good and the Beautiful that appear to be "non-computable elements—for example, judgement, common sense, insight, aesthetic sensibility, compassion, morality" (Penrose 1997: 125) as the attributes of the psyche. Does it mean that the Magician as symbolic of an expansive mode of thought that integrates the unconscious is capable of paradoxically *computing the apparently incomputable*?

References

Barrow, J. D. (2000). The book of nothing. New York: Vintage Books.

Bertalanffy, v. L. (1972). Foreword. In: E. Laszlo (ed.), Introduction to systems philosophy: toward a new paradigm of contemporary thought, pp. xvii-xxi. New York: Gordon and Breach Science Publishers.

Bohm, D. (1980). Wholeness and the implicate order. London and New York: Routledge.

Bohm, D., and Hiley, B. J. (1993). The undivided universe: An ontological interpretation of quantum theory. London: Routledge.

Cilliers, P. (1998). Complexity and postmodernism: Understanding complex systems. London; New York: Routledge.

Deely, J. (1990). Basics of semiotics. Bloomington: Indiana University Press.

Deely, J. (2001). Four ages of understanding: The first postmodern survey of philosophy from ancient times to the turn of the twenty-first century. Toronto: University of Toronto Press.

Deleuze, G. (1986). Cinema 1: The movement-image (H. Tomlinson and B. Habberjam, Trans.). Minneapolis: University of Minnesota.

Deleuze, G. (1988). Foucault (S. Hand, Trans.). Minneapolis: University of Minnesota Press.

Deleuze, G. (1989). Cinema 2: The time-image (H. Tomlinson and R. Galeta, Trans.). Minneapolis: University of Minnesota.

Deleuze, G. (1990). The logic of sense (M. Lester and C. J. Stivale, Trans.). New York: Columbia University Press.

Deleuze, G. (1991). Bergsonism (H. Tomlinson, Trans.). New York: Zone Books.

Deleuze, G. (1993). The fold: Leibniz and the Baroque (T. Conley, Trans.). Minneapolis: University of Minnesota Press.

Deleuze, G. (1994). Difference and repetition (P. Patton, Trans.). New York: Columbia University Press.

Deleuze, G. (1995). Negotiations, 1972–1990 (M. Joughin, Trans.). New York: Columbia University Press.

Deleuze, G. (2000). Proust and signs (R. Howard, Trans.). Minneapolis: University of Minnesota Press.

Deleuze, G. (2003). Desert islands and other texts (1953–1974). New York: Semiotext(e).

Deleuze, G. (2007). Mathesis, science and philosophy. In: R. Mackay (ed.), Collapse III, pp. 141-155. Falmouth: Urbanomic.

Deleuze, G., and Guattari, F. (1983). Anti-Oedipus: Capitalism and schizophrenia (R. Hurley, M. Seem and H. R. Lane, Trans.). Minneapolis: University of Minnesota Press.

Deleuze, G., and Guattari, F. (1987). A thousand plateaus: Capitalism and schizophrenia (B. Massumi, Trans.). Minneapolis: University of Minnesota Press.

Deleuze, G., and Guattari, F. (1994). What is philosophy? (H. Tomlinson and G. Burchell, Trans.). New York: Columbia University Press.

Deleuze, G., and Parnet, C. (1987). Dialogues (H. Tomlinson and B. Habberjam, Trans.). New York: Columbia University Press.

Dewey, J. (1922/1988). Human nature and conduct. In: J. A. Boydston (ed.), The middle works of John Dewey, 1899–1924 (Vol. 14). Carbondale and Edwardsville: Southern Illinois University Press.

Dewey, J. (1925/1958). Experience and nature. New York: Dover Publications.

Dewey, J. (1998). Time and individuality. In: D. Browning and W. T. Myers (eds.), Philosophers of process, pp. 211-226. New York: Fordham University Press.
Gebser, J. (1991). The ever-present origin (N. Barstad and A. Mickunis, Trans.). Athens, OH: Ohio University Press.
Griffin, D. R. (1986). Physics and the ultimate significance of time: Bohm, Prigogine, and process philosophy. Albany: State University of New York Press.
Guattari, F. (1995). Chaosmosis: An ethico-aesthetic paradigm (P. Bains and J. Pefanis, Trans.). Bloomington and Indianapolis: Indiana University Press.
Hoffmeyer, J., and Emmeche, C. (1991). Code-duality and the semiotics of nature. In: M. Anderson and F. Merrell (eds.), On Semiotic Modeling, pp. 117-166. Berlin, New York: Mouton De Gruyter.
Jantsch, E. (1980). The self-organizing universe: Scientific and human implications of the emerging paradigm of evolution. Oxford, New York: Pergamon Press.
Juarrero, A. (1999). Dynamics in action: Intentional behavior as a complex system. Cambridge, MA: The MIT Press.
Jung, C. G. (1953–1979). Collected works (R. F. C. Hull, Trans.). Princeton University Press: Princeton, NJ. [cited as CW].
Kauffman, L. H. (1996). Virtual logic. Systems Research 13(3): 293-310.
Kauffman, L. H. (2010). Virtual logic—an introduction. In: I. Semetsky (ed.), Semiotics education experience, pp. 175-202. Rotterdam: Sense Publishers.
Kelso, J. A. S. (1995). Dynamic patterns: The self-organization of brain and behavior. Cambridge, MA: The MIT Press.
Kelso, J. A. S., and Engstrøm, D. A. (2006). The complementary nature. Cambridge, MA: The MIT Press.
Kennedy, J. B. (2003). Space, time and Einstein: An introduction. Chesham, UK: Acumen.
Laszlo, E. (1991). The age of bifurcation: Understanding the changing world. Philadelphia: Gordon and Breach.
Lloyd, S. (2006). Programming the universe: A quantum computer scientist takes on the cosmos. New York: Alfred A. Knopf.
Merrell, F. (1996). Signs grow: Semiosis and life processes. Toronto: University of Toronto Press.
Noddings, N., and Shore, P. J. (1984). Awakening the inner eye: Intuition in education. New York and London: Teachers College, Columbia University.
Nöth, W. (1995). Handbook of semiotics. Bloomington: Indiana University Press.
Peat, F. D. (1997). Infinite potential: The life and times of David Bohm. Reading, MA: Addison-Wesley Publishing Company.
Peirce, C. S. (1931–1935; 1958). Collected papers of Charles Sanders Peirce (Vol. 1–6 ed. C. Hartshorne and P. Weiss; Vol. 7 and 8 ed. A. Burks). Cambridge, MA: Harvard University Press. [cited as CP].
Penrose, R. (1997). The large, the small, and the human mind. Cambridge, UK: Cambridge University Press.
Penrose, R. (2004). The road to reality: A complete guide to the laws of the universe. London: Jonathan Cape.
Rockwell, W. T. (2007). Neither brain nor ghost: A nondualist alternative to the mind-brain identity theory. Cambridge, MA: The MIT Press.
Rotman, B. (1987). Signifying nothing: The semiotics of zero. Stanford, CA: Stanford University Press.
Rucker, R. (1982). Infinity and the mind: The science and philosophy of the infinite: Birkhauser.
Semetsky, I. (2001a). The adventures of a postmodern fool, or the semiotics of learning. In: C. W. Spinks (ed.), Trickster and ambivalence: The dance of differentiation, pp. 57-70. Madison, WI: Atwood Pub.
Semetsky, I. (2001b). Self-organization in Tarot semiotics. In: W. Schmitz (ed.), Sign processes in complex systems: Proceedings of the 7th International Congress of the International Association of Semiotic Studies, pp. 11-28. Dresden: Thelem.

Semetsky, I. (2001c). Signs in action: Tarot as a self-organized system. Cybernetics and Human Knowing: A Journal of Second-Order Cybernetics, Autopoiesis and Cyber-Semiotics. Special issue: Peirce and Spencer-Brown 8(1–2): 111-132.
Semetsky, I. (2005). The phenomenology of Tarot, or: The further adventures of a postmodern fool. Trickster's Way 4(1): 3.
Semetsky, I. (2008). The transversal communication, or: reconciling science and magic. Cybernetics and Human Knowing: A Journal of Second-Order Cybernetics, Autopoiesis and Cyber-Semiotics 15(2): 33-48.
Semetsky, I. (2011). Re-symbolization of the self: Human development and Tarot hermeneutic. Rotterdam: Sense Publishers.
Semetsky, I. (2013). The edusemiotics of images: Essays on the art~science of Tarot. Rotterdam: Sense Publishers.
Semetsky, I. (2015). Interpreting Peirce's abduction through the lens of mathematics. In: M. Bockarova, M. Danesi, D. Martinovic and R. Núñez (eds.), Mind in mathematics: Essays on mathematical cognition and mathematical method, pp. 154-166. Munich: Lincom.
Shimony, A. (1993). Search for a naturalistic world view: Scientific method and epistemology (Vol. I–II): Cambridge: Cambridge University Press.
Smolin, L. (2001). Three roads to quantum gravity. New York: Basic Books.
Spencer-Brown, G. (1979). Laws of form. New York: E. P. Dutton.
Spinks, C. W. (1991). Semiosis, marginal signs, and trickster: A dagger of the mind (1st ed.). London: Macmillan Academic and Professional Ltd.
Stapp, H. P. (2007). Mindful universe: Quantum mechanics and the participating observer. Berlin: Springer-Verlag.
Varela, F. J. (1979). Principles of biological autonomy. New York: North Holland.
Whitehead, A. N. (1966). Modes of thought. New York: Free Press.
Whitehead, A. N. (1978). *Process and reality: An essay in cosmology* (Corrected ed.). D. R. Griffin and D. W. Sherburne (eds.). New York: Free Press.

Chapter 13
Geometric Cognition

Walter Whiteley

Introduction

My chapter title asserts that "*mathematical cognition*" must include "*geometric cognition"* either as a stand-alone process or in a cognitive blend with other forms of mathematical cognition. I aim to shift our gaze to geometric cognition—or the more general equivalent "*spatial/visual reasoning in mathematics*." Of course, the spatial representations and visual representations can be experienced across all parts of mathematics and statistics. Related cognition is found across all of the sciences and engineering and for centuries mathematics and sciences were part of the same community, with shared cognitive processes (SIGGRAPH 2002). However, geometry offers the clearest and often unavoidable expression of this aspect of mathematical cognition.

I am a geometer. I apply geometry in my funded applied mathematics research across a range of problems in multiple disciplines in science and engineering. I have also been teaching geometry to third-year mathematics majors, many of whom are preparing to teach high school. Beginning with my Ph.D. Thesis on the logical foundations of discrete geometry (*invariant theory* in the language of the nineteenth century), and then my growing collaborations in discrete applied geometry, my life has been immersed in geometry for about 50 years. That immersion means I have been working with spatial reasoning as my sources of insight; my reasoning; my sharing of mathematics; my teaching; and my communication of results across multiple disciplines. Reflections evolved within this immersion has also encouraged my research, collaborations, and writing on spatial reasoning within mathematics education.

As I tell my undergraduate geometry class: "I see geometry everywhere, and want you to share that experience. Learn to recognize where a geometry lens can provide surprising insights." By the end of the yearlong course, most students report that they

W. Whiteley (✉)
Mathemtics and Statistics, York University, Toronto, ON, Canada
e-mail: whiteley@mathstat.yorku.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_13

do not quite see geometry everywhere—but far more places than they ever had before. They also say that this class is "not like any other math class they have taken—in a good way." The class is filled with mathematics explored through manipulatives, drawing, and spatial reasoning, in multiple representations. For examples of these activities, see the text I used in this course for several decades (Henderson and Taiminia 2004). As the student end-of-year reflections confirm, such a hands-on spatial/kinesthetic exploratory approach is rare as a focus within university mathematics courses—and can be missed as a conscious focus of mathematical processes. The unusual approach is appreciated by the students as challenging and sometime altering their sense of what mathematics is—or at least can be.

This chapter draws on decades of presentations and workshops with students, teachers, mathematics educators, and also conversations with a range of collaborators (Whiteley 1999, 2002, 2005, 2010, 2012, 2014, 2019). The feedback from these discussions and collaborations has been invaluable to my evolving reflections on all these issues.

Geometric Cognition and Spatial Reasoning

There are so many exciting and significant experiences which connect spatial reasoning or geometric cognition to mathematical cognition. As an initial focus on geometric cognition, including the related spatial cognition, I will have to be selective: picking a few illustrative examples. For a broader survey of the *Big Ideas and Procedures in Geometry* where I find geometric reasoning, see Whiteley (2019).

What are some key features of geometric cognition and spatial reasoning as I have lived them in research, teaching, and learning? I offer a condensed list:

(a) Transformations to support evolving questions, conjectures, and geometric reasoning
(b) Symmetry, and invariance, as core concepts in many areas (including across physics and geometry)
(c) Shifting dimensions: using 3D to understand 2D
(d) Multiple representations with cognitive blending: switching among representations—learning to see and recognize switching

Let me offer a few examples of how and where these features appear in geometric cognition.

(a) The modern definition of geometry is Klein's Erlanger Program of 1872 (Klein's Erlanger Program 1872, Klein 1924): we have a space of objects and a group of transformation of that space. The study of that geometry is the study of properties which are invariant (unchanged) under these transformations. This is illustrated in Fig. 13.1, with one classical strand of geometries, ordered upwards by inclusion of their expanding groups of transformations.

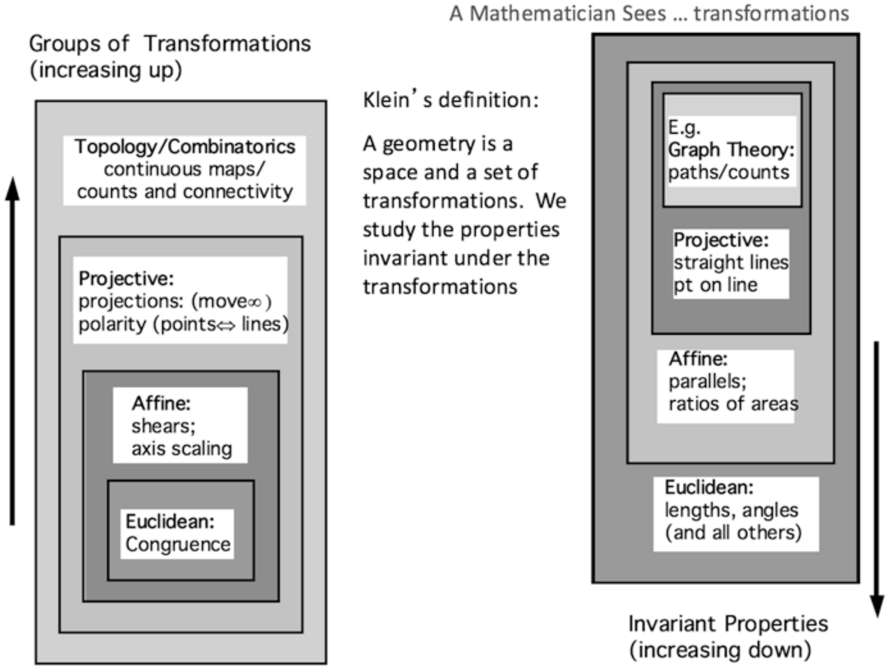


Fig. 13.1 Transformations

In my own problem-solving with discrete applied geometry, I begin the study of a problem with the key question (unusual among applied mathematicians): "which geometry should I use?" I explore which transformations leave the solutions to the problem invariant. I have learned the importance of this question by observing (1) cases where the problem was cast at too high a level (too many transformations) so that no coherent answer is possible, and (2) other cases where the problem is cast at too low a level in the hierarchy—and there are too many properties and details which are not relevant and the important patterns within the "forest" are lost among the "trees." For example, the static rigidity of spatial frameworks is not topological (too many transformations which lose key properties), but requires some further level of geometry. On the other hand, Euclidean geometry of distances is too low (too many irrelevant details) for statics. Static rigidity belongs to projective geometry, though most modern North American trained structural or mechanical engineers do not know those transformations (Schulze and Whiteley 2018). With more available projective transformations, many "different examples" are now recognized as "the same" under the transformations and one can focus on some key geometric properties and corresponding projective methods.

This grounding in transformations, embedded in a hierarchy of groups and subgroups, offers practitioners (researchers and students) thinking tools. Given an example, we play around transforming it and seeing what else is "the same."

We find some surprises and make new conjectures. We focus on what is invariant and find representations that “forget” the details that are changing and which bury the appropriate information.

A series of geometry problem-solving books (written for Russian high school students) explore this hierarchy with effective problems, figures, and solutions (Yaglom I–IV). Part I works with the geometry of rigid motions of the plane (isometries). Part II uses the geometry of shape-preserving transformations of the plane (similarities). Part III focuses on the geometry of transformations of the plane that map lines to lines (affine and projective transformations) and introduces the Klein model of non-Euclidean geometry. Part IV focuses on conformal mappings that take circles to circles. The introduction to Part III gives a nice introduction to the hierarchy, which I regularly used with future teachers. Two striking images adapted from the book are recalled by students many years later (Fig. 13.2).

(b) Symmetry offers a rich and engaging playground for thinking with transformations. Look at the subgroup of transformations which leave a specific object or example unchanged (invariant)—for example the symmetries of a platonic solid, or of a quadrilateral (Whiteley and Paksu 2015). When we find two symmetries, we should look for the composition of the two—filling out the table of group multiplication for the symmetries. We notice that two mirror reflections compose to form a rotation, and a rotation and a reflection compose to form another reflection. I know from students and teachers that these simple connections are lying around, underdeveloped, within the elementary curriculum as well as the university curriculum. Adding geometric thinking tools boosts the interest and richer connections of symmetry for further learning.

(c) Children live in 3D, but the Western math (and science) curriculum gives an early focus to 2D. It is a big shift for students to return to 3D—often needing to overcome weakness to survive in engineering and science (engage) and to thrive in mathematics. Even in my research in the rigidity of frameworks, playing among dimensions revived valued techniques, such as reworking the reciprocal

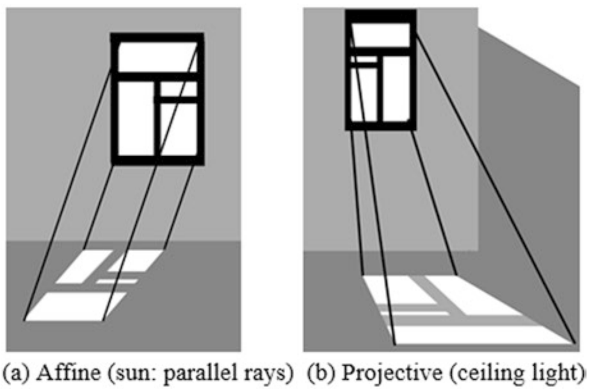


Fig. 13.2 Affine and projective transformations. (**a**) Affine (sun: parallel rays). (**b**) Projective (ceiling light)

diagrams of James Clerk Maxwell which connect the geometry of 3D and spherical frameworks with the statics and rigidity of plane frameworks (Schulze and Whiteley 2018).

(d) For the geometric/spatial thinker, transformations, including symmetries, offer a bundle of connected embodied representations: visual, spatial, kinesthetic, as well as shared patterns of group operations. These concepts form a network of embedded experiences which invite conceptual blending—and together become more richly cognitively linked as we move from one representation to another, and one brain network to another. Rapid switching among brain areas (and representations) is typical of top students, in brain scans around age 14.

These are just a few samples of the *Big Ideas in Geometry and Geometric Cognition* and themes from my decades of talks (Whiteley 2019). We will see below how some of these themes are connected both to historical developments and to how we and our students learn to reason.

Blends with Geometric Cognition, Spatial Reasoning

I am confirming that mathematical cognition is a blend which includes geometric cognition and spatial cognition, along with other patterns of mathematical and scientific reasoning (Fauconnier and Turner 2002; Turner 2014).

It has been distracting that some earlier researchers in education, such as Howard Gardner in Project Zero (Gardner 1985), separated "mathematical intelligence" as "logical/symbolic intelligence" from "spatial/visual intelligence" (see Fig. 13.3). All the recent research confirms that for children, and for many practicing mathematicians,

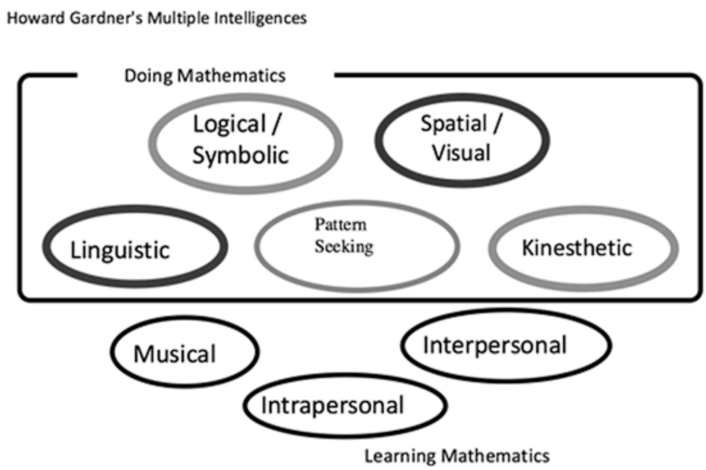


Fig. 13.3 Multiple intelligences

mathematical cognition includes spatial cognition, as well as Gardner's "kinesthetic intelligence." These multiple intelligences of Gardner are commonly discussed in Faculties of Education (Gardner 1985, 2006) to open up diverse options for teaching and learning. This is a valuable emphasis for teachers, but this makes it more important that they have an inclusive vision and wide sample of experiences of mathematical intelligences (see Going Forward section). In my experiences in teaching, learning, and practicing mathematics, we blend from all of the identified intelligences in the dark box below (a graphic I use in my class with future teachers). Gardner's recent addition of "pattern seeking" as an eighth intelligence is easily recognized as at the very core of our activities as mathematicians. Pattern seeking is something we ask all students to practice and develop in mathematics classrooms and well beyond.

Often the use of physical manipulatives, supporting kinesthetic reasoning, is used in close association (a blend) with visual spatial reasoning. For example, consider the back-and-forth process combining dynamic geometry (a tool developed for teaching, but now used in geometry research), with paper folding, which is deeply geometric and kinesthetic (Whiteley and Paksu 2015).

One of the images I use to help future teachers notice the back-and-forth switching involved in geometric problem-solving (and noticed in brain scans) is the zigzag in Fig. 13.4. This figure images the mental shifts from whole to part and parts back into the whole which is required in multistep problem-solving. This type of shift of focus within geometric reasoning is often hidden from students, as the teacher's blackboard notes and gestures focus primarily on the details of "right side" (b). This leaves unexplained jumps for students to puzzle out, as indicated in Fig. 13.4b. We may not publically share the larger processes illustrated in (a)—or even be consciously aware of them to raise them up for the students to reflect on.

One recent exploration of how blending is core to mathematical modeling appears in Whiteley (2012). This analysis reflects on the classroom spatial reasoning/geometric optimization popcorn box activity described for elementary teachers

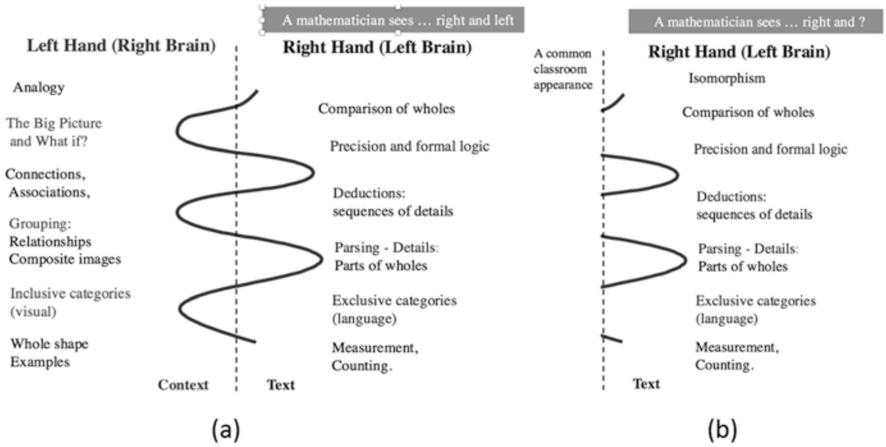


Fig. 13.4 Mental shifts

in Mamalo and Whiteley (2012). This student activity was further analyzed as “a network of and for geometric reasoning” which is explicitly connected to students developing cognitive blends in Mamalo et al. (2015).

As proposed in this analysis (Mamalo and Whiteley 2012; Mamalo et al. 2015), the blend is developed through a sequence of back-and-forth simulations of “generic examples”—examples for which the reasoning does not rely on specific details but is generalizable over a wide range of variations (Mason and Pimm 1984). This “seeing the general within a particular example” is characteristic of a lot of diagrammatic reasoning (reasoning with diagrams) and of reasoning with manipulatives. Grade 4 and 5 students could identify that key choices, such as changing the scale of the model had no impact on the “shape optimum box” (using proportional reasoning)—in the context of multiple physical models. When the activity was done with in-service teachers, they could also explore the analogies with a corresponding 2D problem and some could even explore 4D versions of the problem. Such “generic reasoning” was well developed in centuries of careful geometric practice during the centuries after Euclid. It is however a cultural practice that must be learned, and if this geometric reasoning culture is missing in the classroom or in the visible shared practices, this support for geometric cognition risks being lost to the next generation, along with other ways of doing mathematics (Whiteley 2010).

Several recent papers speak to neural reuse for conceptual mappings, and the reuse of spatial brain maps for memories, ideas, and reasoning about time (Cánovas and Monzanares 2014; Copelewicz 2019). This overlap within neural representations reinforces the claim that these spatial metaphors are fruitful for blending and that spatial reasoning is widely used beyond just for “space.”

Recognizing the Importance and Centrality of Spatial Reasoning

For centuries, geometry was central to mathematics and practitioners integrated diagrams in shared mathematical practices (some of which have been lost). For example, one historical reading of the 13 books of Euclid’s Elements is that the entire program is centered on symmetry, and the books all build, with strong use of diagrams, to the 3D theorem characterizing the five Platonic solids. This suggests that the Elements are based on geometric reasoning as spatial reasoning. The common focus on teaching this “elementary” geometry through “formal—symbolic proofs” ignores these aspects of spatial geometry—to the loss of students and teachers (Tall et al. 2012; Whiteley 1999, 2010). We return to this theme in Development of Geometric Reasoning and Going Forward sections below.

A striking and well-documented example of effective geometric cognition is the work of Michael Faraday on electromagnetism (West 2009; Goodings 2006). Faraday’s 1821 notebooks for the day he built the first electric motor displays his problem-solving in an essentially visual/diagrammatic form. The notebooks record

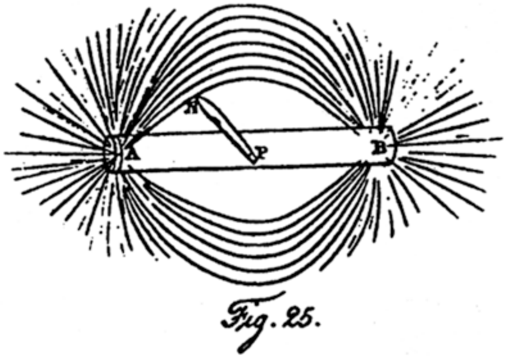


Fig. 13.5 Faraday's lines of force

a back and forth between (1) the laboratory experiments (recorded with a diagram and then a sequence of observations recording changes as more diagrams), and (2) planning using some doodling followed by a new diagram for the next experiment. In his work, Faraday invented some new diagrammatic forms to capture his quantitative reasoning with figures, as illustrated in the now standard diagram of "lines of force" (Fig. 13.5).

All the historical evidence is that Faraday did not, and could not, work with equations (West 2009). Faraday worked with what I am calling geometric cognition. Nevertheless, he reasoned out results such as Faraday's laws of electromagnetism which we now write with equations. This way of reasoning was recognized as "mathematical cognition" by one of the great figures of this field—James Clerk Maxwell: "As I proceeded with the study of Faraday, I perceived that his method of conceiving phenomena was also a mathematical one, though not exhibited in the conventional form of symbols" (James Clerk Maxwell, as quoted in West 2009).

Maxwell wrote that this use of lines of forces shows Faraday "to have been in reality a mathematician of a very high order—one from whom the mathematicians of the future may derive valuable and fertile methods." I agree that there are fertile methods here, but unfortunately, if current students rely on such visual methods, today's schooling often identifies them as failing mathematics—and not suitable to become engineers.

In the mid-twentieth century, the eminent mathematician Jacques Hadamard interviewed a number of leading mathematicians of his generation asking them to describe how they "invented" their mathematics (Hadamard 1945). A key observation was that they wrestled with a problem (exploring the pieces that might become a blend) and then set the problem aside. Their first conscious awareness of the new insight of how to solve the problem was in visual (spatial) form. The recognition of what I call "a cognitive fitting together of pieces towards the solution with the problem" was in the visual form. This provided a platform to support further exploration and re-presentation in a problem-solving blend of connections.

George Polya is widely recognized as a combinatorist, as a student of mathematical problem-solving and as a mathematics expositor and teacher (Polya 1954a, 1954b). In his review of "Plausible Reasoning," the mathematician Paul Halmos (1955) summarized the central thesis in this way: "a good guess is as important as a good proof." To me, this is capturing the insights of visual reasoning—which are prior to the formal, logical, and symbolic reworking of the problem. In his expository paper "On Picture Writing" (Polya 1956) Polya makes visible a good sample of his reasoning when shifting from diagrammatic representations of a counting problem to a symbolic algebraic generating function. The paper contains two, full-page, samples with a sequence of "equivalent representations" of the problem, one line each. Each step records a shift of both the representation and the associated operations, in reversible steps. It is a wonderful expression which makes visible to all of us the otherwise invisible reasoning of this master problem solver—something we do not encounter often enough.

While not the reflections of a mathematician, Temple Grandin expresses well what working primarily with spatial reasoning is like in her autobiographical book "*Thinking in Pictures*" (Grandin 2006). This book echoes stories of other historical figures who relied on vivid and effective spatial reasoning, such as Nicholas Tesla, and Faraday (West 2009). This way of working becomes at least one option within a broader blended mathematical cognition which includes geometric cognition. The very ways we share our work in publications overemphasizing words, symbols, and formulas (which are easier to put down on the page) gives priority to later formal reasoning over also presenting the geometric reasoning which were the basis for our insights. We have limited tools for sharing spatial reasoning, and we often lack enough shared conventions for sharing spatial/visual reasoning.

Burton (2004) describes a more recent study in which she interviewed a wide range of researchers in mathematics and statistics about how they did their work. One of the themes was the wide range of approaches, including analytic, conceptual, and visual thinking. Visual thinking was documented as central by some participants, and an important option among several by others. Insights from self-reflection are necessary sources, as just reading the published articles gives a skewed impression. As mentioned above, mathematicians often select the analytic (algebraic, computational) presentation, in preference to the more difficult-to-present visual/geometric presentation. As I referee research articles, I often recommend more figures and more examples. To paraphrase the responses of some colleagues to why they do not include more figures: "when I read an article, I draw my own pictures—doesn't everyone"? Again, this self-selection is rendering the spatial/visual basis for the work invisible.

Over the centuries, there has been changing emphasis on spatial reasoning. In Whiteley (1999), I draw on experiences in my community to propose a narrative for how geometry faded in North America, particularly in the second half of the twentieth century. This culture shift contributed to the breaking of continuity of geometric practices based on the central role of spatial/reasoning which was obvious in earlier periods, and has almost been lost by the twenty-first century. What I also claim, drawing on evidence from the curriculum in both undergraduate and graduate programs across disciplines, and evidence in current scientific research, is that geometry still remains essential to solving problems in many areas of applied mathematics. However, this geometry may now only surface in other disciplines, when the required geometric cognition is not supported within pure and applied mathematics programs.

In my own research and teaching experience, spatial/visual reasoning is more salient in applied mathematics than in pure mathematics. It was also more salient in mathematics prior to the twentieth century (Tall et al. 2012; Whiteley 1999). Visual presentation of examples and results still remains the standard for sharing mathematical and statistical reasoning across multiple disciplines. For sample resources which support communication: see Howard Wainer's *Visual Revelations*, and *Graphic Discovery, Picturing an Uncertain World* (Wainer 2000, 2007), as well as the discussion and appendices in the White Paper: *Visual Learning for Science and Engineering* (SIGGRAPH 2002).

In talks and workshops, I often present examples where the spatial reasoning becomes visible "with eye and hand." In "*The Case for Mental Imagery*" Kosslyn et al. (2006) addresses the controversy within cognitive science and philosophy of whether images (and therefore spatial reasoning) are actually found in our internal cognition. Kosslyn presents strong evidence that what we call images are also present in the brain—and we think with and operate on these mental images as we do on external images. Spatial reasoning can be done with our eyes closed and our hands not moving—including operations like mental rotation. For example, mental rotations develop early for children (e.g., when learning infant sign language) and this ability remains an important spatial reasoning skill which continues to be tested through to mechanical reasoning tests. Mental rotation regularly occurs entrance tests for medical and dental school (Davis et al. 2015). Weak abilities in spatial reasoning become a negative filter for many careers—so developing such abilities or reasoning is an important challenge.

Development of Geometric Reasoning

There is now a wide recognition of the key role of geometric cognition in the larger development of mathematical cognition, at least for young children. There is a large literature on using spatial reasoning in the learning of mathematics (see the multiple chapters in Davis et al. (2015)), where our spatial reasoning group reviewed a range of the literature and described examples. A recent book with the title "*Visualizing Mathematics; The Role of Spatial Reasoning in Mathematical Thought*" also describes the key role of geometric cognition (Mix and Battista 2018). I will not repeat the references and links from these books.

Children are born into space, learning to see and to move, even from before birth. By the time they enter school, they have learned varying amounts of 3D spatial reasoning—depending on the activities they did, and what the adults around them direct their attention to, in part through use of spatial language (Davis et al. 2015). One of the chapters this book specifically explores the connections between 2D and 3D reasoning. Burke et al. (2017) make the connection of 3D work to embodied cognition—a connection also found in Davis et al. (2015) and implied by my earlier phrase of "working with eye and hand." Unfortunately, this prior knowledge of 3D is often neglected in early schooling. I sometimes say that we take the students who have lived in 3-space, and then "flatten their reasoning into the plane" as they enter grade 1!

In *Visual Intelligence: How We Create What We See* (Hoffman 2000), Donald Hoffman speaks of how we can change what we see, based on experience. As the neurologist Oliver Sacks has said, "when we open our eyes each morning, it is upon a world we have spent a life-time learning to see." This means that what I see is not the same as what you see—and we can change what we see. This is true for how we see geometry, and more generally what we notice in mathematics and statistics (Whiteley 2005, 2012, 2014). In an analogy to Betty Edwards' insights into learning to draw by first learning to see (Edwards (1999), I claim that for many students "learning to see like a mathematician" opens a new door to success in mathematics and statistics (Whiteley 2005, 2012, 2014). Supporting this type of brain changing learning is also explored in the appendices of the White Paper (SIGGRAPH 2002).

Froebel, the inventor kindergarten, began his sequence of activities (Gifts) with a series of 3D spatial activities using rich precursors of our now simplified building blocks (and now somewhat richer Lego) (Brosterman 1997). Drawing on his prior work in crystallography, and in hands on learning, Froebel engaged children in 3D activities and the very name "kindergarten" reminds us that each child had a small garden plot (Brosterman 1997)! The activities included symmetry and transformations, still common in much of the hands on work with manipulatives. It does not seem to be a coincidence that Frank Lloyd Wright's mother was a Froebel Kindergarten teacher—and he continued to play with spatial visual cognition in his design work (Brosterman 1997). Later gifts in Froebel's sequence included 2D activities, still with manipulatives, and often with symmetry.

As Davis et al. (2015), Burke et al. (2017), as well as Froebel notice, 3D comes before 2D for children. This is a critical concern identified by a number of mathematics educators. The eminent Canadian geometer Donald Coxeter contributed to the draft of a rich *Geometry K-13 OISE Report* (Geometry 1967) which started with 3D, and with visual reasoning, and built from there. These far-sighted curriculum drafters included activities such as work with vectors in grades 4–5. This is the age at which children are actually working with maps and compasses to navigate in their world outside of school (a basic geometric task). Sadly, this curriculum was never implemented. A few years ago when I proposed vectors as sample activity to some curriculum writers, there was full agreement that children could do this mathematical activity at age 10. The barriers to enriching the curriculum were (1) that the current teachers probably could not handle it, and (2) such a spatial theme did contribute immediately to the otherwise "calculation and algebra based curriculum" which was driven to get students "ready for calculus" by the end of high school. The general result of these obstacles in North America is an impoverished exposure in school to only a few aspects of mathematical cognition and effective problem-solving, with little spatial/visual reasoning!

The mathematics educator Jo Boaler highlights the damaging myth in mathematics education that visual/spatial reasoning (with hands to count, and diagrams to reason with) is only for young students, weak students, or lay people (nonmathematicians) (Boaler et al. 2016). She notes that West (2009) describes this myth as centuries old—and that there is compelling brain science to contradict this myth. The examples and studies in this section also contradict this myth. In our terms, expert brains contain the key blends that retain these visual/spatial connections, even when we appear to be working with another, more abstract level of the blend. We continue present school mathematics with only one part of the rich, thick cognitive network of concepts and representations, weakening any support for spatial cognition, and for students who depend on this part of the blend.

To return to Klein's Hierarchy from Geometric Cognition and Spatial Reasoning section, Piaget's trajectories for children's stages of learning geometry fit well with moving down Klein's Hierarchy. Simple topology is explored early (George 2017)—with questions about "what is connected to what" as the child explores: "what can I reach"? "Can I go out one door of a room and come back through another door?" Then learning extends to straight lines: anticipating where a toy train that entered a tunnel will emerge. What is the shortest path to follow—all part of (projective geometry). The last concepts mastered by children in schools are concepts like volume and area (Euclidean), around ages 8–10 (grade 4).

While children learn spatial reasoning early, they can, and often do, lose these abilities during puberty and high school. In some longitudinal studies following improved spatial reasoning from piano training developed at age 4, it was found that the improved spatial reasoning was scrambled to just chance during puberty. (See Rauscher et al. (1997) and their following studies.) Other studies confirm a decline of 3D spatial reasoning during high school years, when the curriculum does not practice spatial reasoning or make evident to students that spatial reasoning is a critical prerequisite skill for many university programs and careers. These students graduate to face the shock of the clear necessity of spatial reasoning to succeed in their chosen programs across engineering, sciences, and mathematics.

We offer a quote from work on the urgency of developing Spatial Reasoning for Engineering (Engage 2019): "Most engineering faculty have highly developed 3D spatial skills and may not understand that others can struggle with a topic they find so easy. Furthermore, they may not believe that spatial skills can be improved through practice, falsely believing that this particular skill is one that a person is either 'born with' or not. They don't understand that they probably developed these skills over many years. We don't encourage students not ready for calculus to enroll in calculus in their first semester. Shouldn't spatial skills training be available for those who need the help?" [Sheryl Sorby, quoted at ENGAGE: Spatial Reasoning for Engineering (2019)].

There is convincing evidence that spatial reasoning is malleable for people over a range of ages from zero through middle age (Davis et al. 2015; Uttal et al. 2013). My experience, in multiple classes and workshops over the last 30 years, is that students, including future teachers, are concerned about their weakness in spatial reasoning—and they are very encouraged to learn that they can still continue to improve spatial skills. The same observation is true for classes and workshops with in-service teachers, as they recognize they can then both be better in their own mathematics, and can better support students who may be primarily approaching mathematics (and other subjects) through spatial/visual reasoning.

For undergraduate mathematics majors, this encounter with their gap spatial reasoning often happens, even to high-achieving mathematics students, in their third course in calculus: multivariable calculus. The ENGAGE website and Sorby (2019) speak to both the importance of spatial reasoning and describe short programs (10 h) which (re)-build spatial reasoning for retention and success in key programs such as engineering. The gap in support for this critical skill continues into many university programs—as spatial reasoning becomes a filter for who will remain in programs, rather than another important skill to be developed.

Papers such as those of Mann (2005) and Silverman (1995) describe gifted students who rely on spatial reasoning and are underserved and under-supported in programs and classrooms which omit spatial reasoning. These papers also describe effective classroom techniques to support such students. Faraday would have been such an underserved student—and would not have survived our schooling to enter an engineering program in our century. The losses for mathematics and sciences and our larger society for undervaluing geometric cognition are large.

Going Forward

I highlight two promising themes as directions for future work on geometric cognition.

Aging and Losing Our Mathematical Minds

A lot of research has focused on how we learn different mathematical abilities when we are young or even middle aged. At the other end of life, during aging, we may experience critical stages of cognitive decline. In particular, various mathematical abilities are lost, most visibly in neurodegenerative diseases such as Alzheimer's, Parkinson's, and more generally dementia (Possin 2010). This survey includes an exploration of different facets of visual/spatial cognition—in terms of brain areas and networks that can be disrupted in the changing brain. The different forms of loss in spatial reasoning are diagnostic of different neurodegenerative diseases and different forms of dementia. These losses in spatial cognition are often documented with tests such as the *Montreal Cognitive Assessment* (MoCA) which has a surprising focus on spatial reasoning. Overall, the study of loss of spatial reasoning promises additional insights into how spatial reasoning is processed in our brains, at all ages.

The loss of spatial reasoning, such as navigation and spatial (distance) perception, can have a big impact on people, including the loss of a driver's license and associated independence. I conjecture that the loss of spatial reasoning is also related to other losses in aging—such as a sense of which day it is (the loss of a mental calendar as spatially organized), and what the daily schedule is (the timeline as a spatial orientation). Given how much spatial ability contributes to learning arithmetic, one has to also wonder whether the loss of spatial reasoning contributes to decline in the ability to calculate or do proportional reasoning (e.g., figure out a tip in the restaurant).

With the research emphasis on how mathematical cognition is developed, and our aging population, it would be timely to study how mathematical cognition declines. "Losing our mathematical minds" is an important life stage. Finding strategies for maintaining and supporting spatial cognition should be a parallel to finding ways to maintain language abilities. The losses are commonly identified in the current tests, and some of the computer-based programs being tested are spatially focused. However I have not seen a good survey of support systems for geometric cognition, or more generally for all forms of mathematical cognition, as we age.

More generally, changes in aging brain pose the question of understanding how established cognitive blends are broken, and perhaps how to thicken the connections within a blended network, so that the blend is more resilient as we age. A "good-enough blend" may not last as we age and change. We may also blur across some blends and incorporate other connections which are not helpful, as the brain ages. We "learn how to see"—and we can also lose the ability to perceive connections and metaphors well. As I write this, I recognize I am showing my age and describing my current community!

Geometrizing/Spatializing the Curriculum

Davis et al. (2015) and Boaler et al. (2016) present the overall challenge of spatializing the curriculum, and enriching it with visual activities—a challenge that applies to all stages of education. This can only happen when many people decide geometric cognition or spatial/visual skills are important. For the future, we depend on a shared recognition of how this spatializing of the curriculum will strengthen the mathematical cognition of all students, and the urgent recognition that failure to include this spatial content means the exclusion of many students. The diversity of cognitive styles and prior knowledge among students excludes many who could be major contributors in a more supportive classroom context.

An essential part of such an enrichment is educating a generation of teachers who themselves have experienced the value of spatial/visual reasoning, and therefore are eager to incorporate these activities in their classrooms. In my own classrooms and within my wider community of university geometry teachers, Henderson's "Experiencing Geometry" has provided such support for decades of future teachers (Henderson and Taiminia 2004). When the future, or current, teachers reflect on how they have been learning mathematics with eye and hand, they consistently want their own students to experience that support and engage in spatializing their classrooms.

I am hopeful of a future with a rising of geometry and support for geometric cognition during the twenty-first century

References

Boaler, J. Chen L., Williams, C., Cordero, M. (2016). Seeing as understanding: The importance of visual mathematics for our brain and learning. Journal of Applied Computational Mathematics 5: 325.

Brosterman, N. (1997). Inventing kindergarten. New York: Harry N. Abrams.

Burke, H., Gardony, A., Hutton, A., and Taylor, H.. Thinking 3d! Improving mathematics learning through embodied spatial training; Cognitive Research: Principles and Implications (2017), 2: 1-18.

Burton, L. (2004). Mathematicians as enquirers: Learning about learning mathematics. New York: Springer.

Cánovas, C. P. and Monzanares, J. V. (2014). Conceptual mappings and neural reuse. Frontiers in Human Neuroscience 29 April 2014 doi: https://doi.org/10.3389/fnhum.2014.00261

Copelewicz, J. (2019). The brain maps out ideas and memories like spaces. *Quanta Magazine*. https://www.quantamagazine.org/the-brain-maps-out-ideas-and-memories-like-spaces-20190114/

Davis, B. et al (2015). Spatial reasoning in the early years: Principles, assertions, and speculations. New York: Routledge.

Edwards, B. (1999). Drawing on the right side of the brain. New York: Tarcher.

Engage (2019). Spatial reasoning for engineering (accessed January 2019) https://www.engageengineering.org/spatial/whyitworks/

Fauconnier, G. and Turner, M. (2002). The way we think: Conceptual blending and the mind's hidden complexities. New York: Basic Books.

Gardner, H. (1985). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Gardner, H. (2006). Multiple intelligences: New horizons in theory and practice. New York: Basic Books.

Geometry (1967). *Geometry K-13 OISE Report 1967* www.math.yorku.ca/~whiteley/geometry.pdf Posted with permission of the Ontario Institute for Studies in Education of the University of Toronto.

George, W. (2017). Bringing van Hiele and Piaget together: A case for topology in early mathematics learning. Journal of Humanistic Mathematics 7: 105-116.

Goodings, D. C. (2006). From phenomenology to field theory: Faraday's visual reasoning. Perspectives on Science 14: 40-65.

Grandin, T. (2006). Thinking with pictures: My life as an autistic. London: Bloomsbury.

Hadamard, J. (1945). The psychology of invention in the mathematical field. New York: Dover.

Halmos, P. (1955). Review of Polya. Bulletin of the American Mathematical Society 61: 243-245.

Henderson, D. and Taiminia, D. (2004). Experiencing geometry: Euclidean and non-Euclidean geometry with history. Boston: Pearson.

Hoffman, D. (2000). Visual intelligence: How we create what we see. New York: W. W. Norton and Co.

Kosslyn, S., Thompson, W., Ganis, G. et al (2006). The case for mental imagery. Oxford: Oxford University Press.

Klein, F. (1924) Elementary mathematics from an advanced standpoint: Geometry. New York: Dover.

Klein's Erlanger Program (1872) https://en.wikipedia.org/wiki/Erlangen_program.

Mamalo, A., Ruttenburg-Rozen, R., Whiteley, W. (2015). Developing a network of and for geometric reasoning. ZDM: International Journal on Mathematics Education 47: 483-496.

Mamalo, A. and Whiteley, W. (2012). The popcorn box activity and reasoning about optimization. Mathematics Teacher 105(6): 420-426.

Mann, R. (2005) Gifted students with spatial strengths and sequential weaknesses: An overlooked and under-identified population. Roeper Review 2005.

Mason, J. and Pimm, D. (1984). Generic examples: Seeing the general in the particular. Educational Studies in Mathematics 15: 277-289.

Maxwell, J. Clerk (2003). The scientific papers of James Clerk Maxwell. New York: Dover.

Mix, K. and Battista, M. (eds.) (2018). Visualizing mathematics; The role of spatial reasoning in mathematical thought. New York: Springer.

Polya, G. (1954a). Mathematics and plausible reasoning. Volume I: Induction and analogy in mathematics. Princeton: Princeton University Press.

Polya, G. (1954b). Mathematics and plausible reasoning. Volume II: Patterns of plausible inference. Princeton: Princeton University Press.

Polya, G. (1956). On picture writing. American Mathematical Monthly 63: 687-697.

Possin, K. L. (2010). Visual spatial cognition in neurodegenerative disease. Neurocase 16(6): 466-487.

Rauscher, F. H. et al (1997). Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. Neurology Research 19(1): 2-8.

Schulze, B. and Whiteley, W. (2018). Rigidity and scene analysis. In: C. Toth, J. Goodman and J. O'Rourke (eds.), Handbook of Discrete and Computational Geometry.

SIGGRAPH (2002). White paper: Visual learning for science and engineering. http://education.siggraph.org/conferences/other/visual-learning

Silverman, L. K. (1995) Effective techniques for teaching highly gifted visual-spatial learners. https://eric.ed.gov/?id=ED418535 (with full text).

Sorby, S. (2019). Higher education services: Visualizing success https://www.higheredservices.org/ (accessed Jan 2019). See also the TedX talk: https://www.youtube.com/watch?v=cJZIhl28HFI

Tall, D. et al (2012). Cognitive development and proof. *Proof and Proving in Mathematics Education,* Michael de Villiers and Gila Hanna (eds.), April 2012.

Turner, M. (2014). The origin of ideas: Blending, creativity, and the human spark. New York: Oxford University Press.

Uttal, D. H. et al (2013) The malleability of spatial skills: A meta-analysis of training studies. Psychological Bulletin 139: 352-402.

West, T. (2009). In the mind's eye: Creative visual thinkers, gifted people with dyslexia, and the rise of visual technologies. New York: Random House.

Wainer, H. (2007). Graphic discovery: A trout in the milk and other visual adventures. Princeton: Princeton University Press.

Wainer, H. (2000) Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot. New York: Psychology Press.

Whiteley, W. (1999). The decline and rise of geometry in 20th century North America. *Proceedings of the 1999 CMESG Conference*. www.math.yorku.ca/~whiteley/cmesg.pdf.

Whiteley, W. (2002). Teaching to see like a mathematician www.math.yorku.ca/~whiteley/Teaching_to_see.pdf.

Whiteley, W. (2005). Learning to see Like a Mathematician. In: G. Malcom (ed.), Multidisciplinary approaches to visual representation and interpretation, pp. 279-292. Oxford: Elsevier.

Whiteley, W. (2010). As geometry is lost—What connections are lost? What reasoning is lost? What students are lost? Does it matter? *Plenary Talks PIMS Changing the Culture* 2010. https://www.pims.math.ca/files/AsGeometryIsLost_0.pdf.

Whiteley, W. (2012). Mathematical modeling as conceptual blending: Exploring an example within mathematics education. In: M. Bockarova, M. Danesi, and R. Núñez (eds.), Cognitive science and interdisciplinary approaches to mathematical cognition. München: Lincom Europa.

Whiteley, W. (2014). Seeing like a mathematician: There is a diversity of ways to support the active learning of mathematics, Paper and Presentation at Forum for Action: Effective Practices in Mathematics Education' https://mathforum1314.wordpress.com

Whiteley, W. (2019). *Big ideas in geometry*, http://wiki.math.yorku.ca/index.php/Big_Ideas_Concepts_Procedures (accessed January 2019).

Whiteley, W. and Paksu, A. D. (2015). *Reasoning with quadrilaterals using hands, eyes, and technology*, Workshop, Ontario Association of Mathematics Educators Annual Conference, May 2015 https://www.researchgate.net/publication/322505303_Reasoning_with_Quadrilaterals_using_hands_eyes_and_technology.

Yaglom, I. M. (I–IV): *Geometric transformations* MAA New Mathematical Library Vol 8, 21, 24, 44; 1962, 1968, 1973, 2009.

Chapter 14
Using Evidence to Close the Achievement Gap in Math

John Mighton

Introduction

I'm a professional mathematician and a playwright, but I didn't show much aptitude for math or creative writing until I was an adult. When I was growing up I sometimes struggled with math at school and in middle school I began to wonder if I would be able to learn the subject at an advanced level. Because I had doubts about my abilities, I became obsessed with my intellectual capabilities and with the way I learn. When I started to teach in my 20s, first as a graduate student in philosophy and later as a math tutor, I also became fascinated with the way other people learn. Now, after founding a charity that develops mathematical resources for students (JUMP Math) and teaching thousands of learners of all ages, I am convinced that our society vastly underestimates the intellectual potential of children and adults. And my conviction appears to be well supported by evidence from many fields.

A wide body of recent research, from early childhood development to neurology, suggests that mathematics should be accessible to almost any student. For example, research has shown that for young children the strongest predictors of later achievement in math involve skills and concepts that every person will almost certainly develop, no matter how much they struggle in math in their early years or how delayed they are in acquiring these skills. These indicators involve very simple tasks that humans have evolved to perform with relatively little instruction, including the task of counting to ten or the task of correctly associating a numerical symbol (1, 2, 3, and so on) with a quantity (for example an array of dots) or a position on a number line. The fact that these basic capacities are such strong predictors of later achievement suggests that it is not a lack of innate ability that prevents the majority of students from becoming proficient at mathematics, but rather something that happens to learners in the course of their education.

J. Mighton (✉)
Fields Institute for Research in Mathematical Sciences, Toronto, ON, Canada
e-mail: jmighton@fields.utoronto.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_14

Over a hundred years ago, logicians investigating the logical foundations of mathematics showed that virtually all of mathematics can be reduced to trivial logical steps that draw on the same basic skills and concepts that predict success in math (such as counting or grouping objects into sets). As well, brain scans of mathematicians have revealed that working mathematicians tend to process math in parts of the brain that evolved to perform very basic functions. As David Tall explains in his paper in this collection:

> when Almeric and Dehaene studied mathematicians working in very different research areas (abstract algebra, analysis, geometry, topology), they found that all of them activated areas of the brain related to spatial sense and number which are present in young children before they develop language and are also found in many other non-human species.

To understand why so many people struggle in math, when evidence from so many fields points to the fact that should be easy to learn, we need to look closely at the systemic problems that make it so hard for teachers to nurture the full potential of their students. Until recently, very few of the math programs and resources that are used across North America have been tested in rigorous scientific studies, so educators and parents have often been seduced into adopting instructional approaches that sound very progressive and appealing, but that lack any strong empirical evidence.

For example, when parents, teachers, or administrators choose resources for students, they will try to find materials that they believe are engaging or interesting for children. But the choices they make are rarely informed by rigorous research. In 2013 psychologists Jennifer Kaminski and Vladimir Sloutsky of Ohio State University taught two groups of primary students to read bar graphs using two different types of graph: one had pictures of stacked shoes or flowers in the bars of the graph, and the other, more abstract graph had solid bars. The researchers asked teachers to say which kind of graph they would use with their students. The significant majority chose the graphs with pictures because they were more engaging and represented the objects in the problem. However, the study showed that students learned better from the gray monochrome bars. Students who learned with the bars were better at reading graphs when the scale of the graph changed to reflect some multiple of the number of objects. Students taught with pictures tended to be distracted by counting the objects and so did not look at the scale on the graph.

A great deal of recent research suggests that popular approaches to teaching, including "reform"- or "discovery"-based teaching (in which students are expected to develop their own explanations and approaches to math, often working with other students with little guidance from the teacher), haven't worked for most students because they don't take account of the limited capacity of the brain to absorb and process new information.

In addition to overwhelming students with extraneous information, discovery methods also burden them with too much material at once. Because of this heavy "cognitive load," pure discovery-based strategies do not appear to work as well as those in which a teacher helps a student navigate the complexities of a problem. According to Paul

Kirschner of the Open University of the Netherlands, 50 years of research on learning has consistently shown that instructional approaches where a teacher actively guides the students' learning are more or less effective and efficient than approaches where the teacher provides minimal guidance. In a 2011 review of 164 studies of discovery-based learning psychologist Louis Alfieri of City University of New York and his colleagues concluded: "Unassisted discovery does not benefit learners, whereas feedback, worked examples, scaffolding and elicited explanations do."

An Approach to Problem-Solving

In most math textbooks today there are more words than numbers. The books are typically full of "word problems" that try to make math relevant by asking students to apply their mathematical knowledge in real-world contexts. On state and provincial exams, these are usually the problems that separate students who do well in math from the ones who don't.

Sometimes students struggle with word problems because the text is too hard for them. But even when this isn't the case, students may still have trouble seeing the mathematical structure that is buried under the words. Teachers often try to help students who struggle with word problems by giving them more word problems. This remedy can have the same effect as pouring gas on a fire—it reinforces a student's sense of failure and makes harder for them to develop the confidence and ability to focus they will need to solve the problems.

In grades two and three, students sometimes struggle with word problems that involve a collection of things (or a "whole") that is comprised of two different kinds of things, or two "parts." If we are told that Bob has 4 marbles and Alice has 3 marbles, it's not hard to see that we can add 4 plus 3 to find how many marbles they have altogether. But if we are told how many more marbles one person has, the problem becomes more difficult. Some teachers like to tell students to look for key words in a problem and that when they see the word "more" it means they need to add to find the answer. But this isn't always the case. If Bob has 6 marbles and Alice has 2 more marbles than Bob, we would add 6 plus 2 to find out how many marbles Alice has. But if Bob has 6 marbles and has 2 more marbles than Alice, we would subtract 2 from 6 to find out how many marbles Alice has.

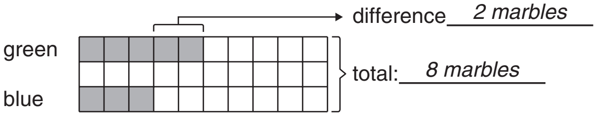
When students are required to balance the cognitive demands of reading a series of word problems—where the vocabulary and context may change from problem to problem—with the demands of recognizing what problem type they are given, they can easily suffer from cognitive overload. The more elements of a problem a teacher varies at one time the more likely it is that they will leave students behind.

One way to address this problem is to allow students to practice finding solutions to various types of part-whole problems in a series of exercises where only the problem type varies but where the numbers are small and the language is minimal and doesn't change. Rather than being asked to read whole paragraphs, about animals, then cars, and then vegetables, students might be given short phrases that are always about, for example, green and blue marbles. They might also be given the easiest problem type (where you are given the part and the part) first and then progress to harder types.

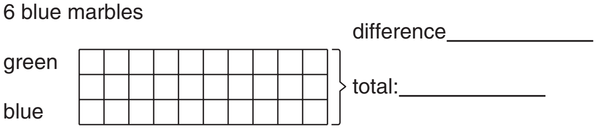
The questions below show the easiest problem type: the part and the part.

1. Shade boxes to show the number of marbles. Then find the total and the difference.

a) 5 green marbles
3 blue marbles



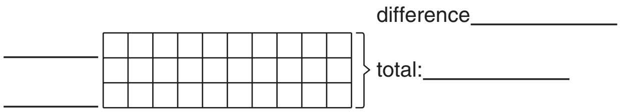
b) 4 green marbles
6 blue marbles



Students might also be allowed to practice solving each type as many times as they need to, so they understand one type before they are introduced to the next type.

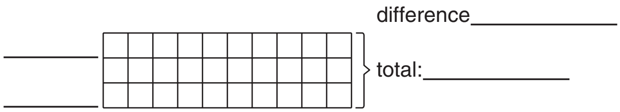
The questions below show the other problem types.

b) 6 marbles altogether
2 green marbles



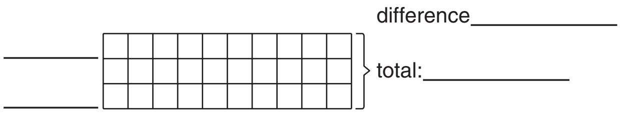
The part and the whole.

c) 3 green marbles
4 more blue marbles than green marbles



The part and the difference: When you have the smaller part and you know how many more.

d) 8 green marbles
3 fewer blue marbles than green marbles



The part and the difference when you know the bigger part and how many fewer.

Once students have mastered the various types of part-whole problem individually, they will still need practice recognizing the various types when they are presented at random. Switching between different problem types is not an easy task for some children, even when they have mastered each type separately. If the problems are presented in full paragraphs, it can interfere with the acquisition of this skill. So the teacher could put the information for each problem into a chart, as shown below.

	Green Marbles	Blue Marbles	Total	Difference
a)	3	5	8	*2 more blue marbles than green*
b)	2	9		
c)	4		6	
d)		2	7	
e)	6		10	
f)	3			1 more blue marble than green
g)		2		1 more green marble than blue
h)		4		1 more blue marble than green

Providing information in a table.

For students who are capable of handling bigger numbers, the teacher could include in the charts some numbers that are larger than the number of squares in their grids. This will stretch students a little, by forcing them either to draw their own sketch or to rely on their knowledge of numbers to find the answer mentally. The research on deliberate practice suggests that students learn most efficiently when they are continually pushed a little outside of their comfort zone but not too far.

Sometimes students who have been taken through this series of exercises will still panic when they see a word problem presented in a full paragraph. They will revert to guessing answers, even if they were able to solve the problems when they were presented with minimal language. To break this guessing reflex, one of the JUMP writers, Anna Klebanov, came up with an ingenious solution. She put fragments of word problems (in which marbles were replaced by fish, so students could get used to changing contexts) on the left-hand side of the chart. Rather than asking students to write the answer to the problem, she asked them to fill in all of the missing pieces of information in the chart and then circle the answer. Students are forced to develop a complete (mental or physical) picture of the situation before they are allowed to answer the question. This stops the student from guessing.

		Red	Green	Total	Difference
a)	Kate has 3 green fish and 4 red fish. How many fish does she have altogether?	*4*	*3*	*(7)*	*1*
b)	Bill has 4 green fish and 6 red fish. How many fish does he have altogether?				
c)	Mary has 8 green fish and 2 more green fish than red fish. How many fish does she have?				
d)	Peter ha 19 fish. He has 15 green fish. How many red fish does he have?				
e)	Hanna has 8 green fish and 3 fewer red fish than green fish. How many fish does she have?				
f)	Ken has 22 red fish and 33 green fish. How many more green fish does he have?				

After assigning the kinds of exercise I described above, the teacher can introduce problems with more text where the context varies from problem to problem.

The most challenging type of part/whole problem is one where the student is given the total and the difference. For example: You have 20 marbles. You have four more green marbles than blue marbles. How many blue marbles do you have? This is a perfect bonus question for students who are ready to be stretched a little further.

Only the most challenged student would encounter any real barriers of pitfalls in the progression of exercises I just outlined. When I teach in this way, I usually find that all of my students can move at roughly the same pace and no one is left behind. I can also cover material fairly quickly, because students are engaged and their brains aren't being overloaded. I always have a stock of incrementally harder bonus questions, so no one is bored. If students are ready to be stretched more, I can skip steps and let them struggle more. Although I haven't had the opportunity to test this series of exercises in a rigorous study, I predict that this approach would yield better results than an approach that involves giving students full word problems at the beginning of instruction.

When the language is stripped away from part-whole problems, you can see how easy the math is. You only have to know the two parts, the total and the difference, in such problems in order to know everything. Once you are able to form a mental representation of the situation (for example by visualizing two bars representing the parts) you have plumbed the depths of the mathematics. There is no hidden mystery here that only the most brilliant people can understand. Fortunately, this is the case with all of the mathematics students are required to learn at school: it can all be taught in a series of steps that almost anyone can understand. That's because the underlying structure of the math is invariably simple and accessible to virtually anyone, as long as that structure isn't obscured by language or the learning is not made difficult by too many cognitive demands imposed at the same time.

Avoiding Cognitive Overload

In the last section, I sketched an approach to problem-solving that reduces the learner's "cognitive load" by limiting the amount of extraneous verbal and visual information that the learner is required to process and by providing simple, semi-abstract models that help the learner see the deeper underlying structure of problems. It is also possible to design lessons that provide students with the scaffolding and practice they need to understand and perform the various mathematical procedures or algorithms they learn at school, such as rounding, addition, and subtraction.

When teachers teach mathematical procedures they sometimes teach more than one step at once without noticing how different those steps might seem to a novice. For example, when teachers teach addition with two-digit numbers, they often show students how to add the ones and the tens at the same time, without allowing them to practice each step in isolation. This is fine when the question doesn't involve regrouping (or "carrying" as regrouping used to be called) because in both the ones and the tens place you just add the bottom digit to the top digit. But when addition involves regrouping, the methods students use to add the ones are very different from the methods they use to add the tens. If the sum of the numbers in the ones place is greater than 9, students have to "carry" or "regroup" a ten. Students often make the mistakes shown below:

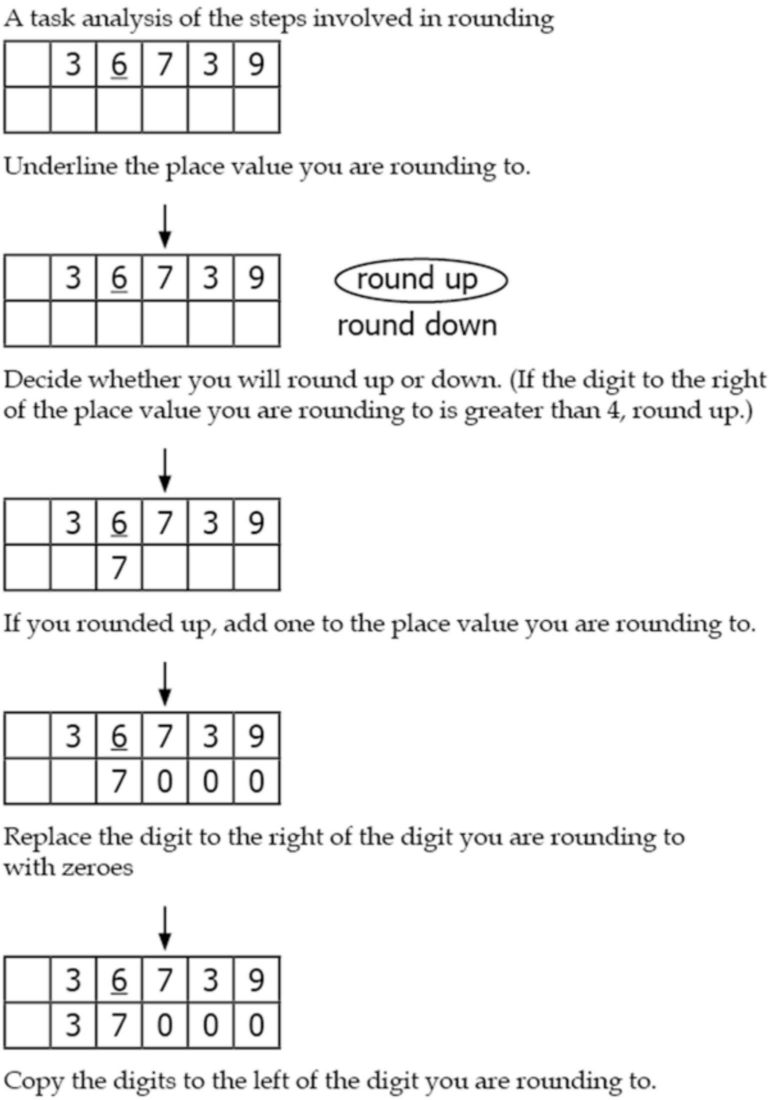
$$\begin{array}{r} 37 \\ +\ 25 \\ \hline 12 \end{array} \quad \text{or} \quad \begin{array}{r} {}^{2} \\ 37 \\ +\ 25 \\ \hline 1 \end{array}$$

To help student understand how to regroup, I will sometimes play a game in which I hold some pennies in my hand and ask the student to predict if I have enough pennies to make a dime. I also ask them to say how many pennies will be left over if I exchange the pennies for a dime. After repeating this exercise several times, I will hold some of the pennies in one hand and some in the other, so students have to mentally add the numbers and then decide if they can make a dime. Then I write various two-digit sums on the board and I tell the students to imagine that the numbers in the ones column represent pennies. In each case I ask them to say if I have enough pennies to make a dime and how many pennies will be left over. Then I tell them that in cases where they can make a dime I write a one above the numbers in the tens column to indicate that I have an extra ten or an extra dime.

Even when students understand why they write a 1 in the tens column when they want to regroup, they still need to practice this step. If I ask students to regroup the ones and add the numbers in the tens column (which involves adding three numbers) at the same time they will often forget how to add the ones. In the JUMP student books, in the exercises where student practice adding the ones, we put black boxes in the tens column to show students that they don't have to think about the tens yet.

$$\begin{array}{rcc} & \square & \\ & 4 & 5 \\ + & 2 & 6 \\ \hline & & \square \end{array}$$

Whenever I want to teach a particular mathematical procedure I first do a *task analysis* by listing every step (no matter how minute) I would follow to perform the procedure. For example, if I wanted to round 36,739 to the nearest thousand, here are the steps I would follow:



I hope you can see how easy each step of the procedure is when the steps are isolated. Even the most challenged student can perform the individual steps if they are given enough time to practice. Of course, in an actual lesson, I would help students understand why each step works, as I demonstrated in the lesson on part-whole problems. But I intentionally isolated the steps here, without focusing on the concepts, to show that the individual steps of the procedure people follow to round a number are absolutely trivial and could be mastered by anyone.

Every mathematical procedure that you learned (or didn't learn) at school can be reduced to steps that are as simple as the ones I outlined for addition and rounding.

This is true even for the advanced algebraic procedures that you had to learn in high school. As well, a competent teacher can create a series of Socratic questions, exercises, activities, and games that allow students to figure out why all of the steps of these procedures work themselves. That is why it is especially tragic that so many students fail to master and understand the mathematical procedures they are expected to learn at school or in training for their jobs. The JUMP writers and I wrote the JUMP Math teachers guides so teachers and parents could learn how to teach all of the math students learn between kindergarten and grade 8 using a method of instruction we call “structured inquiry.” Students are given many opportunities to explore concepts and figure things out for themselves, but the teacher provides enough rigorous guidance to ensure that all students master the material being taught.

Some teachers are reluctant to break instruction into manageable chunks because they think that this kind of teaching is “rote.” The term “rote” refers to a style of teaching where students are taught to blindly follow rules and procedures without any understanding of why those rules and procedures work. I hope it is clear from the lessons on part-whole problems that structured inquiry is not rote. In fact, research suggests that students who are led to discover concepts in manageable steps will develop a much deeper understanding of math than students whose brains are constantly overwhelmed.

Some teachers are averse to teaching in manageable chunks because they think students should struggle in math classes so they learn to persevere. But, as cognitive scientist Daniel Willingham points out, no one likes to struggle too much: “People like to solve problems but not to work on unsolvable problems. If school work is always just a bit too difficult for a student, it should be no surprise that she doesn’t like school much.” Psychologist Carol Dweck made a similar point after watching a JUMP lesson: she said the lesson incorporated “growth mindset principles” because the progression of exercises looked hard to the students but weren’t too hard.

Even though I believe that teachers should learn to teach in manageable steps, I do not advocate that they only teach in steps. The JUMP lessons include exercises, games, and activities that are less structured than the long division lesson. And every JUMP lesson ends with a set of “Extension” questions that are more difficult than the questions covered in the lesson. As well, I recommend that teachers skip steps and give students more challenging or open-ended problems when they have developed the confidence and acquired the knowledge they need to struggle productively. The goal of JUMP is to help all students become creative and independent thinkers who no longer depend on the teacher. But if teachers don’t know how to reduce instruction into small steps or to unravel concepts into their conceptual threads, then they are unlikely to help their entire class reach this level of achievement. As well, they won’t know how to assist students who are capable of more advanced work but who meet a momentary obstacle.

Some teachers are reluctant to guide the learning of younger students too much because they believe that children will naturally learn mathematical concepts on their own, by playing with “concrete materials” (blocks, toys, measuring instruments, and so on). But research has shown that this view is overly simplistic. While students certainly benefit from playing with concrete objects, they usually need assistance seeing the math that is embodied in the objects. As well, concrete materials sometimes appear to impede learning. In a recent study, one group of students was instructed to use play money resembling real paper money to solve a problem while another group was given more abstract money (rectangles with numbers printed on them). The group with the abstract money was more successful at solving the problem. Jennifer Kaminski, whose work I mentioned in Chap. 1, found that grade 1 children learn fraction concepts more readily with grey and white circles than with pictures of objects (for instance flowers with different-colored petals). According to Kaminski, her findings suggest "that concrete, perceptually rich instantiations of fractions may hinder children's acquisition of basic fraction knowledge in comparison to simple, generic instantiations for fractions."

Another body of research, called "concreteness fading," suggests that students learn some concepts better if the teacher starts with a concrete model but gradually and systematically "fades" to abstract symbols.

The Psychology of Success

As early as kindergarten, children start comparing themselves to each other and deciding who is smart and who isn't in various subjects. Most parents and teachers seem to think these comparisons are innocent and natural and so do nothing to discourage them—in fact some teachers actively encourage them. But based on my observations of thousands of students, I believe that our willingness to tolerate visible academic hierarchies is one of the main root causes of students' problems in math. As soon as students decide they are not in the talented group their minds stop working efficiently—they are no longer engaged in learning and stop paying attention, taking risks and remembering things. They may even develop anxieties or behavioral problems that make it even harder to learn.

When I have had the opportunity to observe teachers who have high expectations of every student and who follow the principals of structured inquiry, I've frequently seen entire classes of students become caught up in an intense collective excitement about math—similar to the "collective effervescence" that the sociologist Durkheim observed when crowds of people all experience the same positive emotions at the same time. Children in these classrooms often ask for extra work or even beg to stay in for recess to do more math. Their brains work more efficiently because they stop competing against each other in negative ways and start to direct all of their enthusiasm and mental effort to competing against the problem at hand.

Melanie Greene is a fourth-grade teacher who used the method of guided inquiry to achieve stunning results on the New York state tests in 2014. That year, she and her fellow grade 4 and 5 teachers at the Manhattan Charter School produced the greatest gains on the state test of all schools in New York City. In a blog that appeared in Achieve, Melanie described the high levels of engagement she saw in her students when they were taught through structured inquiry:

> What I saw in those scores wasn't even the full picture. Something special was happening in my classroom. Each day, my students could not wait to begin math. Even my lowest-achieving students were jumping out of their seats to answer questions. I will never forget one student in particular who cried at the beginning of the school year because math was so difficult for her. She quickly got on board with JUMP Math and received a four (the highest rating) on the New York State Test that same year. Thinking of her achievement still brings tears to my eyes.

Psychologists like Carol Dweck have shown that our beliefs about intellectual ability matter more than we think, in part because we can change the trajectory of a students' academic career with a single ill-informed remark. I've heard parents say, in front of their children, that the children can't possibly be good at math because their parents weren't, and I've heard teachers tell students that math is just not their subject. A growing body of evidence also suggests that our beliefs about ability can have an even wider impact, beyond our immediate families: for example, a recent study found that the less people believe that everyone is born with high intellectual potential the less likely they are to support free public education or the redistribution of educational funds from wealthy to poor districts.

People who underestimate the intellectual capacity of children are more likely to tolerate inequitable classrooms—classrooms where the teacher has low expectations of students who are perfectly capable of learning, where the methods of instruction are not designed to close the gap between weaker and stronger students, where students can easily rank themselves and compare their rates of progress, and where students with normal brains (who do not have severe learning disabilities) receive work that is substantially easier than the work given to more advanced students. Fortunately, teachers can now look to a large body of research in cognitive science to help them understand the true intellectual capacity of their students. They can also use this research to reduce visible academic hierarchies in their classrooms and to design lessons that allow students to learn in an efficient and enjoyable way, so that every student can realize their full potential in math.

References

Jennifer A. Kaminski and Vladimir M. Sloutsky, "Extraneous Perceptual Information Interferes with Children's Acquisition of Mathematical Knowledge," *Journal of Educational Psychology* 105, no. 2 (May 2013): 351–63.

Paul A. Kirschner, John Sweller, Richard E. Clark, "Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-based, Experiential, and Inquiry-based Teaching," *Educational Psychologist* 41, no. 2 (June 2006): 75–86.

Louis Alfieri, et al, "Does Discovery-based Instruction Enhance Learning?" *Journal of Educational Psychology* 103, no. 1 (February 2011): 1–18.

All images in the section "An approach to problem solving" are copyright JUMP Math. The section is adapted from the author's forthcoming book *All Things Being Equal: Why Math is the Key to a Better World* (Knopf Canada, 2019).

Chapter 15
Knowledge Building, Mathematics, and Creative Thinking: An Overview on Ontario Elementary Mathematical Teaching Beyond Twenty-First-Century Skills

Miss Stacy A. Costa

Elementary students in Ontario are bombarded with various mathematical classroom instructional methodologies, in order to assist them in improving their mathematical cognition. Children's first formal years of schooling can determine a student's awareness regarding their mathematical achievement, their math anxiety, and motivational stance regarding mathematics (Gunderson et al. 2018). While formal schooling is the most prominent way to disperse mathematical knowledge, children do discuss and embody math and its properties outside of the classroom environment. Mathematical knowledge can be acquired outside of formal school instruction and become a positive influence on mathematical performance throughout one's lifetime (Brownell 1941). Ideas such as financial literacy, probability, and patterning are easy topics in which young students can encounter within their everyday lives. Nonetheless, students in Ontario are still experiencing difficulties when applying their understanding of mathematics to given problems within standardized testing. While the author of this chapter is aware of the problematic nature of standardized testing, it is important to note these results as a basis for the argument for the rest of this chapter. Ontario's Education Quality and Accountability Office (EQAO 2016/2017) released their annual results of standardized testing on mathematical skills. The 2017–2018 results showed that fewer than half of the province's grade 6 pupils—49%—met the provincial standard in math in the 2017–18 academic year. According to the EQAO standardized test in the 2016–17 academic year, only 49% of grade 3 girls in Ontario agreed with the statement they are good at math compared to 62% of boys. The difference widened in grade 6, where 46% of girls said that they were good at math compared to 61% of boys. While these statistics are problematic for Ontario educational partners, they only demonstrate

M. S. A. Costa (✉)
Curriculum, Teaching & Learning & Collaboration in Engineering Education, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada
e-mail: stacy.costa@mail.utoronto.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*, Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_15

the surface of the problem. Students are doing poorly not only on their results of testing, but also on their self-confidence, and specifically math anxiety within females is also distressful. This chapter examines methodologies to go beyond twenty-first-century skills, and how future mathematical instructional methods must become more rigorous in order to change and make an impact on Ontario student's mathematical cognition. This chapter explores possible instructional methodologies in which educators should explore in order to find best practices for their students to build upon.

Students are currently studying for a twenty-first-century workplace wherein skills are defined by student's success in being able to tackle and negotiate unforeseeable problems within global, technological, and the modern world. These challenges need to be taken on by innovative citizens. Such innovative citizens need to have mathematical knowledge to understand facts, reporting, and statistical information. Student's acquisition of math skills is essential, as they are an indicator and possible predictor of future prospective careers. Through the use of twenty-first-century skills within a mathematics classroom, students will embody skills such as collaboration, mathematical reasoning, and preparation for STEM (Science, Technology, Engineering and Mathematics) majors. While not all students may fall into this same path, it is still vital so as students do not terminate mathematics studies earlier in their learning trajectory.

Mathematics is needed for problem-solving, reasoning, questioning, computational strategies, and creative application, well beyond the work required to solve well-defined problems characterized by "instructor input → student problem-solving → verification" model. Students can arguably be trained on how to complete irrelevant math problems to understand formulaic math queries but not to comprehend outside of imaginary textbook questioning. Arryo et al. (2014) found that students can progress through higher level math classes yet are missing foundations of mathematical reasoning, thus providing teachers with numerous math difficulties that vary across students within the class and making it difficult for teachers to meet all the needs of every student.

Students acquire mathematical knowledge and procedures through a variety of instructional methods. However, within educational reforms, students are now streamlined into utilizing only one correct method of acquiring mathematical texts, and this can become a stepping stone to exacerbating the problem of mathematical knowledge being validated. Streamlining can lead to concerns of math anxiety, or poorer academic performance. This is one explanation as to the statistic nightmare reported earlier by EQAO and Ontario student's scores. By age 4 or 5 most children seem to have both the conceptual foundation needed for the acquisition of complex mathematical knowledge (Balfanz 1990, p. 45). Through classroom instruction, there is difficulty in transferring mathematical knowledge to novel contexts. Bredekamp (2004) argues that educators should examine research and practices on how children learn mathematics and integrate relevant methods into school mathematics. Most mathematical instruction needs to be assessed on how to integrate mathematical competencies best to be well understood by students. As mathematical concepts are precursors to the next set of mathematical knowledge, it is easy to have skills become too tricky or misunderstood, leading students to abandon the subject during upper years of schooling entirely. This then leads into a problem within high school of academic versus applied streaming. Unfortunately, these two categorical choices are made for students and can easily perpetuate disparities in educational outcomes. Understanding why students are being streamed within their education not only disadvantages them but sets them with a mindset that they must stay within such streams and what their capabilities are at such a young age.

In the Liberian schools examined by Brenner (1985), teachers promoted the use of indigenous methods in concert with school-taught procedures. By incorporating what Brenner (1985) examined provide space and importance for student's voice and interpretation to incorporate student's understanding and theory. When students are introduced to textbooks, and other authoritative sources, a student can begin to understand and relate to their own unique context, relational understanding of math to answer inquiries they may be facing at their age. However, when referencing these materials, if not scaffolded correctly, problems that were initially introduced are structurally identical to the ones they will continue to work on during the lesson. No unique problem-solving or twenty-first-century skills are incorporated.

"Most children's environmentally acquired knowledge is obtained orally and developed mentally, without the use of writing" (Balfanz 1990, p. 46). Children learn their basic numeracy skills through math talk, sharing, manipulating, and trading. Understanding mathematical cognition would go beyond problem-solving. We need to understand that mathematical understanding incorporates more than just symbolic reasoning. Balfanz (1990, p. 46) found that as a result, young students begin to cease understanding mathematical as a thinking process but instead a mechanical one due to their classroom instruction.

One way in which mathematical classrooms in Ontario could benefit would be through the use of innovative principle-based learning like knowledge building (Scardamalia and Bereiter 2014). In line with the World Economic Forum (2016) focus on collaborative work on complex problem-solving at the elementary level, collaborative knowledge building (Scardamalia 2002) supports building ideas in an online space accessible to all participants. By providing students with ample opportunities to interact with their peers and play with math ideas, knowledge building has the potential to address this perceived math anxiety, especially for the most vulnerable population, Ontario girls, and enhance their student achievement. Thus, let's go beyond and empower our learners within the math domain. Not unlike Brenner's (1985) approach, knowledge building allows for student voice and classroom promotion of relevant methods to be incorporated with school-taught curricula. This pedagogy goes beyond problem-based learning and has previously been studied and implemented successfully on elementary students' mathematical studies in geometry (Costa 2016). Knowledge building also aligns with its principle-based software knowledge forum which is an innovative virtual space. Knowledge forum is one method to improve students' mathematical communication ability, and studies have shown the potential of using computers (Costa 2016).

It is possible to note that knowledge building as an educational technology can tailor metacognitive support to individual student needs, including within the discipline of mathematics. Studies have shown that knowledge developed in school may only be used in school (Greenfield and Lave 1982; Carraher et al. 1987); where students are often deprived of the information, they need to understand why their invented methods are incorrect. It is important that students go beyond correct or incorrect ideas or theories and instead be able to incorporate more mathematical talk, explanations, and theories in order to better understand beyond just demonstration of procedure of the mathematical concepts being studied. It is essential to recognize and have students share the knowledge they develop outside of the classroom and legitimize their knowledge. By building upon student's mathematical interests, it segues to more in-depth math talks and concepts for students to grasp as oppositional theories or ideas meant to be built upon. By allowing students to have multiple representations of their math talk, students were able to type text, use speech, correlate new math vocabulary to manipulatives, draw and annotate images to demonstrate their understanding, scaffold ideas, and also discuss how being taught the same idea providing different theoretical perspectives on the same terminology. These examples all allowed students various opportunities to make errors and to have other students assist in catching those errors or providing rationale and explanations to support their reasoning and understanding. Furthermore, it provides students with multiple methods to express and demonstrate their mathematical understanding. Knowledge building pedagogy also allows students to promote their own graphical literacy. Students built shapes and had discussions surrounding what they considered a shape, and what properties a shape possessed. Students took agency over importing images and describing the symbolic renderings if images were shapes, by incorporating conversations surrounding ideas such as hearts, being a shape versus being a symbol rendering (Costa 2016).

Ontario students' standardized testing scores also demonstrate the problems instruction is facing within the province. The elementary math curriculum cannot only be focused on skills that are not connected and are solely drill tasks or are solely discourse conversations of the technique used. It is evident that mathematical methods which are being introduced lack modular use and applicability to problems that are perceived to be more difficult or provide alternative context. Students are grounded in the text examples they are provided initially and cannot maneuver between new tasks to apply their knowledge to. Students are not grasping nor metacognitively understanding that their newly introduced math knowledge is applicable in multifaceted methods.

Conversely, discovery math is often referred to as inquiry-based mathematics or open-ended math. Many are critical of this technique as it may facilitate only classrooms that do not emphasize drill-set learning or focus on procedural knowledge. One area that can be more prevalent within a discovery math classroom is the usage of math talk. During these sessions students articulate their thoughts, expand their understanding of mathematical terms and concepts, listen to their peers, and describe the math task they are presented. Discovery math should be used in tandem with skill-set-building procedures in math classrooms to allow young students to be comfortable and confident in using these terms.

Understanding how we can create mathematical understanding stemming from a young age seems to remain a mystery still, as many youths are still struggling with math cognition and choosing not to enjoy math as it is perceived as difficult, unrelatable, or pointless to student's self-learning goals. Given that math is a conceptually rich and complex subject, students often perceive that math is more difficult to understand than other subjects in school (Costa 2015). Many students are hesitant to voice confusion or ask for help when stuck on a mathematical problem, which deters from their learning. Gunderson et al. (2018) found that students' mathematical achievement and reciprocal math anxiety were related and found effects present in grade schools' student within their first 2 years of formal schooling, with each demonstrating that at a young age a positive or negative trajectory in math can occur. Therefore, if students are performing with difficulties at the grade 6 level in Ontario, it could be led to believe that these are cascading effects which can continually determine and perpetuate difference in low math achievement and avoidance in math-related tasks. We need to have skills applied in order to have students cognitively be aware of how to apply their mathematical learning to additional problem sets, and not just create understanding and skills for specific problem sets. "From the known to the unknown through hypothesis, experimentation and analysis, solving a hard problem by looking for ways to simplify it and thinking by analogy" (Balfanz 1990, p. 54). Student mathematical thinking that emerges must go beyond just math talk, but also acquiring understanding to apply these theories to solve problems and apply the thinking to other problem sets.

As Greeno et al. (1996, p. 20) note, "knowledge is distributed among people and their environment, including objects, artifacts, tools, books and the communities they are a part of." Students must take collective responsibility for advancing community knowledge, using resource material to extend their work beyond the ideas in the local community. Students can shift between the role of providing peer feedback or asking for information in order to advance their own thought processes and as well of that of their classroom Knowledge building mathematical community. In this scenario, students are forced to disclose their thoughts, and see value in the importance of sharing knowledge, as well as understanding metacognitive thought of self and peer. This then lends to additional feedback from learning between students and teacher interjection to facilitate and assist when needed. Costa (2016) found that by using symbols and graphical representation of math ideas, students can build representations of mathematical structures and designs that incorporate other multimedia components that can be attached and uploaded onto the group view, allowing students to build on ideas in various formats in which the technology is flexible as to cognitive learning style. Moreover, all uploads and representations can be co-authored, allowing for collaborative production with potential for stronger motivation and engagement. This example can not only help foster student's

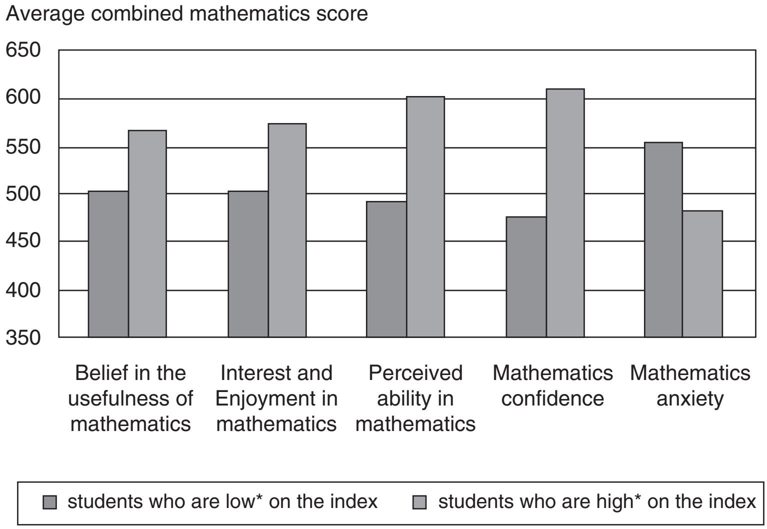


Fig. 15.1 Combined mathematics score for students with high mathematics engagement compared to students with low mathematics engagement (Ministers of Education 2004). Students low on a given index are defined as those falling one standard deviation below the average, and students high on a given index are defined as those falling one standard deviation above the average

learning and understanding but also be a point of instruction to lead the teacher to incorporate areas that are unclear for students. Mathematical misconceptions may have the potential of the student thinking, furthering the mathematical understanding of the class if incorporated into instruction (e.g., Stockero and Van Zoest 2013).

As we can see from Figure One (Fig. 15.1) below (Ministers of Education 2004), students who are not engaged are correlated to have higher mathematics anxiety, with lower mathematics confidence, and do not have an interest and enjoyment in mathematics performed the equivalent of one proficiency level lower. These statistics are concerning as those who have higher scores are paradoxically higher in mathematics confidence and lower in mathematics anxiety. These students who positively related to achievement also had higher motivation to learn math. As noted, these scores are prior to EQAO scores from 2017, and of Canadian country-wide averages not just the province of Ontario. These scores still emphasize a shift and concern with mathematical instruction.

Problem-based learning (PBL) is a "teacher facilitated, student-driven approach" (Bell 2010). Problem-based learning can be seen as an "object of consideration by the class in order to engage the class in making sense of that thinking to better understand an important mathematical idea" (Van Zoest et al. 2017, p. 36). Boaler (1999) found that students at a PBL school were equally able to answer procedural questions that used formulas to a traditionally taught math program, as well as they were superior in answering applied and conceptual problems. Boaler's (1999) research supports that students using problem-based learning perform better on standardized tests and that they had learned the real-world application and analytical thinking; thus, they could see the application of their learning and were less likely to view math as an isolated or useless skill. We want our students also to be active learners within the math classroom. If a student then begins to feel spoken to, it is easier for them to disengage and daydream, if they are aware that they are only told facts, and are not provided with any learning opportunities to engage with mathematical concepts. By allowing students to engage in twenty-first-century collaboration and communication skills this honours students' agency of learning styles or preferences.

Students are passive observers and with demands to meet curricula expectations, or teaching needing to be conducted, students cannot be challenged or think about mathematical application in creative ways or entrepreneur ways that match mathematics to familiar applications within their own live trajectories. Novita and Putra (2016, p. 35) define creativity within mathematics as "abstraction, connection and research." Now, while there is no universally accepted definition of mathematical creativity (Haylock 1997) this chapter would declare an abstraction of mathematical creativity as a methodological to see and understand mathematics in multiple ways, and to understand various demonstrations to represent their understanding. Writing allows children to make connections between the spoken and written word (Cohen et al. 2015). Students can write about their reasoning, develop math vocabulary, organize their thoughts, and develop problem-solving methods (Furner and Berman 2012). One aspect which is not mentioned throughout testing or within math instruction is creativity. Creativity is a skill in which will serve students as a fundamental skill in twenty-first-century evolving global economy. Unfortunately, as it seems, the concept of creativity being sidelined within lessons also removes student's natural curiosity and interest in mathematics beyond rules and formulaic measures. The need for mathematics as a tool for creative work, in a technology-rich knowledge society, is widely recognized (Wagner and Dintersmith 2015; Ritchhart 2015).

In mathematical learning, students need to understand two knowledge types: procedural and conceptual (Baroody and Dowker 2003; Sidney and Alibali 2015). Sidney and Alibali (2015) define procedural knowledge as "knowledge of sets of actions that can be used to solve a particular type of problem" (p. 162) while conceptual knowledge is knowledge surrounding meaning and processes (Rittle-Johnson et al. 2001). When referring to standardized testing, is EQAO testing solely distributed to collect procedural knowledge? If so, then can a student succeed if they choose to display conceptual knowledge yet are not tested for this? Again, the author is aware of the controversy surrounding standardized testing as a whole, but it begs the question: Do successful math-fluent students need to possess both types of knowledge to be adequate or successful mathematicians?

As this chapter notes the inconsistencies and concerns with elementary mathematical learning, it is more important to emphasize the importance of student interest and curiosity of mathematics at a young age to stem on long-term outcomes instead of short-term outcomes of EQAO scores or mastering a specific unit. It is concerning as if students draw on existing knowledge, to understand to build on future mathematical questions, how do students transfer successfully? Are students failing to identify that the questions are similar and relate to their current knowledge?

Effective instruction cannot occur for students with mathematical difficulties if a teacher is halted due to such logistical concerns. As noted in this chapter, it is crucial that teachers and classrooms be open to understanding how student ideas are dealt with and identified within the math classroom. These ideas need to be constructively understood by all participants in the classroom, and they should open up new avenues of instruction and inquiry to support teachers to respond, according to the various challenging and unique inquiries students may possess. Furthermore, while classrooms import various ideas, this will not mean that the instruction will be linear but instead will need to provide a more in-depth and highly flexible routine that must be open to shift in teacher belief and differentiated instruction. Teachers must go beyond a set method but instead be fluid based on math topic, and classroom student response.

Murata (2015) describes a concept seen as "instructional width," in which a classroom incorporates a wider variety of ideas that occur during instruction to support student's individual and varying learning goals to converge towards classroom instructional goals gradually. This concept would be a possible solution for Ontario math teachers to allow for students' understanding in curricular mandated areas to be met but to also provide individual needs that can be overwhelming for one teacher.

Henning et al. (2012) found that one useful methodology for elementary mathematics instructors is to facilitate math lessons through three sections: framing, conceptual, and application. This methodology allowed for teachers to maintain their explanation and allowed for more complexity to be described by the teacher. The rest of the lesson allows for student clarification of understanding, and to encourage and evaluate other's ideas. In summary, teachers must be open to several instructional strategies in order to allow students to benefit in the long-term understanding of mathematical cognition to occur.

While standardized testing scores are problematic measures, it is crucial that we look at these results and question and try to understand new methodologies and instruction and be open to mathematical learning in a less rigid fashion. As noted, we need to continually work on the multifaceted approach of student's mathematical learning trajectory. It is important to note that both informal, basic quantitative competencies and domain-general abilities contribute to formal mathematics learning (Bull et al. 2008). This overview can provide Ontario educators, policymakers, and parents with some knowledge and facts to consider in instructing mathematics to Ontario students, primarily to provide vulnerable populations, a shift in student's self-mathematical cognition. This chapter has concluded that mathematical instruction for Ontario classrooms needs to be open to differentiated practices and further review.

References

Arryo, I., Woolf, B., Burelson, W., Muldner, K., Rai, D., and Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence Education* 24: 87–426.

Balfanz, B. (1990). Elementary school quality, the mathematics curriculum and the role of local knowledge. *International Review of Education* 36 (1): 43–56.

Baroody, A. J. and Dowker, A. (2003). *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Mahwah, NJ: Erlbaum.

Bell, S. (2010) Project-based learning for the 21st century: Skills for the Future. *The Clearing House* 83: 39–43

Boaler, J. (1999). Mathematics for the moment, or the millennium? *Education Week* 17(29): 30–34.

Bredekamp, S. (2004). Standards for preschool and kindergarten mathematics education. In: H. Clements and J. Sarama (eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education*, (pp. 77–87). Mahwah, NJ: Lawrence Erlbaum Associates.

Brenner, M. (1985). The practice of arithmetic in Liberian schools. *Anthropology and Education Quarterly* 16: 177–186.

Brownell, W. A. (1941). *Arithmetic in grades I and II*. Durham, NC: Duke University Press.

Bull, R., Espy, K. A., and Wiebe, S. A. (2008). Short-term memory, working memory, and executive functions in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology* 33: 205–228.

Carraher, T. N., Carraher, D. W., and Schliemann, A. D. (1987). Written and oral mathematics. *Journal for Research in Mathematics Education* 18: 83

Cohen, J., Casa, T., Miller, H., and Firmender, J. (2015). Characteristics of second graders' mathematical writing. *School Science and Mathematics* 115 (7): 344–55.

Costa, S. (2015). Exploring student "Math Talk" in 21st Century Classrooms. In: M. Bockarova, M. Danesi, D. Martinovic, D., and R. Núñes (eds.), *Mind in mathematics: Essays on mathematical cognition and mathematical method*, (pp. 176–182). Münich: Lincom Europa.

Costa, S. (2016). *Math discourse in a Grade 2 Knowledge Building classroom*. Master's thesis, University of Toronto.

Furner, J., and Berman, B. (2012). Review of research: Math anxiety: Overcoming a major obstacle to the improvement of student math performance. *Childhood Education* 79 (3): 170–174.

Greenfield, P. and Lave, J. (1982). Cognitive aspects of informal education. In: H. Stevenson (ed.), *Child development in cross-cultural perspective*. San Francisco: Freeman Press.

Gunderson, E., Park, D., Maloney, E., Beilock, S., and Levine, S. (2018). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. *Journal of Cognition and Development* 19 (1): 21–46.

Haylock, D. (1997). Recognizing mathematical creativity in schoolchildren. *Zentralblatt fuer Didaktik der Mathematik* 29(3): 68–74.

Henning, J. E., Mckeny, T., Foley, D., and Balong, M. (2012). Mathematics discussions by design: Creating opportunities for purposeful participation. *Journal of Mathematics Teacher Education* 15 (6): 453–478.

Human Resources and Skills Development Canada, Council of Ministers of Education, Canada and Statistics Canada (2004). Measuring Up: Canadian results of the OECD PISA Study: The performance of Canada's youth in mathematics, reading, science and problem solving, 2003, Vol. 2, Catalogue number 81–590-XIE2004001.

Greeno, J. G., Collins, A. M. and Resnick, L. B (1996). Cognition and learning. In: D. C. Berliner and R. C. Calfee (eds.), *Handbook of educational psychology*, pp. 15–46. New York: Macmillan.

Murata, A. (2015). Interactions between teaching and learning mathematics in elementary classrooms. In: D. Scott and E. Hargreaves (eds.), *The Sage handbook of learning*, (pp. 233–242). London: Sage.

Novita, R. and Putra, M., (2016). Using task like PISA's problem to support student's creativity in mathematics. *Journal of Mathematics Education* 7 (1): 31–42.

Ontario. (2016/2017). EQAO: Education Quality and Accountability Office. Toronto: The Office.
Ritchhart, R. (2015). *Creating cultures of thinking: The 8 forces we must master to truly transform our schools*. Hoboken: Wiley.
Rittle-Johnson, B., Siegler, R. S., and Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology* 93: 346–362.
Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In: B. Smith, and C. Bereiter (eds.), *Liberal education in a knowledge society*, (pp. 67–98). [Berkeley, CA.]: Distributed by Publishers Group West.
Scardamalia, M., and Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In: K. Sawyer (ed.), *The Cambridge handbook of the learning sciences,* pp. 397–417. New York: Cambridge University Press.
Sidney, P., and Alibali, M. (2015). Making connections in math: Activating a prior knowledge analogue matters for learning. *Journal of Cognition and Learning* 16(1): 160–185.
Stockero, S. L. and Van Zoest, L. R. (2013). Characterizing pivotal teaching moments in beginning mathematics teachers' practice. *Journal of Mathematics Teacher Education* 16: 125–147.
Van Zoest, L. R., Stockero, S. L., Leatham, K. R., Peterson, B. E., Atanga, N. A., and Ochieng, M. A. (2017). Attributes of instances of student mathematical thinking that are worth building on in whole-class discussion. *Mathematical Thinking and Learning* 19: 33–54.
Wagner, T., and Dintersmith, T. (2015). *Most likely to succeed: Preparing Our kids for the innovation era*. New York: Scribner.
World Economic forum (2016). 10 skills you need to thrive tomorrow—and the universities that will help you get them. Directors of Research and Evaluation, US Department of State. https://www.weforum.org/agenda/2016/08/10-skills-you-need-to-thrive-tomorrow-and-the-universities-that-will-help-you-get-them/

Chapter 16
Crypto-Mathematics in Ethnography: Estimation and Approximation via Ballparks and Eyeballing

Myrdene Anderson

Foreplay

Of course, I know that counting isn't really maths, but the cognitive play in estimation and approximation might qualify! To set this up, some relations pertaining to mathematics and arithmetic, and to anthropology and ethnography, are in order, as ethnography is the quintessential way humans study humans, whether they're aware of the process or not.

The Intersection of Mathematics and Anthropology

Mathematics and ethnography overlap at several angles. Ethnography, as one approach to the scientific study of humans, is associated with the discipline of anthropology. However, anthropology overall employs any manner of "hard" technologies as well as "soft" observation—everything from microscopes to telescopes, as well as our primal sensational, perceptive, and cognitive aptitudes of humans as ethnographic investigators, singly or in groups, about their own conspecifics. Furthermore, the scope of the entire anthropological discipline leaves no holds barred when it comes to tackling space and time … in the quest to understand humans from their evolutionary emergence to the present and into the future, in time, and from pole to pole and into any future spheres, in space.

Anthropology in Boasian North America inherits a century-long tradition of four subdisciplines, or subfields, that a century ago emerged like Topsy from holism—that commitment to cover as much as possible about our species, or at least not to willfully overlook let alone erase any perspective. These four fuzzy subdisciplines

M. Anderson (✉)
Department of Anthropology, Purdue University, West Lafayette, IN, USA
e-mail: myanders@purdue.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_16

can't be sequenced, as in fact they overlap, with fallow interstices and redundancies, all seeming to be splitting out of their seams.

To be explicit, the four subfields' labels specify the link to anthropology, itself a young discipline, thus in some cases allowing for parallel formations in sibling disciplines: archaeological anthropology, biological anthropology, linguistic anthropology, and sociocultural anthropology. Orthogonal to these four subfields is a social-problem-oriented approach leading to other applied anthropologies such as activist, advocacy, practicing, and more. On occasion, an interstitial subfield (for example, medical anthropology, overlapping potentially with all four subfields) may itself be a label for a project in applied anthropology. As in the case of applied mathematics (Halmos 1984), applied anthropology must also first be anthropology.

While ethnography—as participant-observation in a quotidian human setting—typically unfolds from sociocultural anthropology, its practice may also be found in the other three subfields as well as in the applied approaches. Ethnography has also been adopted by a number of other fields, including sociology and education. The adventitious roots of anthropology lie in the cultural practices and the individual habits already available to its earliest self-conscious practitioners. Justifying the four subdisciplines may be either or both futile and irrelevant; the following substruction may honor some of the shaping factors, both cognitive and material (Fig. 16.1).

Ethnography as "Participant-Observation"

Zeroing in onto ethnography (sense one), this participant-observation research tack is carried out in an actual experiential space and time—a natural habitat of the studied humans, saturated by habits of the insiders' culture and language (be it in a kindergarten playground or a scientific laboratory, or even an individual subject)—where the researcher, typically an outsider, documents the quotidian goings-on. More recently, one hears this boiled down to "deep hanging-out." The ethnographer is never invisible, so it goes without saying that doing ethnography has at least the tacit permission of the studied persons, while inevitably but inadvertently influencing them and the setting, and being affected in return. Consequently each ethnographic venture will be unique in every way, beyond just the nature and extent of "participation" of the outside researcher in the language and culture of the activity and events of the insider population, and respecting the nature of the "observation"—which in every case would include many other senses alongside vision.

I am inclined to be inclusive when it comes to what qualifies as ethnography. Sadly, no longer will Darwin's adventures qualify, if he has been properly quoted (in Shermer 2019, p. 67) as writing, "… all observation must be for or against some view, if it is to be of any service." The ethnographer might be in a too-familiar situation, as would pertain to autoethnography, or in one so alien that reality disputes believability. The ethnographic endeavor may be carried out in several settings and be of any length, even interminable, as longitudinal projects may effectively

<table>
<tr><td colspan="2"></td><td colspan="3">Spatial phenomena</td></tr>
<tr><td colspan="2"></td><td>Focused</td><td>General</td><td rowspan="5">General systems</td></tr>
<tr><td colspan="2"></td><td colspan="2">Often elucidating transactions of …</td></tr>
<tr><td colspan="2"></td><td>Energy/ecology</td><td>Information/ethology</td></tr>
<tr><td rowspan="3">Temporal phenomena</td><td>Span or general
Diachronic bias
Often elucidating process and function // evolution and development</td><td>Archaeological and prehistoric anthropology</td><td>Biological anthropology, formerly called physical anthropology</td></tr>
<tr><td>Focused
Synchronic bias
Often elucidating system and structure // communication and exchange</td><td>Ethnography, cornerstone in sociocultural anthropology</td><td>Languaging, fusing aspects of linguistic and cognitive anthropology and more</td></tr>
<tr><td colspan="3">General systems</td><td>Semiotics inclusive of biosemiotics, semiotics of culture, and literary, cognitive, computational semiotics</td></tr>
</table>

Fig. 16.1 Substruction of the four subdisciplines of anthropology

continue beyond one lifetime. For the ethnographer, ever-emerging context, that is relations within and across time and space, will provide the justification for Darwin's "observation" (Deely 2012).

Some ethnographic projects involve a team, and any solo ethnographer may also have personal accompaniment in the field, perhaps family. Given the cognitive tilt of any human, and other living things (given markedness, Waugh 1982, given error detection, Noritake et al. 2018), the ethnographer may first notice some familiar expected phenomena, but upon recognition will be struck by distinctions; this process—since the second wave of Russian Formalism—is now often called ostranenie, or defamiliarization. The researcher may reflect on the particulars at hand as data, with or without relating them to analogous data or larger patterns pertaining to human cognitive constraints (Ludwig 2018) or even matters of scale (West 2017).

The outsider ethnographer-researcher, acknowledging his/her own background, focuses on noticing, eventually tentatively understanding, even collaborating in performing, the inferred lived realities of the others, those insiders. Terminologically clumsily, the ethnographer's pursuit of ethnography (sensu one), once summarized, analyzed, and perhaps compared with other research, results in publication of formal documents, articles, and perhaps books, any of these also called ethnography (sensu two).

In this naturalistic research, the ethnographer gathers data and captures capta—however possible, often summed up as participant-observation—in a summative process called "writing-down." The writing may entail media other than pencil and paper, from computing to film, any and all of which requiring digestion [often "analysis" before the production of any, even interim, result, the ethnography (sensu two)].

Every field of inquiry—from art to zoology—may call on mathematics, or at least arithmetic, in the carrying out of a research project. In the case of anthropology's commitment to holistic research, diligently inclusive of all research subjects, their actual and imagined material and mental conditions, anthropology is in a position not only to employ mathematics in collection, analysis, and writing-up stages, but also to discover the mathematics indigenous to other linguïcultures (sensu Anderson and Gorlée 2011; Fabbri 2013, Overmann 2015, Van Bendegem and Van Kerkhove 2004).

Although some ethnomathematical systems had been documented, those other mathematics went largely unstudied by mathematicians, anthropologists, and philosophers, alike, until the 1970s, when scholars were equipped to discuss the variety, similarities, and distinctiveness among these systems (Ascher 1991, Berland 1982, D'Ambrosio 1997, Lave 1988, Selin 2000, Wilder 1981, Zavlasky 1973; D'Ambrosio's work foundational from 1985). Their variety should not be any more surprising than the variety in form and in substance manifest across cultures and, especially, languages, all deriving from but also shaping cognition itself. Much research continues to contribute to the Sapir-Whorf conjecture (from 1929; surely not a hypothesis, though called one elsewhere) (Hill and Mannheim 1992, p. 386), which conjecture proposes that the structure of a language shapes its speakers' worldview or cognition. Finally, anthropologists perhaps more than linguists have been captivated by J. Willard Gibbs' (1839–1903) quotable quote that "mathematics is a language" (in Silver 2017, p. 364). At the same time, mathematicians like Daniel S. Silver (2017) now probe whether or not there can be mathematics without worded language, illustrating the promiscuity among and between cognition, communication, and culture.

Introducing "Discrete" and "Indiscrete" Anthropology

Data and capta come in two flavors, even for Peircean semioticians: the digital (discrete or discontinuous/episodic/categorical) amenable to counting, and the analogic (continuous in space or time or substance) that, unless reduced by convention for enumeration, can only respond to measurement. Differences also come in two flavors: in kind or in degree.

That is, digital data may be either countable if of one kind, with numerals ("discrete"), or clumpable and then manipulated arithmetically (Bateson 1978; Conner 2016;

Kauffman n.d.). Analogic data, reflecting the infinite variables and infinitely varying cosmos that immerse us in the so-called real world, require intervention before quantification, thus inviting the fanciful moniker “indiscrete.”

> Hence, clearly, qualification precedes quantification, even though, in sciencing, the numerical may appear to trump narrative, especially when the quantitative is enhanced through statistics … Whatever we choose to count, measure, or weigh, or manipulate in any way, has already been masticated by our linguïculturally-saturated human minds and habits (Anderson 2012, p. 296).

Too many persons within and beyond the academy have heard about, and perhaps are concerned with, a quantitative versus qualitative “debate” across the human sciences (Tenenbaum et al. 2011). This discourse has been particularly poignant in anthropology, given its inherently interdisciplinary, multidisciplinary, and even transdisciplinary stance. Currently, some assume that these incommensurable paradigms can be resolved through “mixed methods” (Johnson et al. 2007), these often decorating the edge of ethnographic projects with token quantitative data. Any textual or numerical data can be repackaged, especially visually, and fed into big data, including digital humanities (Aiden and Michel 2013; Drucker 2014), but these researchers may not be involved in original documentation at all.

Ethnography—a methodology based on the most fundamental and foundational qualitative approach in anthropology and summed up beyond participant-observation as naturalistic or qualitative inquiry—is now also practiced in the “pure” and “applied” social and behavioral sciences, even for research in mathematics education (Anderson et al. 2003). It may not be obvious how mathematical cognition will be relevant to such a naturalistic approach. But consider the assumption that mathematical cognition cannot be excised from any participatory or observational enterprise! That would include living as a subject in one’s ordinary environment, from birth to death. This view assumes that ethnography is not just “naturalistic inquiry” but integral to our normal faculties, in fact suturing our variously attuned senses into patterns and even narratives, tracing the relations within our Umwelt.

First, in broadest strokes, a Rorschach of ethnography would specify that it is carried out by a human investigator or investigators, called ethnographers, in a spatiotemporal setting with other humans—or even other creatures, but thus far not with plants as their behavior doesn’t sync with ours—who allow or at least tolerate the ethnographer to observe, and even to clumsily participate. That’s at the initiation of what is usually a longer project, even years. As the ethnographer acclimates to the language, culture, and especially personalities of the subject group, the very quality of the observation and of the participation will shift. With experience the relevance, direction, and quality (if not quantity) of both participation and observation are shaped by habit and by the emerging setting. It is important to consider the unique and open dynamics of each ethnographic situation, even holding constant the ethnographer.

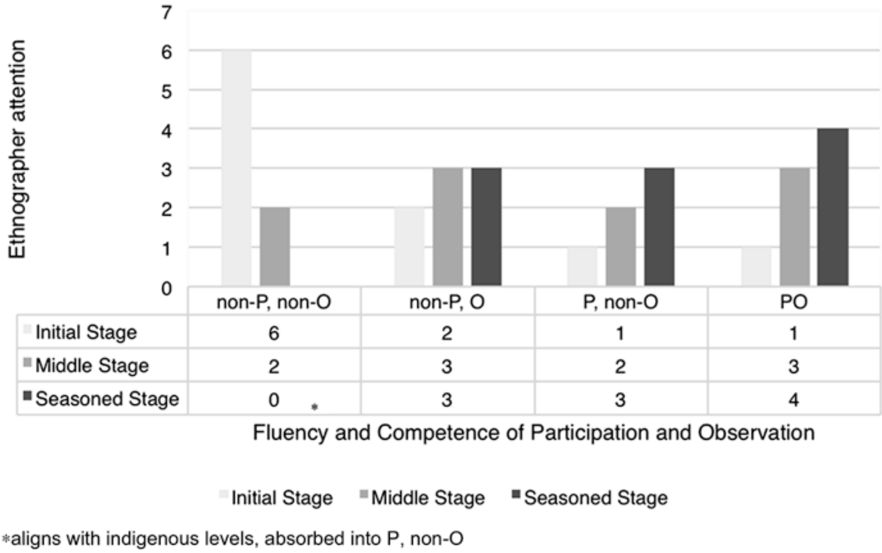
I have first decoupled the basic variables of participation and observation in Fig. 16.2 (adapted from Anderson 1999–2000, p. 184), showing their permutations in a static Punnett square. Then, in Figs. 16.3, 16.4, and 16.5, the relative inputs of participation and observation are put into time and motion, showing how inputs

		Scale of fluency Depth and intensity of involvement	
		Nonparticipation = $\overline{\mathbf{P}}$	Participation = **P**
Scale of competence Breadth and exhaustiveness of documentation	Non-observation = $\overline{\mathbf{O}}$	$\overline{\mathbf{PO}}$ = Dissociated state or trance; culture shock, nonanalytic “hanging out” in an unfamiliar setting	$\mathbf{P}\overline{\mathbf{O}}$ = Ordinary pole of “subjective” practice in familiar languaculture, operational understanding taken for granted
	Observation = **O**	$\overline{\mathbf{P}}\mathbf{O}$ = Extraordinary pole of “objective” practice in partitioned “normal” science setting. Hypotheses guide operational inquiry	**PO** = Merging state of ethnographic practice in dialectical, dialogic setting. Translation-quadrant may be further substructed

Fig 16.2 A first-level substruction of the field of participant-observation

change from the initiation of a research project—for either an experienced or inexperienced ethnographer—through to what would obtain for a seasoned, more competent, and more confident, researcher.

What is documented by the ethnographer, and “written down,” qualifies as data (or “given”), whether description, quotation, numerical, sketch, or in another modality altogether. Collections of data may lend themselves to refinement through selection as capta (or captured) (Drucker 2014; Weissner 2016), before further interpretation and analysis. Moreover, the ethnographer will be aware as to whether the data would be considered “etic,” objective from the observer’s vantage, and/or subjective from the projected insider’s vantage, and whether these distinctions will be productive.



	non-P, non-O	non-P, O	P, non-O	PO
Initial Stage	6	2	1	1
Middle Stage	2	3	2	3
Seasoned Stage	0 *	3	3	4

Fig. 16.3 Changing allotment of ethnographer's time in participation and observation

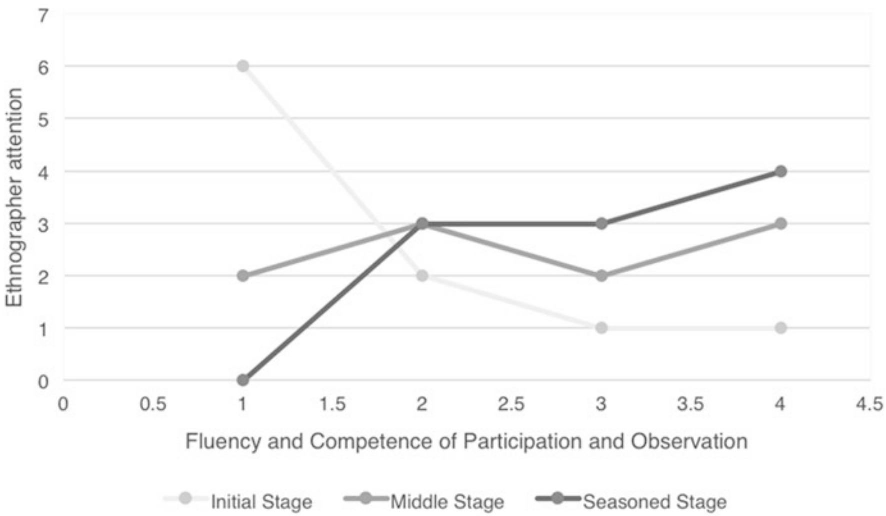


Fig. 16.4 Direction of investments from early to later stages in ethnography

In any form—for instance language or numerical—data face similar judgments as to validity and reliability regardless of the research being classed as qualitative or quantitative or mixed methods (Davis 1992). Validity (data-extrinsic) for qualitative approaches rests on explanatory coherence, rather than being measurable, and reliability (data-intrinsic) particularly through triangulation points to dependability.

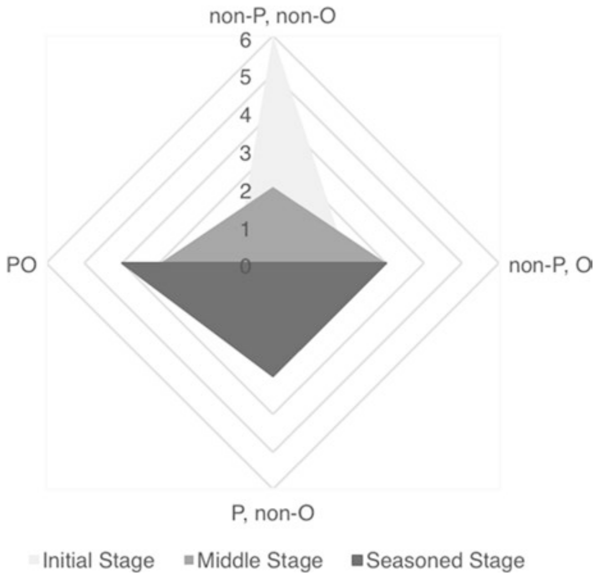


Fig. 16.5 Trajectory of ethnographic competencies

Certain collections of data have also been characterized as rich and "thick" (Geertz 1973), or rich and "fine" (Kuipers and McDermott Conklin 2007, re Conklin), the point being that extended qualitative inquiry will net more data than can be exhausted by analysis in any given period. All these perhaps esoteric considerations aside, no one would argue with the assumption that the seasoned researcher will be the more competent in documentation of both qualitative and quantitative data. The seasoned researcher will also take more risks in apportioning of time and selection of activities.

Estimation and Approximation

This discussion will not deal with ethnomathematical systems and their cognitive and/or linguistic correlates, discerned from ethnography (sensu two). About these, we can confidently expect continuing insights from ordinary ethnography and ethnology (cross-culture comparisons), and especially from the extraordinary cognitive and neurological sciences that now flourish with numerous new sensing aids such as MRIs and CT scans and with new ways to represent all data and their cumulations (Drucker 2014).

Rather I pick out two rudimentary arithmetical methods (at the level of techniques, not methodology), singled out in the above subtitle—estimation and approximation via ballparks and eyeballing—as these can contribute to ethnography (sensu one). Beyond this perhaps trivial detail, I assert that these techniques may be practiced for reasons more important than their coming up with tentative results … of the number of people in a group or the size of a garden plot or the weight of a joint of meat. The ethnographer must remain alert even when nothing is demanding attention. That's when jotting down some numbers can feel very rewarding; it may even be that some of these numbers, if properly grounded, will become meaningful down the line, in and of themselves or when associated with others.

This acknowledgment of the manifold value in making "educated" guesses encounters a complication from research suggesting that children's ability to estimate may associate positively with performance on math achievement tests (Halberda et al. 2008; Halberda 2018; Gallistel and Gelman 2005), this concluding that individuals may be differentially endowed with an ingrained sense of number. This has been contested by cross-cultural research that found no evidence of differences in general arithmetic abilities (Reeve et al. 2018).

An ethnographer who regularly makes serial estimates of any countable environmental variable, whether interesting in and of itself, will be convinced that accuracy (or relative congruence) accrues with experience (and patience).

Estimation and Approximation via Ballparks and Eyeballing

Guessing ranks high in getting anyone through the day, and the ethnographer "counts on" guessing as well, and not just about counting, measuring, weighing, and the like. The guess may even be dressed up as "abduction," freed of the constraints of the other modes of reasoning, induction (from specific to general) and deduction (from general to specific).

The ethnographer will never attain the routinized confidence of insiders in the studied population, but those insiders are handicapped in other ways. Insiders also guess, but without conscious labor, as they can move smoothly through the day relying on habit (West and Anderson 2016). Little may depend on each of the guesses made by either insider or ethnographer, and guesses are apt to be about different things as well. The ethnographer may be aware of being unaware about those routine habits, and even hesitate before activating any part of their body, especially the tongue. This early stage of inhibition should relax as the ethnographer becomes adept in focusing, simultaneously, on the figure as well as on the ground, becoming seasoned in their ethnographic setting.

Theoretically the ethnographer's project will be shaped by theory, that theory coupling with a suitable methodology, in this case ethnography, to focus the data collection. But since context is always relevant, and since the outsider will be tentative in identifying and attending to what is what and what relates to what (and how, not to mention why), the earliest data collected in a project will be somewhat spurious—some in irrelevant detail, some suffering from misplaced concreteness, and other data still obscured by higher order distractions. The visual metaphor of observation inadequately covers what the ethnographer deals with: sights and sounds and other senses in the moving scene, whether immediately interpretable or not, and their absences, their stillnesses, and their silences. All this will be "written down" (and/or typed, recorded, photographed, filmed …), for compilation and analysis and eventual "writing up" as an ethnography.

Up to this point the ethnographer may consider it prudent to capture some numerical data, such as how many persons are in the room/bus/farmyard, even though names may be unknown. This calls for head-counting. Along with any head-counting that is feasible, an estimate will be prudent, in fact, serial estimates, even though persons may not be coming and going. Eyeballing may not seem like a distinguished technique, but any researcher can become as skilled as the journalist, without instruction, particularly as occasionally there may be reassuring confirmations of estimates via actual counting. Often accuracy is not even important, but relative numbers may be at least interesting. Any collection of discrete and probably labeled objects, whether gathered or dispersed, may be so estimated.

The same exercises obtain with regard to "weights and measures," these eventually to respect both the indigenous and the investigator's units of analysis. Weights correspond to mass, while measures apply to lines, areas, and solid volumes.

Eyeballing is more apt to be applied to assemblies of discrete objects, than to three-dimensional objects to be weighed or (typically) two-dimensional expanses to be measured. Handheld objects lend themselves to direct approximation of relative weight, by hefting, and volumes may also be assessed as weighing more or less than an equivalent volume of water, the most familiar of substances. Measurements of modest areal size may submit to approximation by imagining, or deploying, a convenient known unit, perhaps a matchbox or sheet of A4 paper, or for volumes, a "breadbox"; for larger expanses, some Westerners use a "football field," without saying whose "football." Areas will be more confidently estimated in acres or hectares than in thousands of square feet or square meters. In each case, the assumptions of the insiders, whether known or estimated, will also count as data.

In English, "estimation" suggests some inductive effort, while "approximation" suggests some deductive assumptions. One is more apt to be estimating digital units, and approximating more analogue aggregates, sizes, volumes, and weights. Reflecting on either category of informed guess, one gauges whether the result is "in the ballpark," whether or not "eyeballing" was used in the assessment.

Time, though, even in the terms of the ethnographer's units of analysis, presents itself as two separate systems, equally important to document, along with the units of the insiders (Lebebvre 2013). For the Western ethnographer, there will be the "objective" conventional calibrated time dictated by clocks, alongside a subjective experienced time or times, that challenge description, as all qualia do by definition.

Teaching Ethnography: Why We Count, Measure, and Weigh

And why do we usually use another sequence: weigh, count, and measure, if there's evidence for that? Western countries' regulations seem to specify "weights and measures," or "coinage, weights, and measures" (Linklater 2003, p. 103–142). Interesting in this regard is the proliferation of folk units especially when it comes to often perishable odd-shaped or small-grained staples that figure in the commerce of daily life.

I for one question whether anything “interesting” can be ‘“taught.” That would include subjects like swimming, creative writing, biology, and anthropology, especially ethnography. When teaching such a course, I state in the syllabus and discuss in the first class period that, in fact, ethnography cannot be taught, per se, but yet we must at least pretend. Maybe this corresponds with another assertion posted on my office door: Nothing interesting can be defined; corollary, everything is interesting.

While humans without hesitation go about studying things larger and smaller than they are, when it comes to conspecifics (and allomammals in particular), humans studying humans involves nothing less than our languaging studying languaging, culture investigating culture, and cognition researching cognitions. It helps to be able to license the naive habits of the proto-ethnographer, inclusive of senses, perceptions, and lurking cognitive biases. Even unaltered, uncorrected, and unknown biases are worthy of documentation, since when it comes to qualitative research, “closure” may be a culture-bound illusion.

Alongside embracing effortless natural habits, albeit edited along the way, the proto-ethnographer should be willing to exercise even marginal skills in representation by drawing, and documenting through the tools and toys commercially available, but never should technological aids impinge on the “writing down” as mediated by cognition; this interpretation through writing is relatively robust, while sound recording and film documentation, however valuable for expert analysis, are absent in the knowledge of someone present in the context—even though that ethnographer will likely overlook, underinterpret, and overinterpret in the encounter with the living and moving subject matter.

Traditionally the student of ethnography is encouraged to read other ethnographies, but these can be more methodologically opaque than a syllabus on methodology containing instructions that sound like paint-by-number, in other words, that sound doable, and doable by anyone so trained. That sounds like a task without enough challenges to be worth doing!

Asides from an Ethnographic Case History: Counting with Tensho-Kotai-Jingu-Kyo

As a mature but naïve student at the University of Hawaii, I opted to do an ethnography for a B.A. thesis in anthropology. The ethnography was carried out at meetings of a new war-time Japanese religious group in Honolulu, Tensho-Kotai-Jingu-Kyo, the movement arriving to Hawaii in 1952. Tensho is also called Odoru Shukyo, or Dancing Religion, as one of their prayer forms involves bodily movement. Like some other new religions, or sects, arising or being awakened during the hardships of World War II, this group’s charismatic founder was a woman.

Known as the Dancing Goddess, Ogamisama's philosophy evolved through the 1940s, the period during which she was visited by various spirits, and the movement flourished sufficiently to send out missionaries after the war, hence the Honolulu congregation. In the middle of the 1967–1968 academic and ethnographic year, while preaching in Japan, Ogamisama died unexpectedly. This event turned out to be an interesting variable, while my original interest in health and healing in the sect gained no traction once I was familiar with their priorities (Anderson 1968).

I had permission to attend the usually evening meetings that sometimes were more frequent than weekly. Most meetings were 2-h indoor chanting-prayer gatherings, but on third Sundays there were gatherings in a park where they practiced their dance-prayer, each individual wandering with eyes closed, chanting the same (Buddhistic) nam-myoho-renge-kyo. Members of this English-speaking congregation understood that I would be studying them, but no doubt hoped that I would also convert. I was encouraged to sit, that is, kneel, in the first row of meetings often held in a home. I realized that I would be a distraction in the front row, even if I reined in the "writing down." Therefore, I consistently settled in the back row, in a room with up to several dozen other persons, eager to capture absolutely everything, especially if numerical.

I compulsively counted (nothing available to be measured or weighed) every variable imaginable (participants' age, sex, ethnicity, occupation, history in sect …), and every inferred category of voluntary act or even word in speech acts [testimonies classed as confessions, recollections, exhortations; words such as Ogamisama, Wakagamisama (her son), Himegamisama (her granddaughter and successor Goddess), "god"], and not to overlook documenting presumably involuntary indications of possession by good or dangerous spirits.

This is not about guessing. As a former actuary and sometimes student of statistics I was after numbers, even if I could never squeeze any stats out of them. At the same time, I did not feel at liberty to keep a notepad out at all times, so I concentrated on memorizing everything I would have otherwise written down. But I couldn't figure out how to remember all the numbers I was intent on gathering, unless I resorted to wearing a lightweight fly-fisherman's vest.

There was no room in the thesis for all the data I collected, whether narrative or numerical, and not even place to detail how I used the vest. I cannot confidently recite at this time, 50 years later, exactly which object in which pocket would undergo manipulation or relocation to register what behavior by which class of person, as those records are not accessible. This slender description will convey the basics.

The "glass beads" in this "scientific" investigation were both natural seeds (often diverse dried legumes) and manufactured pastas (for example, long spaghetti or noodles that could be gently broken while remaining in the same pocket), to tally particular events or behaviors. The congregation broke down into male leaders, other males, and females. Each category was also represented by a distinctive bean, and each particular behavior would dictate the relocation (to pockets above or below, on the same or the other side of the vest). Because each category of participant (numbering, say, 1 to 30-plus) could chalk up dozens of acts in any category, it was important to have a good supply of each bean at the outset. At home after each 2-h meeting, though exhausted, I had to carefully excavate each pocket to record all the events captured by each category of congregants, and also others that were more general. But first I had to type up all the narrative notes I would ordinarily have written down, leaving space for expansions for future recall. Only then, the bean-counting is done, defending the calculations against resident cats.

I also kept track of the length of the verbal testimonials, and the incidence of some key terms, and these I would remember until it was easy to jot them down inconspicuously. Shorthand is a defense against roving eyes, besides being expedient. Each meeting room had a large round analogue clock on the wall; the leader of each prayer-chant—10 min at the inception and close of each 2-h meeting—would be the only one looking at this clock. After 10 min, the leader would clap twice, the chanting would cease as the congregation clapped back, twice. If chanting continued, it could be that someone had been captured by an evil spirit, and the congregation returned to louder chanting to bring the afflicted person back to consciousness.

These data were natural to inspect for any change in behavior after the death of the Dancing Goddess. For instance, the average-per-meeting incidence of the utterance of "Ogamisama" after her death went up 125%; the increase for "god," however, went up 557%. This would be more interesting, had I similar data from the Japanese-speaking congregation that I also only occasionally visited. Overall, these tabulations also confirmed any impression that the male leaders were manipulating their public (although supposedly everyone kept their eyes closed), and perhaps were deceiving themselves, too, by being possessed by only good spirits, and rather frequently!

In this ethnography, by a naïve and serious ethnographer, the collection of numerical data was an end in itself. Keeping indices of the data in movement inside or between pockets, and then holding both qualitative and other quantitative data in memory, the arduous process left no room for guessing, estimating, or approximating. The size of the group, though, being modest, allowed for the precision, however spurious these data may be. It's more relaxing, perhaps, to undertake projects in a whole village, when eyeballing and ballparks can take up the slack of natural limits to ethnographic finesse.

Coda

Leaving aside this introduction to the tender role of ethnographer as proto-mathematician, some social scientists and mathematicians have actually made a case for intuitive judgments and quick approximations (Gigerenzer and Brighton 2009; Kahneman 2011; Nisbett and Ross 1980); sometimes these are inevitable and not just expedient. Unaided by ample samples for statistical penetration, however, humans are handicapped when it comes to discerning probabilities, "causation," correlation, and more convoluted relationships (E.N. Anderson 1996, 2013; Kahneman et al. 1982; Piattelli-Palmarini 1994). Humans get by in judgments of similarity and difference, including interpreting metaphor and simile, despite a tendency to under-predict the amount of variability

(Tversky 1977). Tversky also points out that humans can perceive things that do not appear in the data at all, for example clusters, patterns, and winning/losing streaks, as when gambling.

Being so undisciplined about their disciplines, it is not surprising that students and researchers can be so easily recruited to elevate the simpler quantitative at the expense of the ephemeral qualitative. In privileging the quantitative, many assume it to be the more "scientific" approach for inquiry (delivering data) and for convincing meaning-making (kneading those data statistically); they also assume the quantitative to be the more difficult, when in fact it's quite the opposite. The quantitative researcher starts, and ends, outside of the subject/object of study; from that vantage point, a refined set of observations is selected to elucidate a finite number of questions, or even hypotheses (Anderson 2012, p. 299).

Acknowledgments The quality of figures and text was greatly enhanced with the assistance of my former student, Jamie Kruis.

References

Aiden, E. and Michel, J. B. (2013). *Uncharted: Big data as a lens on human culture*. New York: Riverhead Books.

Anderson, E. N. (1996). *Ecologies of the heart: Emotion, belief, and the environment*. New York: Oxford University Press.

Anderson, E. N. (2013). What shapes cognition? Traditional sciences and modern international science. In: M. Quinlan and D. Leopfsky (ed.), *Explorations in ethnobiology: The legacy of Amadeo Rea*, 47–77. Denton, Texas: Society of Ethnobiology.

Anderson, M. (1968). *Impact of the death of a charismatic leader, incorporating: An ethnography of Tensho-Kotai-Jingu-Kyo*. B. A. honors thesis in anthropology, University of Hawaii, Honolulu.

Anderson, M. (1999–2000). Anthropology as translation. *Athanor* 10 (2): 181–187.

Anderson, M. (2012.) How qualification and quantification meet, or don't, in ethnography. In: M. Bockarova, M. Danesi, and R. Núñez (eds.). *Semiotic and cognitive science essays on the nature of mathematics*, pp. 296–329. Munich: Lincom Europa.

Anderson, M., Sáenz-Ludlow, A., Zellweger, S. and Cifarelli, V. V. (2003). *Educational perspectives on mathematics as semiosis: From thinking to interpreting to knowing*. New York: Legas Publishing.

Anderson, M. and Gorlée, D. L. (2011). Duologue in the familiar and the strange: Translatability, translating, translation. In: K. Haworth, J. Hogue, and L. G. Sbrocchi (eds.), *Semiotics 2010*, pp. 217–228. Ottawa: Legas Press.

Ascher, M. (2002[1991]). *Mathematics elsewhere: An explanation of ideas across cultures*. Princeton: Princeton University Press.

Bateson, G. (1978). Number is different from quantity. *CoEvolution Quarterly* 18: 44–46.

Berland, J. C. (1982). *No five fingers are alike: Cognitive amplifiers in social context*. Cambridge: Harvard University Press.

Conklin, H. C. (2007). Fine description. In: J. C. Kuipers and R. McDermott (eds.), *Ethnographic and linguistic essays*. New Haven: Yale University Southeast Asia Studies Monograph Series.

Conner, S. (2016). *Living by numbers: In defence of quantity*. London: Reaktion Books.

D'Ambrosio, U. (1997). Foreword. In: A. B. Powell and M. Frankenstein (eds.), *Ethnomathematics: Challenging eurocentrism in mathematics education*, pp. 13–24. Albany: State University of New York Press.

Davis, K. A. (1992). Validity and reliability in qualitative research on second language acquisition and teaching. *TESOL Quarterly* 26: 605–608.

Deely, J. (2012). The semiosis of mathematical thought. In: M. Bockarova, M. Danesi, and R. Núñez (eds.). *Semiotic and cognitive science essays on the nature of mathematics*, pp. 211–224. Munich: Lincom Europa.

Drucker, J. (2014). Graphesis: *Visual forms of knowledge production*. Cambridge: Harvard University Press.

Fabbri, M. (2013). Finger counting habits and spatial-numerical association in horizontal and vertical orientations. *Journal of Cognition and Culture* 13 (1–2): 1568–5373.

Gallistel, C. R. and Gelman, R. (2005). Mathematical cognition. In: K. J. Holyoak and R. G. Morrison (eds.), *The Cambridge handbook of thinking and reasoning*, pp. 559–588. Cambridge: Cambridge University Press.

Geertz, C. (1973). *The interpretation of cultures: Selected essays by Clifford Geertz*. New York: Basic Books.

Gigerenzer, G. and Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science* 1 (1): 107–143.

Halberda, J., Mazzocco, M. M., Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455 (7213): 665–668.

Halberda, J. (2018) Logic in babies. Science 359[6381]: 1214–1215.

Halmos, P. R. (1984). Applied mathematics is bad mathematics. In: D. M. Campbell and J. C. Higgins (eds.), *Mathematics: People, problems, results*, Volume 3, pp. 193–200. Belmont, California: Wadsworth International.

Hill, J. H. and Mannheim, B. (1992). Language and world view. *Annual Review of Anthropology* 21: 381–406.

Ingold, T. (2000). *The Perception of the environment*. New York: Routledge.

Johnson, R. B., Onwuegbuzie, A. J., and Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research* 1: 112–133.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press

Kauffman, L. H. (n.d.). What is a number? 13-page document, Mathematics Department, University of Illinois at Chicago.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge: Cambridge University Press.

Lebebvre, H. (2013). *Rhythmanalysis: Space, time, and everyday life*, trans. G. Moore and S. Elden. London: Bloomsbury Academic.

Linklater, A. (2003). *Measuring America: How the United States was shaped by the greatest land sale in history*. New York: Plume Publishing.

Ludwig, D. (2018). Does cognition still matter in ethnobiology? *Ethnobiology Letters* 9 (2): 269–275.

Nisbett, R. E. and Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, New Jersey: Prentice-Hall.

Noritake, A., Ninomiya, T. and Isoda, M. (2018). Social reward monitoring and valuation in the macaque brain. *Nature Neuroscience* 21: 1452–1462.

Overmann, K. A. (2015). Numerosity structures the expression of quantity in lexical numbers and grammatical number. *Current Anthropology* 56 (5): 638–653.

Piattelli-Palmarini, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds*. New York: Wiley.

Reeve, R. A., Reynolds, F., Paul, J., and Butterworth, B. L. (2018). Culture-independent prerequisites for early arithmetic. *Psychological Science* 22: 1–10.

Selin, H. (ed.) (2000). *Mathematics across cultures: The history of non-western mathematics*. Dordrecht: Kluwer Academic Publishers.

Silver, Daniel S. (2017) The new language of mathematics. *American Scientist* 105.6: 364–371.

Shermer, M. (2019). Stein's law and science's mission. *Scientific American* 320 (1): 67.

Smit, B. and Onwuegbuzie, A. J. (2018). Observations in qualitative inquiry: When what you see is not what you see. *International Journal of Qualitative Methods* 17: 1–3.

Tenenbaum, G., Gershgoren, L., and Schinke, R. J. (2011). Non-numerical data as data: A positivistic perspective. *Qualitative Research in Sport, Exercise, and Health* 3 (3): 349–361.

Tversky, A. (1977). Features of similarity. *Psychological Review* 84 (4): 327–352.

Van Bendegem, J. P. and Van Kerkhove, B. (2004). The unreasonable richness of mathematics. *Journal of Cognition and Culture* 4 (3): 525–549.

Waugh, L. L. (1982). Marked and unmarked: A choice between unequals in semiotic structure. *Semiotica* 39: 299–318.

Weissner, P. (2016). The rift between science and humanism: What's data got to do with it? *Current Anthropology* 57: 154–166.

Werner, O. J. and Schoepfler, G. M. (1987). *Systematic fieldwork: Foundations of ethnography and interviewing*, Volume 1. Thousand Oaks, California: Sage Publications.

West, D. E. and Anderson, M. (eds.) (2016). *Consensus on Peirce's concept of habit: Before and beyond consciousness* New York: Springer.

West, G. (2017). *Scale: The universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies*. New York: Penguin.

Wilder, R. L. (1981). *Mathematics as a cultural system*. Oxford: Pergamon.

Zavlasky, C. (1999/1979/1973). *Africa counts: Number and pattern in African cultures*. 3rd edition. Chicago: Chicago Review Books.

Chapter 17
"A Mathematician, a Physicist and an Engineer": The Meaning of "M" in STEM

Dragana Martinovic

Introduction

In the recent decades, the "STEM" acronym became a consistent part of different governments' calls for revamping education, so that it equips workforce with skills appropriate for the new century. While the original intention was to increase enrollment into the science, technology, engineering, and mathematics (STEM) fields of study, in time, it opened up discussions about the nature of STEM-related skills, and a possible creation of the unified field of study and its place in the school curriculum.

To this end, some recommended fully integrating "mathematical methods in science and scientific methods into mathematics such that it becomes indistinguishable as to whether [the subject] is mathematics or science" (Berlin and White 1992, p. 78). Technology could support this integration, especially in cases when mathematics and science concepts are misaligned in the curricula (e.g., students usually learn the physics of motion before they study differential calculus, but with graphics calculators and motion detectors they can conceptualize displacement, velocity, and acceleration without fully understanding the mathematics behind the formulas). Also, the project-based learning may be an appropriate pedagogy for integrating subjects, since:

> projects, when they encourage thoughtful student exploration, neatly integrate mathematics and science. This is not too surprising, since many projects are based on interesting situations which require a quantitative understanding of cause and effect, that is, they require a mathematical treatment of science topics. (Tinker 1992, p. 51)

D. Martinovic (✉)
University of Windsor, Windsor, ON, Canada
e-mail: dragana@uwindsor.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*, Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_17

The arguments against the integration of science with mathematics are provided by those who fear that "merging of the disciplines might cause people to lose important philosophical, methodological, and historical differences between the two subjects" (Berlin and White 1992, p. 78). As a mathematician and mathematics educator, I am particularly interested in the position of mathematics in the integrated STEM education—the meaning of "M" in STEM. In this chapter, I will explore a folk view of the STEM fields, such that emphasizes their differences, and compare it to the ideas of STEM as a unified field of study. By comparing the epistemological and educational perspectives, I hope to arrive at a more comprehensive understanding of what STEM really is and to provide some recommendations. As the resources for a folk view, I used a limited number of jokes posted on the Internet, and for the education view the literature that deals with STEM education.

Folk View of the STEM Fields: "A Mathematician, a Physicist, and an Engineer" Jokes

The Internet is full of jokes about mathematicians and other professions (e.g., scientists, engineers, doctors, and accountants). Gilkey (1990, p. 215) calls them a "professional slur," as they represent professions in a stereotyped way, but mentions that the originators of these jokes believe that there is some truth in them. Kessel (2013, p. 15) is of the opinion that "this helps to orient mathematicians to features of their own discipline by contrasting them with aspects of other sciences."

The jokes that I present in this chapter were probably written by mathematicians or some witty individuals who spent many hours sitting in the postsecondary mathematics classes. They are not always easy to understand by nonmathematicians, and as such demonstrate some pride in belonging to an elite group that speaks and thinks in its own way. For Renteln and Dundes (2005, p. 34), "the tension between the mathematical universe and the nonmathematical universe … is central to much of mathematical humor," emphasizing that mathematicians live in their own world and think in a way different from other professions.

In the interest of space, I selected four jokes that are exemplary of aspects in which thinking of mathematicians differs from that of professionals in other STEM fields; I named these aspects: (a) points of view, (b) dealing with redundancies, (c) precision of language, and (d) proofs are mathematicians' work.

Joke 1: Points of View

> An engineer, a physicist, and a mathematician are trying to set up a fenced-in area for some sheep, but they have a limited amount of building material. The engineer gets up first and makes a square fence with the material, reasoning that it's a pretty good working solution. "No, no," says the physicist, "there's a better way." He takes the fence and makes a circular

pen, showing how it encompasses the maximum possible space with the given material. Then the mathematician speaks up: "No, no, there's an even better way." To the others' amusement he proceeds to construct a little tiny fence around himself, then declares: "I define myself to be on the outside."

In this joke, a mathematician's thinking is different from that of an engineer and a physicist—while his solution mathematically makes sense, it is completely detached from the reality. Also, while the engineer is led by practicality, the physicist seeks the best solution under the circumstances (see Fig. 17.1).

Probably because building the largest enclosed area was not explicitly required, the engineer in this joke went for practicality, rather than for optimal use of the material. Building a squared fence is pretty straightforward—the fence material could be stacked into four equal piles, one for each side, and the right angles between the sides could be easily determined using different tools. The physicist was not satisfied with this solution. He knew that when a square and a circle have the same border length, the area of the circle is larger by approximately 27%. Knowing the length of the fence (i.e., a circumference of the future circular sheep-pen), it would be enough to divide it by $2\pi\sim 44/7$ to obtain the approximate value of the radius. After determining the center of the circle, the construction could proceed—all the fence material would be used to create the sheep-pen enclosing the maximum area. The mathematician worked conceptually. He modeled the situation, so that the Earth became a 3-D closed, connected, and triangulable surface (Zeeman 1966) and a sheep-pen, a closed curve on this surface. Since the closed fence would separate the surface of the Earth (i.e., "A curve is said to separate [the surface] M if cutting along the curve causes M to fall into two pieces," p. 14), this would create two areas—one inside and another outside the fence. Then he chose to put the sheep in the larger area of the two.

The funny part of this joke is that for the mathematician, who obviously knows his discipline, only he and the sheep existed in this task, so as long as they were in the different regions, the task was completed. He did not take into account the purpose of fencing—to protect sheep and to allow the farmer easy access to them. But

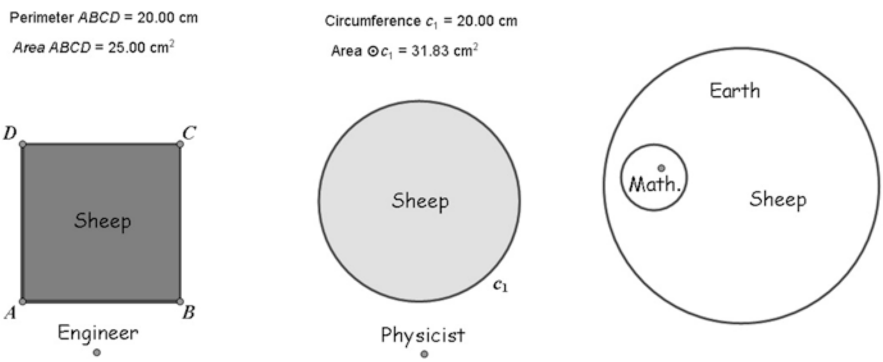


Fig. 17.1 Solutions offered by an engineer, a physicist, and a mathematician

those “details” are not of the concern for mathematicians—in an ideal world, their solutions work ideally! Also, he outsmarted the engineer and the physicist by saving on the fence material.

Joke 2: Dealing with Redundancies

> How does a mathematician or a physicist prepare tea? Both of them pour water into the pot, light the flame, boil the water, and infuse the leaves. What is the difference in the solution, if the pot is already filled with water? The physicist lights the flame, boils the water, and infuses the leaves. The mathematician pours the water out, thereby reducing the problem to the previous, already solved one.

This joke is very relevant to how mathematicians approach problems. They know that once a problem is solved, they can apply its result and/or method in solving another problem (Pólya 2004). This is an important heuristic used in problem-solving which Pólya (2004, p. 110) mentions in “Here is a problem related to yours and solved before.” In this strategy, it is important to determine how the two problems are alike and how they are different, since “trying to link up the two problems…we introduce into the new problem elements corresponding to certain important elements of the old problem” (p. 111). Indeed, mathematicians avoid redundant actions; once they have an agreed-upon knowledge (e.g., axioms, and proved lemmas and theorems), they could use it in new situations, deeming it unnecessary to present how this knowledge was acquired in the first place. However, in this joke, the physicist’s approach is both rational and correct, while the mathematician’s approach is silly; (h)she is conducting a pointless as well as a wasteful action (i.e., first pouring the water out of the pot), while being guided by the principle of avoiding unnecessary effort (see Fig. 17.2).

This is not how mathematicians act. If one starts from the last step, which is to infuse tea leaves into hot water, the steps leading to it could be considered auxiliary elements in Polya’s heuristics for problem-solving. The mathematician can look into conditions leading to this step, one of which is “having the tea leaves,” and the other, “having the boiling water.” The mathematician would then conclude that (h)she needs to boil the water, for which the flame is needed as well as the pot with water. In other words, the mathematician would not pour out the water, but recognize that the problem (h)she has is subservient to the problem solved. Using a “unilateral reduction” of the “more ambitious” (p. 56) solved problem, the mathematician would prepare the tea in the same way as the physicist in this joke.

Joke 3: Precision of Language

> An astronomer, a physicist and a mathematician (it is said) were holidaying in Scotland. Glancing from a train window, they observed a black sheep in the middle of a field. “How interesting,” observed the astronomer, “All Scottish sheep are black!” To which

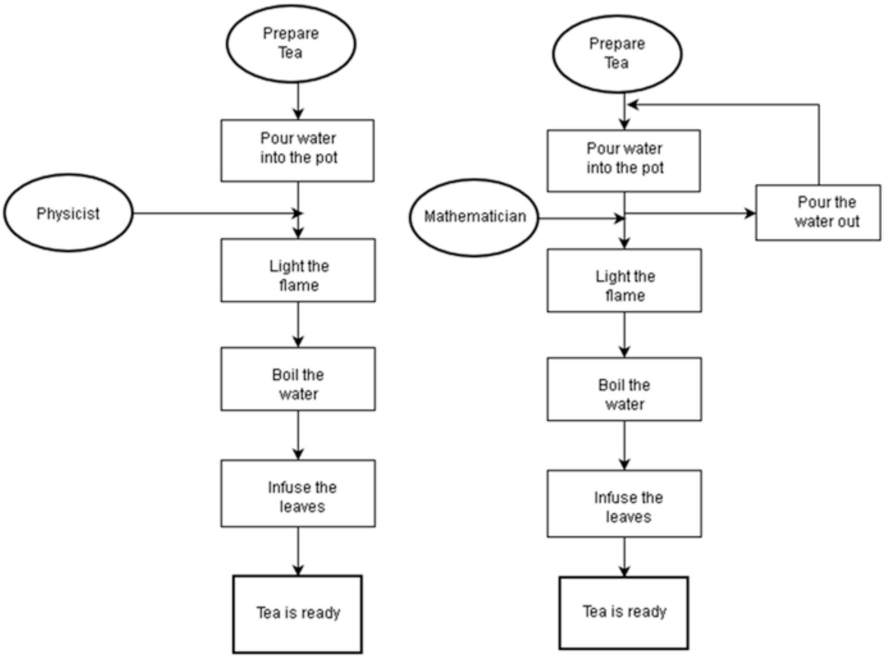


Fig. 17.2 The physicist's and the mathematician's approach when the pot is full

> the physicist responded, "No, no! *Some* Scottish sheep are black!" The mathematician gazed heavenward in supplication, and then intoned, "In Scotland there exists at least one field, containing at least one sheep, *at least one side of which is black*." (Stewart 1995, cited in Kessel 2013, p. 243, italics in original)

This popular joke can be found not only on different Internet websites but also in academic papers and books. It presents mathematicians as obsessed with linguistic precision and accuracy. In their reasoning, mathematicians are very careful to distinguish facts from assertions. Here, the astronomer overgeneralizes, and the physicist makes a reasonable inference, while the mathematician's conclusion sounds ridiculously formal. I believe that this aspect is what makes this joke funny—the mathematician is deliberately using the overly precise language in order to prevent it being "'contaminated' by real-world knowledge" (Schoenfeld 1992, p. 347), confirming that mathematicians live in their own world.

In everyday communication and especially in the context of "holidaying," such language simply does not have place. Halmos (1970, p. 142) advises that "The symbolism of formal logic is indispensable in the discussion of the logic of mathematics, but used as a means of transmitting ideas from one mortal to another it becomes a cumbersome code." The same holds when using language of formal logic in informal speech. The mathematician is reasoning about a real-live situation in a way appropriate for reasoning about some mathematical structure. The distinction between the two is epistemologically what distinguishes science from mathematics.

In their development of ontology of mathematical objects, Font, Godino, and Gallardo (2013, p. 113–114) describe the difference between material and nonmaterial objects as follows:

> Objects outside mathematics, such as oranges, trees, etc. are considered to be particular and to have a material existence [in time and space]. This type of existence means that they are ostensive in the sense that they can be shown directly to another person. … By contrast, this kind of existence is not attributed to non-ostensive objects, which are usually considered to have an ideal existence.

Language (oral, written, or gestural) allows us to describe, define, and represent constructs that exist in our minds; in that way, non-ostensive objects materialize and become sharable with others, workable, and explainable. Font et al. (2013, p. 101) assert that the language of mathematics leads to reification—treating ideas as real—"the process by which we assume, or state linguistically, that there is an object with various properties or various representations." It strikes me in this joke how the mathematician treats the sheep as a non-ostensive object; (h)she describes it stripped from any preconceived notion based on the lived or learned experience. The sheep has sides (and we do not know how many), one of which is black, but there is nothing more to say about it.

Similar to Font et al., Bertrand Russell (1938, Preface) distinguishes "actual" from "hypothetical" objects and writes how:

> In pure mathematics, actual objects in the world of existence will never be in question, but only hypothetical objects having those general properties upon which depends whatever deduction is being considered; and these general properties will always be expressible in terms of the fundamental concepts which I have called logical constants. Thus when space or motion is spoken of in pure mathematics, it is not actual space or actual motion, as we know them in experience, that are spoken of, but any entity possessing those abstract general properties of space or motion that are employed in the reasonings of geometry or dynamics.

For our understanding of epistemological differences between different fields and especially the specificities of mathematics, remembering Russell's words is crucial.

Joke 4: Proofs Are Mathematicians' Work

> A mathematician, a physicist and an engineer are given an identical problem: Prove that all odd numbers greater than 2 are prime numbers. They proceed:
>
> Mathematician: 3 is a prime, 5 is a prime, 7 is a prime, 9 is not a prime —counterexample—claim is false.
>
> Physicist: 3 is a prime, 5 is a prime, 7 is a prime, 9 is an experimental error, 11 is a prime,…
>
> Engineer: 3 is a prime, 5 is a prime, 7 is a prime, 9 is a prime, 11 is a prime,...

This joke is about how different professionals understand and conduct proofs. It presents the mathematician as both "precise and correct …, the physicist [as] willing and eager to overlook error in the name of bigger truth [… and the] engineer [as] ignorant in mathematical matters" (Gilkey 1990, p. 216). Proofs, as Bakker and Hußmann (2017, p. 399) would say, reveal that "reason and inference are the bread and butter of mathematics." Proofs and refutations go hand in hand, as they both contribute to mathematical knowledge. Mathematicians constantly invent and validate statements; if they cannot prove that the statements are correct and consistent with the existing knowledge, then they try to understand in which way they are flawed. One way of doing so is to find a counterexample—an example for which the statement is not true. Stylianides and Al-Murani (2010, p. 21) explain that:

> The process of validating assertions (and mathematical knowledge more broadly) often follows a 'zig-zag' path between attempts to generate proofs for the truth of the assertions and the discovery of counterexamples that refute the assertions and necessitate their refinement before they can be subjected to new proving attempts … A fundamental idea that underpins this validation process is that it is not possible to have a proof and a counterexample for the same assertion.

The funny part of this joke is that one does not have to be very skilled in mathematics to understand that 9 is an odd number that is not prime (i.e., 9 = 3 times 3). This refutes the claim that "all odd numbers greater than 2 are prime." However, neither the physicist nor the engineer refute this statement. How is that possible? Stylianides and Al-Murani found that even good undergraduate and high school students may have very limited understanding of proofs and that some may think that the same statement can simultaneously be true and have a counterexample. The difficulty here may be that, on one side, mathematicians argue that empirical arguments do not constitute proofs because one cannot generalize something based on a proper subset of all possible cases (e.g., "all odd numbers are prime, because 3 and 5 and 7 are such"). Yet, on the other side, the refutation of a statement by finding a counterexample is an accepted technique. For those not skilled enough in mathematical reasoning, this may even appear as constituting a double standard, and consequently can result in thinking that it is possible to have a proof and a counterexample for the same statement.

In this joke, the physicist is presented as a scientist who works with empirical data. Without finding more evidence that the claim is false, his/her first inclination is to suspect that the data were contaminated in some way. The engineer made an oversight, but who knows, errors do not always have big consequences.

Summary

Except for jokes #1 and #4, #2–#3 do not appear to be connected to mathematics, and yet they provide some insight into how mathematicians think. They suggest that, even in casual conversations, mathematicians see "phenomena in mathematical terms" (Schoenfeld 1992, p. 341). Other jokes that Gilkey (1990, p. 219–220) refers to present engineers as "competent in action," mathematicians as "[too precise] in theoretical matters" (p. 219), and physicists somewhere in between, "[unable to] be pragmatic if theory dictates otherwise." For Schoenfeld (1992), enculturation is an important aspect of becoming a professional. He writes that, "The case can be made that a fundamental component of thinking mathematically is having a mathematical point of view—seeing the world in ways like mathematicians do" (p. 340). Indeed, this can be said about other professions as well, providing the base for professional jokes to develop. According to Kessel (2013, p. 250):

> Being "trained by the discipline" involves oral traditions, jokes, and experiences with mathematics. These experiences shape a mathematical perspective which includes views about shared characteristics of the natural sciences and scientific methods as well as distinctions among the sciences. It includes use of heuristics (e.g., reducing a problem to a previously solved problem), precision in use of terms, care with notation, and familiarity with structure-preserving correspondences that coordinate concepts and representations…Precision in definitions helps to delineate the scope of a theory. Details matter, slight differences in wording matter, and assumptions matter.

Examples that we reconstructed in this section point to the specificity of mathematical reasoning and, probably as well, that professionals from different disciplines take pride in their idiosyncratic traditions. So then, how would integrating the STEM fields look like? What would be the place of mathematics in STEM?

Fitting "M" into STEM

The acronym "STEM education" has been in use for a couple of decades already, but its agreed-upon meaning is still lacking. While mathematics and science education are well established in national curricula, the roles and contents of technology and engineering education are much less established. T(echnology) education may mean educational technology that is applicable across the subjects, or may mean computer science education. But then would not computer science be part of S(cience)? Also, some even define mathematics as S(cience), so why would these two be separated? The meaning of E(ngineering) in STEM is even less clear (Assefa and Rorissa 2013), especially in recent calls to start with STEM education as early as in kindergarten. Governments consistently use this acronym in their calls for providing students with the twenty-first-century skills, but many advise moving away from the acronyms (e.g., STEM, STEAM, …), as they are too limiting.

In these debates, the general public remains confused or oppositional. Keefe (2009) found that by STEM, her survey participants mostly understand stem cell research or plant stems. In a survey conducted by Breiner et al. (2012), out of 222 faculty members at the University of Cincinnati, 25% did not know the meaning of the acronym, while about 10% thought that "M" may stand for medicine, mathematics, music, or even management. The researchers concluded that the survey results point to "a challenge in changing the paradigm from compartmentalizing academic disciplines to the integration of these disciplines as advocated by many through the STEM movement" (p. 9). Also, it seems that even the academics, who are probably following media and government calls to enhance STEM-related skills, do not always clearly associate mathematics with other STEM fields.

Mathematicians, like members of other disciplines, have their own epistemological views about the nature, limits, methods, and organization of knowledge. These views are used in judging knowledge of their own kind or that of other disciplines. Also, it is important to understand how to achieve disciplinary literacy as a way to read, write, speak, think, and investigate in the discipline. This prompted Spires et al. (2018, p. 1428) to ask, "If we do not expect students to become experts in each discipline, what should we expect of elementary students, middle grades students, and high school students?" Their study showed that science and mathematics understand literacy differently. Mathematicians have to explain their reasoning using the dense and rigorous language of mathematics, to use models and different representations, and to avoid redundancy, while scientists make connections between data and conclusions within their own fields (e.g., physics, chemistry, biology). Both professions require analytical literacy.

In science as well as in mathematics, there is a distinction between theoretical and applied discipline, rarely mentioning that there may be an engineering part. Frezza et al. (2013) propose distinguishing the "science" of computer science from its "engineering" aspects. They argue that "science aims to explain [i.e., provide universal, reliable, comprehensive and sufficiently precise formulation of knowledge] and technology/engineering aims to create artifices (complete them in a timely manner, with sufficient precision and comprehension]" (p. 1). Although Frezza et al. in many instances interchangeably use the terms "engineering" and "technology," they highlight that the knowledge that engineers have is very specific; it is codified and catalogued in the form of reference books, which they use to solve very specific problems of practice. As Horgan (2013, para. 19) would say:

> [Engineers] don't seek 'the truth,' a unique and universal explanation of a phenomenon or solution to a problem … They seek merely answers to specific, localized, temporary problems, whether building a bridge with less steel or a more efficient solar panel or a smartphone with a bigger memory. Whatever works, works.

When teaching in an integrated STEM curriculum, mathematics may be perceived as a tool for science, engineering, and technology—as a service discipline. Stephen Hawking (2005, p. xi–xii) reminds us that:

> All through the ages, no intellectual endeavor has been more important to those studying physical science than has the field of mathematics. But mathematics is more than a tool and language for science. It is also an end in itself, and as such, it has, over the centuries, affected our worldview in its own right …. the brilliance of the Greeks [was] to recognize the importance of principle plura [fundamental principles] in mathematics, and that in its essence mathematics is a subject in which one begins with a set of concepts and rules and then rigorously works out their precise consequences.

To conclude, nowadays, instructors are challenged, not only when they are expected to integrate the STEM subjects, but also when they are faced with students who have increasingly gone through individualized curricula, which may contain a bare minimum of required courses in addition to a medley of "breadth" or filler courses. Barnett (2003, p. 3) poses that in the era of uncertainty of what the future may bring, we are faced with multiple cognitive, epistemological, and ontological challenges so that:

> [W]e can no longer be sure of our identity. In the contemporary world, what is a doctor? What is a professor? What, even, is a university? As knowing subjects, our hold on the world is loosened, if not broken apart all together.

According to Donald (2009, p. 48–49), "the disciplines tell us that we [as instructors] have to tread carefully,… although there are commonalities in the way we think, the philosophies under which the disciplines operate are distinct and require different navigation patterns." She proposes that courses address aspects such as "what questions does the discipline ask and how are these questions related to those asked in other fields? How does the expert in the discipline function?" (p. 48). While looking within and across disciplinary boundaries would help students and instructors alike to freely explore the possibilities of acquiring knowledge, it is important not to lose sight of social and intellectual specificities of each discipline, as the jokes above clearly point out.

References

Assefa, S. G. and Rorissa, A. (2013). A bibliometric mapping of the structure of STEM education using co-word analysis. *Journal of the American Society for Information Science and Technology* 64(12): 2513–2536.

Bakker, A. and Hußmann, S. (2017). Inferentialism in mathematics education: Introduction to a special issue. *Mathematics Education Research Journal* 29: 395–401.

Barnett, R. (2003). Learning for an unknown future, in Learning for an Unknown Future, *Proceedings of the 26*th *HERDSA Annual Conference*, Christchurch, New Zealand, 6–9 July 2003: pp 3.

Berlin, D. and White, A. (1992). Report from the NSF/SSMA Wingspread conference: A network for integrated science and mathematics teaching and learning. In: D. Berlin (ed.), *A Network for Integrated Science and Mathematics Teaching and Learning Conference Plenary Papers*, pp. 75–81. NSF/SSMA Wingspread Conference (Racine, Wisconsin, April 1991). School Science and Mathematics Association Topics for Teachers Series Number 7. Columbus, OH: National Center for Science Teaching and Learning.

Breiner, J. M., Harkness, S. S., Johnson, C. C., and Koehler, C. M. (2012). What is STEM? A discussion about conceptions of STEM in education and partnerships. *School Science and Mathematics* 112: 3–11.

Donald, J. G. (2009). The Commons: Disciplinary and interdisciplinary encounters. In: C. Kreber (ed.), *The university and its disciplines: Teaching and learning within and beyond disciplinary boundaries*, pp.35–49. New York and Abingdon: Routledge.

Font, V., Godino, J. D., and Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics* 82: 97–124.

Frezza, S. T., Moodey, R. W., Nordquest, D. A., and Pilla, K. (2013). Applying a knowledge-generation epistemological approach to computer science and software engineering pedagogy. 120th *ASEE Annual Conference & Exposition*, June 23-26, 2013, Atlanta, Georgia, Paper ID #6570.

Gilkey, C. F. (1990). The physicist, the mathematician and the engineer: Scientists and the professional slur. *Western Folklore* 49 (2): 215–220.

Halmos, P. R. (1970). How to write mathematics. *L’Enseignement Mathématique* 16: 123–152.
Hawking, S. (2005). *God created the integers: The mathematical breakthroughs that changed history*. Philadelphia, PE: Running Press.
Horgan, J. (2013). Is “social science” an oxymoron? Will that ever change? Retrieved from https://blogs.scientificamerican.com/cross-check/is-social-science-an-oxymoron-will-that-ever-change/
Keefe, B. (2009). *The perception of STEM: Analysis, issues and future directions*. Entertainment and Media Communication Institute, Division of Entertainment Industries Council, Inc. (EIC). Burbank, CA: EIC.
Kessel, C. (2013). A mathematical perspective on educational research. In: Y. Li and J. N. Moschkovich (eds.), *Proficiency and beliefs in learning and teaching mathematics: Learning from experts*. Rotterdam: Sense Publishers.
Pólya, G. (2004). *How to solve it: A new aspect of mathematical method*, 2nd ed. Princeton, NJ: Princeton University Press.
Renteln, P. and Dundes, A. (2005). Foolproof: A sampling of mathematical folk humor. Notices of the AMS, 52(1):24–34.
Russell, B. (1938). *The principles of mathematics*, 2nd ed. New York: Norton.
Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In: D. Grouws (ed.), *Handbook for research on mathematics teaching and learning*, pp. 334–370. New York: Macmillan.
Spires, H. A., Kerkhoff, S. N., Graham, A. C. K., Thompson, I., and Lee, J. K. (2018). Operationalizing and validating disciplinary literacy in secondary education. *Reading and Writing* 31 (6): 1401–1434.
Stewart, I. (1995). *Concepts of modern mathematics*. Courier Corporation.
Stylianides, A. J. and Al-Murani, T. (2010). Can a proof and a counterexample coexist? Students’ conceptions about the relationship between proof and refutation. *Research in Mathematics Education* 12(1): 21–36.
Tinker, R. F. (1992). Integrating mathematics and science. In: D. Berlin (ed.), *A network for integrated science and mathematics teaching and learning conference plenary papers*, pp. 49–51. NSF/SSMA Wingspread Conference (Racine, Wisconsin, April 1991). School Science and Mathematics Association Topics for Teachers Series Number 7. Columbus, OH: National Center for Science Teaching and Learning.
Zeeman, E. C. (1966). *An introduction to topology: The Classification Theorem for Surfaces*. University of Warwick.

Chapter 18
Why a Duck?: A Three-Part Essay on the Mathematics of Cognition

Yair Neuman

Part 1

If one would like to sense the human mind in its full complexity, he should just reflect on the experience of watching one of the comedies produced by the Marx Brothers. Of a specific interest is Groucho Marx, the master of wit, whose signature walk is a "duck walk." Why is it so funny to observe Groucho Marx walking like a duck?

When our mind is amused by Groucho's duck walk, it is not because we believe that Groucho IS a duck but because he walks LIKE a duck despite the fact that he is a human being. This example illustrates the way in which the human mind in its different expressions, from poetry to religion and mathematics, muses with the spectrum ranging from identity to equivalence. When Groucho performs his duck walk our mind may imagine him as a member of the set DUCK. Indeed, the famous Duck Test suggests that if something looks like a duck, quacks like a duck, and swims like a duck, then it is probably a duck. In a limited sense, Groucho Marx looks like a duck when he performs his duck walk. However, he doesn't quack or swim like a duck, and the reasonable conclusion is that he is not a duck, a trivial observation that doesn't require any sophisticated form of abductive reasoning. However, from a set theoretic perspective, it is difficult to understand the humorous aspect of the duck walk and its deep cognitive meaning; Groucho *is not* a member of the DUCK set and there is nothing to laugh about.

Indeed, one of the most basic operations of the mind is working with sets or collections of distinct objects. In this context, it is not a surprise that in "Mathematics, form and function," Saunders MacLane (1986/2012) saw the collection of objects as a basic human activity reflected in the formation of set theory in mathematics. The foundational status of set theory in mathematics can be explained from this

Y. Neuman (✉)
The Department of Cognitive and Brain Sciences, Ben-Gurion University of the Negev, Beersheba, Israel
e-mail: yneuman@bgu.ac.il

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*, Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_18

psychological perspective, and up to quite recently the foundational status of set theory held an almost sacred status in mathematics. The whole palace of mathematics is grounded in set theory. Can we imagine and model the mind as a set theoretic kind of a computational engine? Well, the mind clearly collects objects into sets; however the reduction of the mind to a set theoretic engine may miss something crucial. If this was the case then the mind would have been rigid and humor could not have any place in our life.

As explained by the mathematician Pierre Schapira (online): "A set A is a collection of elements a, b, c... with no relations between them, such as the points of a geometric space. One says that a belongs to A and writes $a \in A$. The only relations between a and b are the equality $a = b$ or its negation $a \neq b$."

From a cognitive epistemological perspective, one may wonder how the unrelated elements of a given set have found their way into the same basket. The mind, whether the mind of a civilized human being or the mind of a less civilized hen, groups elements in a structural and value-laden way (Neuman 2017). In vivo, elements do not simply find their way into the same basket. Therefore, it seems that from a cognitive perspective, set theory cannot serve as a foundational stratum of the mind, the way it is usually conceived in mathematics.

The most basic justification for the above argument is that in the realm of set theory, membership and identity cannot be negotiated. Either Groucho Marx belongs to the set of ducks or not. If he is a duck then no surprise should be expected from observing his duck walk. How else can a duck walk? On the other hand, if Mr. Marx is not a duck, what is so funny in observing him walking in a low squatting position? The idea of the mind as grounded in the same logic as the one of set theory clearly misses the undenied flexibility evident in the mind's different forms of behavior. One may argue that the mind's violation of set theoretic foundations is a shortcoming of our rationality. However, the attempt to revolutionize the foundations of mathematics in the Univalent Foundations Project at Princeton Institute for Advanced Studies suggests that instead of fitting our mind to the Bed of Procrustes formed by our limited and rigid formalizations, we may feel free to consider different foundations. From a cognitive perspective this new venture may not only better describe the mathematical realm but may even better fit the flexibility of our mind. In this context, we may use the Duck case as a point of departure for discussing the way we may better understand and model the human mind with clear implications for mathematical cognition, cognition of mathematics, and education of the mind through mathematics.

Part 2

Let us continue our discussion by considering the possibility that our approach should be relation oriented rather than object oriented. The reason for this move is that if we would like to free our mind from the rigidity of identity then a possible starting point is to imagine the mind as formed through relations rather than objects, whether abstract or concrete. Why should we be concerned by identity? The reason is that the logical notion of identity is that its binary form fixes our object in a rigid way. If Groucho Marx is either a duck or not, then he cannot be "like a" duck. The simple realm where a cigar is just a cigar cannot exist for the mind that continuously seeks the maximum freedom for musing with conceptual structures.

We can explain this idea by adopting the language of category theory (Lawvere and Schanuel 2000). A category may be imagined as a set of objects linked by arrows/morphisms. If we represent the idea of Groucho Marx walking like a duck, then we have two objects: Groucho Marx and Duck, linked by the relation/arrow "walks like a." However, we may consider a category in more abstract terms where *relations have precedence over objects*, an idea discussed by Gregory Bateson many years ago albeit in a different context (Bateson and Bateson 1987).

An object-free definition of a category (Adámek, Herrlich, and Strecker, 1990) is as follows. An object-free category is a partial binary algebra C = (M, °), where the members of M are called morphisms and the sign ° stands for composition. This algebra satisfies the following conditions:

Matching condition: For morphisms f, g, and h, the following conditions are equivalent:
$g \circ f$ and $h \circ g$ are defined,
$h \circ (g \circ f)$ is defined, and
$(h \circ g) \circ f$ is defined.

Associativity condition: If morphisms f, g, and h satisfy the matching conditions, then $h \circ (g \circ f) = (h \circ g) \circ f$.

Unit existence condition: For every morphism f there exist units u_C and u_D of (M, °) such that $u_C \circ f$ and $f \circ u_D$ are defined.

Smallness condition: For any pair of units (u_1, u_2) of (M, °) the class $hom(u_1, u_2) = \{f \in M \mid f \circ u_1 \text{ and } u_2 \circ f \text{ are defined}\}$ is a set.

In sum, this definition proposes that a category is basically a set of morphisms. We can go further by even denying the foundational notion of a set and substituting it with a "type" or a "space." According to this perspective, the "object" "Groucho Marx" has no meaning outside the relations that weave it into to a wider context, such as the observation that he walks like a duck. In fact, this relational perspective isn't new and one can find various expressions of it from the atomic relations proposed by C. S. Peirce to the linguistic theory of Lucien Tesnière (1959/2015).

Emphasizing the precedence of relations (i.e., morphisms) over objects is our first step. In this context, the axiom of identity "a = a" is interpreted as a morphism both originating and ending at *a*. From a more dynamic perspective, we may think about the identity morphism as a function changing nothing in the value of *a* as determined by his position as a point in a space formed by morphisms. This point has crucial consequences for the issue of personal identity as intensively discussed by philosophers, and by adopting it we may introduce a reasonable new understanding of personal identity which is a middleway between hard-headed forms of essentialism or naïve realism and their zealous modern opponents from the postmodernist sect, denying any existence beyond the one gained through social constructions and narratives (Neuman 2019).

Relations don't exist only between objects but can exist between categories too. These morphisms titled "functors" can have relations between them in a way that establishes second-order relations. In other words, relations can exist as morphisms between morphisms or as second-order morphisms. Our mind cannot perform the quantum leap from identity to equivalence unless given the ladder of relations and relations between relations.

This is not a semantic issue per se. Drawing the metaphor with homotopy type theory, we may substitute the expression "*a* is an element of the set A" with "*a* is a point of the space A," where the space A is defined in categorical terms of morphisms. According to this idea, instead of considering Groucho as a member of the set DUCK, we may consider him as a point of the space A, where the path/arrow from point "Groucho Marx" to point "Duck" is titled "walks like a." Here an interesting link is established between morphisms and information. The path from Groucho Marx to Duck is actually telling about Groucho Marx. Knowing that Groucho Marx walks like a duck portrays him in a way we haven't considered before. While we usually consider information in Shannon's terms of probability and surprise, here we may have a new sense of information, an idea worth pondering about. The idea is that information can be represented in terms of morphisms and morphisms of morphisms. In a certain sense, it reminds us of Deutsch and Marletto's (2015) theory of information, where information is expressed in terms of transformations.

This shift in perspective has implications for the idea of identity. In homotopy type theory the logical notion of identity of two objects a = b of the same type is substituted by the notion of a *path* from a to b in the same space. As the objects themselves are categories, it may be better to think about their identity in terms of equivalence rather than in the rigid logical notion of identity. The amusement of observing Groucho's duck walk may be therefore interpreted by understanding that in a certain sense Groucho IS a duck because as a "point" in the "duck space" he walks in a squatting position characterizing ducks. The equivalence formed by our mind between the space of Duck and the space of Groucho is only a specific illustration of the fact that our mind is relational and works in a flexible way at different scales of abstraction. While there is a difference between Groucho and a Duck, as two elements of two different sets, and a similarity between Groucho and a Duck as categories, we may also identify similarity of differences and differences of similarities. In sum, considered as an object in a set, Groucho Marx cannot be clearly related to a duck as they are two elements belonging to two different sets. Their similarity can be established only when conceived as categories.

Part 3: Equivalence and the Duck

Here is the point where we may dwell deeper into the notion of equivalence but without messing with the demanding formalisms elsewhere explained (Neuman 2017). To explain the idea of equivalence, let's recall Shakespeare's *Romeo and Juliet*, and the way Romeo describes Juliet by comparing her to the sun. A simple interpretation is that the same as the sun is rising in the east, Juliet is rising in front of his eyes. However, we should not forget that Romeo is deeply in love. In this context, another interpretation is that "Juliet is the sun" in another deep sense. The same as the sun warms the earth, Juliet warms her lover's heart. Here the analogy, or metaphor, can be simply represented as follows (Fig. 18.1):

What we see is that the functor mapping the relation between the celestial bodies to the two human beings can flexibly move us from identity to similarity. While Romeo denounces that "Juliet *is* the sun," it is clear that the "is" is not used to signify identity. Juliet is not the sun. However, when mapping the relation between sun and earth to the relation between the lover and his loved one, a metaphor is formed. It is a kind of a play where the mind says something like that. I know that the sun is the sun. I also know that the sun warms the earth in a physical sense. Now, I also know that Romeo feels warmth when he observes the approaching Juliet. Therefore, we may reason that Juliet is the sun. This metaphor is a-symmetric. We cannot trivially say that the sun is Juliet … However, our mind may consider the *equivalence* of the two above categories, as another step of its muse. If the two categories are equivalent then the two celestial bodies may be humanized as if they were the two heroes of the play. Indeed, when Romeo says: "Arise, fair sun, and kill the envious moon," he poetically describes two celestial bodies, the sun and the moon, as two human beings having the feeling of jealousy and quarreling over a social status. Moving from objects to relations, from identity to equivalence and from a one-dimensional realm to a high-dimensional space of relations of relations, allows us to better understand the flexibility of our mind.

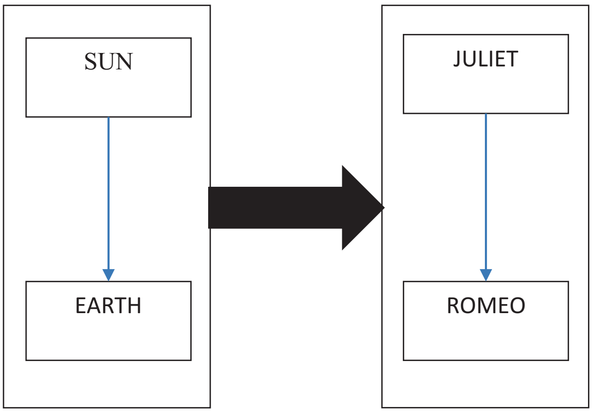


Fig. 18.1 Romeo and Juliet metaphor

References

Adámek, J., Herrlich, H., and Strecker G. E. (1990). *Abstract and concrete categories*. Chichester: John Wiley & Sons.

Bateson, G. and Bateson, M. C. (1987). *Angels fear: An investigation into the nature and meaning of the sacred*. New York: Rider.

Deutsch, D. and Marletto, C. (2015). Constructor theory of information. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 471 (2174): 20140540.

Lawvere, F. W. and Schanuel S. H. (2000). *Conceptual mathematics*. Cambridge: Cambridge University Press.

MacLane, S. (1986/2012). *Mathematics form and function*. Springer Science & Business Media.

Neuman, Y. (2017). *Mathematical structures of natural intelligence*. New York: Springer.

Neuman, Y. (2019, under review). *Personal identity: A semiotic journey from literature to idempotents and fixed points.*

Tesnière, L. (2015). *Elements of structural syntax*. Amsterdam: John Benjamins.

Chapter 19
On Mathematical Ways of Knowing: Musings of a Humanistic Mathematician

Gizem Karaali

Humanistic mathematics is a perspective on mathematics that emphasizes the ways our species creates, interacts with, and lives through it. I summarized this idea elsewhere (see Karaali 2015) by asserting that mathematics is the way our species makes sense of this world and that it is inherent in our thinking machinery; our mathematics in turn is dependent on the way we view our universe and ourselves. Lakoff and Núñez (2000) argue carefully and eloquently for a mathematics inherently based on human cognition.

Cognition is "the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses" (Wikipedia). In this note I attempt to engage with the construct of mathematical cognition through the lens of humanistic mathematics.

Three Questions

Cognition is essentially about mental processes involving knowledge, knowing, and understanding; mathematical cognition therefore raises questions about mathematical knowledge, knowing mathematics, and understanding mathematics. Thus, I first intend to explore broadly three related questions:

1. What does it mean to know something mathematical?
2. How do we come to know a mathematical truth?
3. What does it mean to understand something in mathematics?

G. Karaali (✉)
Department of Mathematics, Pomona College, Claremont, CA, USA
e-mail: Gizem.Karaali@pomona.edu

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*,
Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_19

In what follows, I will not pretend to offer a comprehensive treatment of any of these questions. But in the very least I intend to open up all three questions in productive ways, so that all readers intrigued by these questions will find in the following assertions worth agreeing with and arguing against.

Question 1

The first question is a natural extension of traditional epistemological investigations into mathematics. Philosophers have tinkered with the knowledge question for centuries, or rather, millennia, and mathematical knowledge has often been a part of the equation. Knowledge as justified true belief, a core tenet of epistemology since the Enlightenment, is where I want to start this note.[1]

If (mathematical) knowledge is justified true belief (in mathematical statements), we have multiple avenues to the first question. Or alternatively we have two related questions to attend to:

1A. What does it mean that a mathematical statement is true?
1B. What does it mean that a belief in a mathematical statement is justified?

1A is perhaps on the natural playground of mathematicians. Mathematicians seem to be concerned quite single-mindedly and profoundly in the truth of their assertions. One can even suggest that mathematics is nothing if not true. That is, doing math is making true mathematical assertions.[2] In some sense, therefore, I think that the truth of a mathematical assertion means that it is a part of mathematics, this edifice we human mathematicians are building together. Philosophers have tried to clarify what mathematicians might mean when they say that a mathematical statement is true. Sitting between mathematics and philosophy, the logician Alfred Tarski (1933/1956) proposed a definition of just what truth might mean in "the deductive sciences," which presumably include mathematics. Once again many have commented on Tarski's definition of truth (see Tarski and Vaught (1956) for an extension and elaboration). I will not go into that here but there is indeed much more that can be said along these lines if one is concerned about question 1A.

---

[1] At least since the 1960s with the publication of Edmund Gettier's "Is Justified True Belief Knowledge?" (Gettier 1963), this conventional approach to knowledge has seen many rebuttals and rephrasals, but for this note I will mostly ignore this recent body of work. My interest is most in line with the idea of mathematical knowledge as being justified true belief.

[2] Truth is surely not the only target of mathematicians. It is not even the driving force, according to William Byers, who writes: "Classifying ideas as true or false is just not the best way of thinking about them. Ideas may be fecund; they may be deep; they may be subtle; they may be trivial. These are the kinds of attributes we should ascribe to ideas" (Byers 2007). As a working mathematician, I agree with Byers, but this does not mean that I don't also believe that truth is a prerequisite. Even when mathematicians work with tentative and even patently false assertions, they have a broader truth in mind, and are not done until eventually they can reach that truth.

If we want to address question 1B about justification, we can, like some, invoke idealized conceptions of mathematical justification involving formal systems and proof theory.

But most mathematicians agree that belief in the truth of a mathematical statement is justified once there is a proof of the statement that experts can agree upon. This is quite in tune with Reuben Hersh's various definitions of a "proof" (Hersh 2014), most notably "The 'proof' is a procedure, an argument, a series of claims, that every qualified expert understands and accepts." Though some philosophers reading Hersh might disagree (see for instance Pollard 2014), it is indeed the case that when mathematicians claim that a statement is true, they mean that there is some consensus among the relevant experts that the statement is true. And an argument might have been a proof at a given time and place and afterwards, with contemporary expertise changing sides, it might become invalid. Similarly, proposed proofs do not become proofs until verified and validated by experts. Indeed, one might argue that "it is the provision of [...] evidence, not the endorsement of experts, that makes [a display of symbols, words, diagrams and such] a proof" (Pollard 2014). However, nothing needs to change in the display for an argument to remain an alleged proof until experts deem it is valid, and then and only then is the rest of the mathematical community comfortable in feeling justified to believe that the mathematical statement in consideration has finally been proved. What counts as persuasive evidence is almost always context dependent. In the case of law this is obvious; even Wikipedia knows that there are variations on what counts as proof, what counts as evidence.[3] Why do we expect mathematics to be different?

If we see mathematics as something done by humans, the formalist, proof theory-based understandings of proof and mathematics remain idealized approximations at best. It is the human (or sometimes, and begrudgingly, human-assisted) verification that mathematicians look for in a proof.[4] And this is definitely context based, both in terms of space-time coordinates and cultural makeup of the audience. As Israel Kleiner (1991) quotes—G. F. Simmons wrote "Mathematical rigor is like clothing: in its style it ought to suit the occasion, and it diminishes comfort and restricts freedom of movement if it is either too loose or too tight." This quote captures well how our understanding of just what should count as proof is dependent on the fashions of our times. This aligns with Harel's perspective (Harel 2008): "Mathematics is a human endeavor, not a predetermined reality. As such, it is the community of the creators of mathematics who makes decisions about the inclusion of new discoveries in the existing edifice of mathematics." Among the decisions left to the human creators of mathematics are the truth of a mathematical statement and the validity of its proof.

---

[3] https://en.wikipedia.org/wiki/Burden_of_proof_(law)#Legal_standards_for_burden_of_proof lists a selection of legal standards of evidence and proof.

[4] Do we know if the four-color theorem is true? Yes, we do. Or at least most mathematicians would concede that the human-assisted computer proof (or alternatively, the computer-assisted human proof) is enough for us to agree that it is true. There are still those who want more human proofs of the result, but the truth of the statement does not need further justification.

Question 2

The standard modern answer to question 2 is “by a rigorous proof.” Let us leave aside the historicity and cultural dependency of this response now, and its vagueness (what is proof, what is rigor). I already wrote a bit about all that above. The reader who is not yet convinced may read (Kleiner 1991) for more on rigor and proof. But I want to emphasize here the possible distinction between the doer of the proof and the believer who believes with justification that the proof is valid and that the related statement is true. Does the believer need to understand the proof in order to know that the related statement is true? How similar and how different is this from the calculus student saying that they learned calculus because they passed the final exam? Let us narrow things down a bit more. Should the successful calculus student be able to state the fundamental theorem of calculus? Should they be able to prove it? Should they be able to replicate the argument in their textbook or the one their instructor put on the board? Should they be able to provide a convincing argument for its truth? Alternatively, should they be able to use it in a problem that requires the result? When do we assume that a student has learned or knows the fundamental theorem of calculus?

Mathematics education researcher Guershon Harel thinks that these kinds of pedagogical questions are not independent from the philosopher’s concern about mathematical knowledge. In particular he proposes a definition of mathematics which originates from his pedagogical research that might help us with our endeavor here:

Mathematics consists of two complementary subsets:

- The first subset is a collection, or structure, of structures consisting of particular axioms, definitions, theorems, proofs, problems, and solutions. This subset consists of all the institutionalized ways of understanding in mathematics throughout history. It is denoted by WoU.
- The second subset consists of all the ways of thinking, which are characteristics of the mental acts whose products comprise the first set. It is denoted by WoT (Harel 2008).

Here Harel uses “ways of understanding” and “ways of thinking” as technical terms. According to him “a way of understanding is a particular cognitive product of a mental act carried out by an individual,” while “a way of thinking, on the other hand, is a cognitive characteristic of a mental act.” Here is how he uses them in context:

> Any statement a teacher (or a classmate) utters or puts on the board will be translated by each individual student into a way of understanding that depends on her or his experience and background … The range of ways of understanding a fraction makes the area of fractions a powerful elementary mathematics topic—one that can offer young students a concrete context to construct desirable—indeed, crucial—ways of thinking, such as: mathematical concepts can be understood in different ways, mathematical concepts should be understood in different ways, and it is advantageous to change ways of understanding of a mathematical concept in the process of solving problems.

Consider the mathematician who reads the statement of the four-color theorem and discussions about its proof and as a result is convinced of the veracity of the statement and the sufficiency of the proof. This mathematician, in my opinion, is not that different from the good student who learned of the fundamental theorem of calculus from their instructor, can use it in various scenarios, and can even perhaps outline a convincing argument about why it might be a reasonable thing to assume. In each case I'd say that the person knows the concept and construct in question. They believe the truth of a true statement and are justified in doing so. They put their trust in a relatively trustworthy source of expertise. But does the student really know (that is, understand) the fundamental theorem of calculus? Does the mathematician really know (that is, understand) the four-color theorem? This naturally brings us to question 3: What does it mean to understand something in mathematics?

Question 3

Harel's ways of thinking and ways of understanding are related quite visibly to question 3. For example, Harel offers a handful of ways of understanding the concept of fractions:

(a) The *part-whole* interpretation: m / n (where m and n are positive integers) means "m out of n objects."
(b) m/n means "the sum $1/n+\ldots+1/n$, m times,"
(c) "the quantity that results from m units being divided into n equal parts"
(d) "the measure of a segment m inches long in terms of a ruler whose unit is n inches"
(e) "the solution to the equation $nx = m$"
(f) "the ratio m: n; namely, m objects for each n objects."

Similarly, we can develop a list of ways of understanding for the derivative, in terms of the limit definition; in terms of slopes of tangent lines; in terms of linear approximations; and so on. And it seems reasonable to assume that when we say that a student understands fractions, we mean that they have mastered an indeterminate (but definitely nonzero) number of ways of understanding the concept. The fluency with which they can move from one interpretation to the other can help us if we want to further qualify how much they understand.

Understanding seems to presuppose knowing but is there anything more to it? More specifically, understanding a piece of mathematics does presuppose knowing that piece of mathematics; what we want to know is if it involves anything more.

This is the perfect context for me to bring up my favorite quote from one of my mathematical heroes. John von Neumann was a polymath, a genius mathematician who was instrumental in the development of quantum mechanics, game theory, functional analysis, operator algebras, and computer science, a great mathematical mind. This distinguished mathematician is known to have said to a young scholar asking for advice: "Young man, in mathematics you do not understand things. You just get used to them."

I grant that this is a sharp quote. It hits you hard and shakes you up, especially if you have at least a passing knowledge of the extent of von Neumann's own mathematical contributions. But can you take it seriously? Can you get anything out of it in the context of mathematical knowledge and mathematical ways of knowing?

My personal take on this quote is twofold.

One is that of the optimistic student of mathematics. Even if I feel like I am not understanding something, there is some benefit to pushing forward, if only a bit more. Doubtless the great mathematician is right; sometimes you simply have to move on, after accepting the fact as a fact and see where it leads you. Stubborn patience. Dogged perseverance.

However perhaps von Neumann did not really mean to recommend moving forward without understanding. Perhaps he was saying something else, that what you call understanding is not something subtle or sublime. In fact, when we are learning a new concept, a new theory, don't we start by making mental patterns, charting new pathways in our mind, formatting our minds so certain types of programs run well or smoothly enough? How is this different cognitively from getting used to brushing our teeth before going to bed, splashing our faces before leaving the bathroom, or eating with the fork in our right hand? Habit forming is done by doing something over and over again; aren't mathematical ways of thinking and ways of understanding reinforced by repeated practice as well? And is there a genuinely different, a genuinely distinct, sense of "understanding" that goes beyond "getting used to thinking of the concept in question in a particularly productive manner"?

Imagine a student who learns to think of a complex number first as an ordered pair, then as a point on the complex plane, and then as a linear transformation on the complex plane. When can we say that the student knows complex numbers? When do we say that the student understands them? I agree with Emily Grosholz who argues, using complex numbers as a concrete case study, that "the best way to teach students mathematics is through a repertoire of modes of representation, which is also the best way to make mathematical discoveries" (Grosholz 2013). But this also makes things a lot more complicated. If there are multiple ways of understanding, a la Harel, that need to go into understanding a construct, when do we really know the construct? When do we understand it?

Bloom's taxonomy is just one of many ranking frameworks education researchers use to delineate cognitive tasks and their demands on a learner. In the original taxonomy of Bloom and colleagues, knowledge is the very lowest level of cognition needed and includes knowledge of specifics, terminology, specific facts, ways, and means of dealing with specifics, conventions, trends and sequences, classifications and categories, criteria, methodology, universals and abstractions in a field, principles and generalizations, theories, and structures (Bloom et al. 1956). Understanding is related to the second level, comprehension. (In the revised version (Anderson and Krathwohl 2001), to understand is once again the second level, and there are several other layers of cognition ranked higher.) Mathematics education researchers prefer other frameworks to evaluate the cognitive demand of mathematical tasks; see Smith and Stein (1998) for a commonly used schema distinguishing between memorization, procedures without or with connections, and "doing mathematics." It is quite interesting that learning mathematics involves doing mathematics as a subcategory!

Putting taxonomies aside, knowing and understanding a concept might not be that different from one another after all. There seems to be a psychological difference for sure; understanding always occurs with knowing but sometimes we might feel we know something but do not "really understand." But perhaps von Neumann was not just being cheeky. Perhaps there is really nothing more than getting used to knowing something. Perhaps all we need is a brain formatted in the right way to accept our knowledge as truth and naturally so.[5]

Mathematical Ways of Knowing

In the remainder of this chapter, I want to delineate a construct I will call "mathematical ways of knowing." This is in some sense related to what Harel calls "ways of thinking." But I believe that it is not exactly the same.

I start with the axiomatic definition that mathematics is one of the main systematic bodies of knowledge that formulates and occasionally aims to address questions about human perception of the world and the human endeavors to understand it. The concepts and constructs of number, shape, form, time, change, and chance are fundamental to our understanding of our world as well as ourselves. These concepts and constructs are mathematical in nature, or at least they are naturally amenable to mathematical approaches.

Within this setting, I mean a mathematical way of knowing a way of formulating and addressing a question or a set of questions about our world and ourselves that allows for mathematical inquiry. It might be interesting to try and put these mathematical ways of knowing in contrast to or in conversation with a handful of other ways of knowing: scientific/empirical, faith based, and philosophical. I do no such thing in this note however. Here I merely put down some ideas as placeholders, as tentative yet suggestive notes to a self that might or might not be able to come back to revisit them in a possible future.

Mathematical Ways of Knowing: Rationalism and Imagination

The first two mathematical ways of knowing I would like to consider are rationalism and imagination.

Rationalism in mathematics is the fundamental assumption that we ought to reach our mathematical truths through reason. Experiment and experience might

[5] The formatting of the brain might sound strange but is not that far from the point of education. The point of education is to shape students' minds. What is that if not brain formatting? Indoctrination might also fit in this view. Some might take this to discussions of how mathematics education promotes dogmatic beliefs. Though I am interested in such inquiry, I will not pursue it here.

provide hints towards a truth, but they are never enough to convince us in the final count. Even if we "know something in our guts," we are not convinced that we have mathematical certainty until we can reason our way to that something. And reason and rational thinking are captured effectively in axiomatic thinking. The central tenets of axiomatic thinking are captured by Tarski in the following:

> When we set out to construct a given discipline, we distinguish, first of all, a certain small group of expressions of this discipline that seem to us to be immediately understandable; the expressions in this group we call PRIMITIVE TERMS or UNDEFINED TERMS, and we employ them without explaining their meanings. At the same time we adopt the principle: not to employ any of the other expressions of the discipline under consideration, unless its meaning has first been determined with the help of primitive terms and of such expressions of the discipline whose meanings have been explained previously (Tarski 1946).

Thus, we begin with some initial assumptions, ideas, fundamental beliefs, core values, and axioms. We take certain things for granted. We try to make these as self-evident as possible.[6] And from there we build our argument step by step, using logic as our guide. We define new constructs in terms of older, already accepted ones, and thus attempt to build a new world which has a solid foundation.

Mathematical rationality can be found in the various versions of the ontological argument for the existence of God, as well as in the Declaration of Independence of the United States (see Grabiner 1988 for a convincing argument about how mathematical rationality is built into the Declaration as well as many other illustrative examples of the impact of mathematical ways of knowing on Western thought).

In the history of intellectual thought Rationalism of the European Enlightenment was met with a backlash movement, Romanticism. Today we can but do not have to see these two as directly opposing and mutually exclusive methodologies, each rejecting and invalidating the other. Alternatively, and I believe more productively, we might choose to accept that they point to two distinct ways of knowing, and occasionally certain truths will be more accessible via one way than another.

Mathematical rationalism also has a similar complement, in what I will call mathematical imagination, or mathematical romanticism, if you will. Mathematical imagination is the way we select our axioms, the way we fix our principles. Mathematical imagination is how we determine our target truths. Human mathematicians do not start with a random formal axiomatic system and automatically go through all possible provable truths of the theory determined by it. Instead they engage with the worlds around them, both real and imaginary, and detect what is interesting, conjecture what might be productive to pursue, and then set out. Our human mathematics has a freedom to it and mathematical imagination captures this freedom.

And freedom, broadly construed, may be viewed in these terms as well. Sándor Szathmári's utopian, satiric novel, The Voyage to Kazohinia, might just be the best (fictional) guide to how mathematics is fundamental to a human society. Susan Siggelakis (2019) describes how the protagonist of the novel learns that in a society

[6] But of course, self-evidence, just like beauty, is in the eye of the beholder. There is more that can be said here.

without mathematics, “nothing stable exists with which a human can connect and find meaning in his/her life.” There is only chaos and violence. As Edward Frenkel (2013) declares, “where there is no mathematics, there is no freedom.”

Mathematical Ways of Knowing: Universals and Eclecticisms

Two other mathematical ways of knowing beckon us here: universals and eclecticisms. The tendency of the mathematician to generalize is well known. “Mathematics compares the most diverse phenomena and discovers the secret analogies that unite them,” wrote Jean Baptiste Joseph Fourier. “The art of doing mathematics is forgetting about the superfluous information,” says Hendrik Lenstra. Thus the human mathematician tries to generalize, to abstract from specific examples, and to reach universal statements with “any” and “all” that capture the essence of what is true about a whole slew of eclectic examples. The human desire to “see the big picture,” the human tendency to “find patterns,” is precisely what I mean by mathematical universalism.

Accompanying and complementing (again productively) this tendency is the alternative, what I will call mathematical eclecticism. David Hilbert describes the complementarity of these two tendencies as follows:

> In mathematics, as in any scientific research, we find two tendencies present. On the one hand, the tendency toward abstraction seeks to crystallize the logical relations inherent in the maze of material that is being studied, and to correlate the material in a systematic and orderly manner. On the other hand, the tendency toward intuitive understanding fosters a more immediate grasp of the objects one studies, a live rapport with them, so to speak, which stresses the concrete meaning of their relations.

Thus, by mathematical eclecticism I mean the search for that one representative example on the one hand (“the art of doing mathematics consists in finding that special case which contains all the germs of generality,” wrote Hilbert), and the excitement of the weirdness of eclectic cases on the other. In fact a lot of mathematics concerns itself with concrete examples. Paul Halmos wrote: “the heart of mathematics consists of concrete examples and concrete problems.” John B. Conway wrote, “mathematics is a collection of examples; a theorem is a statement about a collection of examples and the purpose of proving theorems is to classify and explain the examples.” Also, “We think in generalities, but we live in details,” wrote Alfred North Whitehead. In fact, I believe we think in both and we live in both. Once again, these two mathematical ways of knowing complement one another and help us live our human lives.

Mathematical Ways of Knowing: Certainty and Ambiguity

Alan Lightman says it best (Lightman 2018): “We are idealists and we are realists. We are dreamers and we are builders. We are experiencers and we are experimenters. We long for certainties, yet we ourselves are full of the ambiguities of the Mona

Lisa and the I Ching. We ourselves are a part of the yin-yang of the world." Certainty and ambiguity find their way into mathematics and mathematical ways of knowing in a similarly complementary fashion.

Certainty is a part of most people's perception of mathematics. Mathematics, for those people, is made up of questions and answers. Answers are certain once we know them. Indeed mathematics is perhaps the only certain knowledge we will ever have. I do not wish to minimize this perspective. I admit that I too have a romantic attachment to the idea that mathematical knowledge has a quality of certainty that goes farther than any other type of knowledge. And it took us a long time to get over this perspective as the unique way to conceive of mathematical truths and mathematical knowledge.

The loss of certainty in mathematics began more than a century ago (Kline 1980). Lakatos (1976) was also influential in convincing many who cared to listen of the fallibility of mathematics. And Byers (2007) is perhaps the most detailed expositor of the role of ambiguity in the work of mathematics today. The complementarity of these two ways of knowing is rich and, at least to me, inspiring.

Applications of Mathematical Ways of Knowing: Identity and Self-Knowledge

Doing math at school or anywhere else is tied deeply with our views of ourselves. This has good and bad aspects of course. We can relate our mathematical experiences to confidence, resilience, and determination as well as feelings of inadequacy, resistance, and rebelliousness; and as many can personally attest, any combination of the six can occur together. There is much emotion in mathematical engagement: hatred, love, anger, fear, anxiety, surprise, frustration, and anticipation. How we handle mathematical challenges (suffering alone, valiantly standing defeated or undefeated, finding commiserators and conspirators) tells us about ourselves. Many of those who continue to do math after school connect with it at an emotional and personal level. We find aesthetic stimulation and creative joy in mathematical activity, as well as terrible frustration and occasional bouts of tedium.

But can mathematical ways of knowing allow us to reach self-knowledge and a sense of identity mathematically? Andres Sanchez recounts how through an intentional application of set theory to his own personal life he was able to discover his true identity and sexuality (Sanchez 2018). Set theory, more generally, offers us pathways of thinking about belonging and not belonging. A clever student of mine, when asked to form study groups, chose to name his team "the identity element" as he wanted to work alone for the project in question. Mathematical ways of knowing may come in handy when thinking in terms of borders and boundaries of nations, communities, and cultural groups. Mathematical ways of thinking and knowing can indeed allow us to view and understand ourselves in new and insightful ways.[7]

[7] And in turn other ways of knowing can help us understand our mathematics better (Gutiérrez 2012).

Parting Words: Till Next Time

We have come a long way since Ptolemy argued that "mathematics alone yields knowledge and that, furthermore, it is the only path to the good life" (Feke 2018). Mathematical knowledge has lost its certainty somewhere along the way, and primality in the eye of the public a long time ago. Mathematicians eventually learned to be humbler about the reaches of mathematical ways of knowing. But mathematics can still yield powerful knowledge, and not just the kind that can blow up cities and optimize factory production. I urge us to try and open up to the world once again. If we dig deeper into mathematical ways of knowing and the contexts where they might apply, mathematics may yet surprise us.

References

Anderson, L. W. and Krathwohl, D. R. (eds.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Boston: Allyn and Bacon.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: David McKay Company.

Byers, W. (2007). *How mathematicians think: Using ambiguity, contradiction, and paradox to create mathematics*. Princeton: Princeton University Press.

Feke, J. (2018). *Ptolemy's philosophy: Mathematics as a way of life*. Princeton: Princeton University Press.

Frenkel, E. (2013). *Love and math: The heart of hidden reality*. New York: Basic Books.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis* 23 (6): 121–123.

Grabiner, J. V. (1988). The centrality of mathematics in the history of Western thought. *Mathematics Magazine* 61 (4): 220–230.

Grosholz, E. (2013). Teaching the complex numbers: What history and philosophy of mathematics suggest. *Journal of Humanistic Mathematics* 3 (1): 62–73.

Harel, G. (2008). What is mathematics? A pedagogical answer to a philosophical question. In: B. Gold and R. Simons (eds.), *Proof and other dilemmas: Mathematics and philosophy*, pp. 265–290. Mathematical Association of America.

Gutiérrez, R. (2012). Embracing *Nepantla*: Rethinking 'knowledge' and its use in mathematics teaching. *REDIMAT - Journal of Research in Mathematics Education* 1 (1): 29–56.

Hersh, R. (2014). *Experiencing mathematics: What do we do, when we do mathematics?* Providence: American Mathematical Society.

Karaali, G. (2015). Can zombies do math? In: M. Bockarova, M. Danesi, D. Martinovic and R. Nunez (eds.), *Mind in mathematics: Essays on mathematical cognition and mathematical method*, pp. 140–153. Munich: Lincom Europa.

Kleiner, I. (1991). Rigor and proof in mathematics: A historical perspective. *Mathematics Magazine* 64 (5): 291–314.

Kline, M. (1980). *Mathematics: The loss of certainty*. Oxford: Oxford University Press.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge: Cambridge University Press.

Lakoff, G. and Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Lightman, A. (2018). *Searching for stars on an island in Maine*. New York: Penguin Random House.

Pollard, S. (2014). Review of *Experiencing mathematics* by Reuben Hersh. *Philosophia Mathematica* 22 (2): 271–274.

Sanchez, A. (2018). My sets and sexuality. *Journal of Humanistic Mathematics* 8 (1): 359–370.

Siggelakis, S. J. (2019). Sándor Szathmári's *Kazohinia*: Mathematics and the Platonic Idea. *Journal of Humanistic Mathematics* 9 (1): 3–23.

Smith, M. S. and Stein, M. K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School* 3 (5): 344–350.

Tarski, A. (1933/1956). The concept of truth in the languages of the deductive sciences (Polish), *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych* 34, pp. 13–172. Expanded English translation in Tarski 1956 (*Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, edited by John Corcoran, Oxford: Oxford University Press), pages 152-278.

Tarski, A. (1946). *Introduction to logic and the methodology of the deductive sciences*. Oxford: Oxford University Press.

Tarski, A. and Vaught, R. (1956). Arithmetical extensions of relational systems. *Compositio Mathematica* 13: 81–102.

Chapter 20
Epilogue: So, What Is Math Cognition?

Marcel Danesi

Introduction

A key 2005 collection of papers (Royer 2008) showed how complex the study of mathematical cognition (MC) had become already in the early 2000s, incorporating a broad range of scientific, educational, and humanistic perspectives into its *modus operandi*. Studies published in the journal *Mathematical Cognition* have also revealed how truly expansive the field is, bringing together researchers and scholars from diverse disciplines, from neuroscience to semiotics. This volume has aimed to provide a contemporary snapshot of how the study of MC is developing. In this final chapter, the objective is to provide a selective overview of different approaches from the past as a concluding historical assessment.

The interdisciplinary study of MC became a concrete plan of action after the publication of Lakoff and Núñez's 2000 book, *Where Mathematics Comes from*, following on the coattails of intriguing works by Dehaene (1997) and Butterworth (1999). Lakoff and Núñez argued that MC is no different neurologically from linguistic cognition, since both involve blending information from different parts of the brain to produce concepts. This is why we use language to learn math. The most salient manifestation of blending in both linguistic and mathematical cognition can be seen in metaphor (as studies in this volume have saliently shown). If metaphor is indeed at the core of MC then it brings mathematics directly into the sphere of language and culture where it is shaped symbolically and textually. This was the conclusion deduced as well by American philosopher Max Black in his groundbreaking 1962 book, *Models and Metaphors,* in which Black argued that the cognitive source

M. Danesi (✉)
University of Toronto, Toronto, ON, Canada

Cognitive Science Network, Fields Institute for Research in Mathematical Sciences, Toronto, ON, Canada
e-mail: marcel.danesi@utoronto.ca

M. Danesi (ed.), *Interdisciplinary Perspectives on Math Cognition*, Mathematics in Mind, https://doi.org/10.1007/978-3-030-22537-7_20

of science and mathematics was the same one that involved the same kind of metaphorical thinking that characterizes discourse. Indirectly, Black laid the foundations for a humanistic-linguistic study of MC with his radical idea for the era in which it was written.

The interdisciplinary study of MC has produced a huge database of findings, theories, and insights into how mathematics intersects with other neural faculties such as language and drawing. The field has not just produced significant findings about how math is processed in the brain, but also reopened long-standing philosophical debates about the nature of mathematics. In this chapter a general characterization of MC that extends the classic views will be discussed at first. Then, it will selectively discuss various works and findings that can be used to determine whether math is separate or not from language, neurologically and cognitively. Finally, it will revisit the Platonist-versus-constructivist debate on the basis of these patterns, which is intrinsically a cognitive debate.

Mathematical Cognition: A Selective Historical Survey

Mathematical cognition is defined in two main ways—first, it is defined as the awareness of structural patterns among quantitative and spatial concepts; second, it is defined as the awareness of how symbols stand for these concepts and how they encode them (for example, Radford 2010).

A historical point of departure for investigating MC is Immanuel Kant's (1790: 278) assertion that thinking mathematically involves "combining and comparing given concepts of magnitudes, which are clear and certain, with a view to establishing what can be inferred from them." He argued further that the combination and comparison become explicit through the "visible signs" that we use to represent them—thus integrating the two definitions above predictively. So, a diagram of a triangle (a visible sign) compared to that of a square (another visible sign) will show the differentiation between the two concretely. As trivial as Kant's definition might seem, upon further consideration it is obvious that the kind of visualization that he describes now falls under the rubric of spatial cognition, and the claim that visible signs guide it is consistent with various psychological and semiotic theories of MC (see, for example, Stjernfelt 2007, Danesi 2013). As we know today, mental visualization stems from the brain's ability to synthesize scattered bits of information into holistic entities that can be understood consciously through representations such as diagrams.

Kant's main idea that diagrams reveal thought patterns was given a semiotic-theoretical formulation by Charles Peirce's existential graph theory (Peirce 1931–1958, vol. 2: 398–433, vol. 4: 347–584). An existential graph is a diagram that displays how the parts of some concept are visualized as related to each other. For example, a Venn diagram can be used to show how sets are related to each other in a holistic visual way. These do not portray information directly, but the process of thinking about it (Peirce, vol. 4: 6). Peirce called his existential graphs "moving pictures of thought" (Peirce, vol. 4: 8–11). As Kiryuschenko (2012: 122) has aptly observed, for Peirce "graphic language allows us to experience a meaning visually as a set of transitional states, where the meaning is accessible in its entirety at any given 'here and now' during its transformation."

The gist of the foregoing discussion is that diagrams and visual signs might mirror the nature of MC itself—an idea that has been examined empirically in abundance (Shin 1994; Chandrasekaran et al. 1995; Hammer 1995; Hammer and Shin 1996, 1998; Allwein and Barwise 1996; Barker-Plummer and Bailin 1997, 2001; Kulpa 2004; Stjernfelt 2007; Roberts 2009). The main implication is that the study of MC must take semiotic notions, such as those by Peirce, into account in order to better explain the findings of neuroscientists in this domain. In effect, diagrams represent our intuitions about quantity, space, and relations in a visually expressive way that appears to mirror the actual imagery in the brain, or more specifically what Lakoff and Johnson (1980) call *image schemata*—mental outlines of abstractions. The intuitions are probably universal (first type of definition); the visual representations, which include numerals, are products of historical processes (second type of definition).

Algebraic notation, too, is a diagrammatic strategy for compressing information, much like pictography does in reproducing referents in compressed semiotic forms (Danesi and Bockarova 2013). An equation is an existential graph consisting of signs (letters, numbers, symbols) organized in such a way as to reflect the structure of events that it aims to represent. It may show that some parts are tied to a strict order, whereas others may be unconstrained as to sequential structure. As Kauffman (2001: 80) observes, Peirce's existential graphs contain arithmetical information in an economical form:

> Peirce's Existential Graphs are an economical way to write first order logic in diagrams on a plane, by using a combination of alphabetical symbols and circles and ovals. Existential graphs grow from these beginnings and become a well-formed two dimensional algebra. It is a calculus about the properties of the distinction made by any circle or oval in the plane, and by abduction it is about the properties of any distinction.

An equation such as the Pythagorean one ($c^2 = a^2 + b^2$) is an existential graph, since it is a visual representation of the relations among the variables (originally standing for the sides of the triangle). But, being a graph, it also tells us that the parts relate to each other in many ways other than in terms of the initial triangle referent. It reveals hidden structure, such as the fact that there are infinitely many Pythagorean triples, or sets of three integers that satisfy the equation. Expressed in language ("the square on the hypotenuse is equal to the sum of the squares on the other two sides"), we would literally not be able to *see* this hidden implication. Once the equation exists as a graph, it becomes the source for further inferences and insights, which (as is well known) gave rise to a hypothesis, namely Fermat's Last Theorem, whereby only when $n = 2$ does the general formula hold ($c^n = a^n + b^n$) (Taylor and Wiles 1995). This, in turn, has led to many other discoveries (Danesi 2013). To use Susan Langer's (1948) distinction between discursive and presentational cognition, the equation tells us much more than the statement (a discursive act) because it "presents" inherent structure holistically, as an abstract form. We do not read a diagram, a melody, or an equation as individual bits and pieces (notes, shapes, symbols), but *presentationally,* as a totality which encloses and reveals much more meaning. Mathematical notation is visually presentational, which as research has shown, may be the source for how abstract ideas emerge (Barwise and Etchemendy 1994; Allwein and Barwise 1996; Cummins 1996; Chandrasekaran et al. 1995).

Needless to say, mathematicians have always used diagrams to carry out their craft. Some diagrammatic practices, such as Cartesian geometry, become actual fields of mathematics in themselves; set theory, for example, is an *ipso facto* theory of mathematics, based on Venn diagrams (1880, 1881) which were introduced so that mathematicians could literally see the logical implications of mathematical patterns and laws. These are, as mentioned, externalized image schemata (Lakoff and Johnson 1980, 1999; Lakoff 1987; Johnson 1987; Lakoff and Núñez 2000) which allow us to gain direct cognitive access to hidden structure in mathematical phenomena. Actually, the shift from sentential logic to diagram logic started with Euler, who was the first to represent categorical sentences as intersecting circles, embedded circles, and so on (Hammer and Shin 1996, 1998). It actually does not matter whether the schema is a circle, a square, a rectangle, or a freely drawn form; it is the way it portrays pattern that cuts across language (and languages) and allows us to envision a relation or concept in outline form. The power of the diagrams over sentences lies in the fact that no additional conventions, paraphrases, or elaborations are needed—the relationships holding among sets are shown by means of the same relationships holding among the schemata representing them. Euler was aware, however, of both the strengths and weaknesses of visual representation. For instance, in the statement "No A is B. Some C is A. Therefore, Some C is not B," no single diagram can be envisioned to represent the two premises, because the relationship between sets B and C cannot be fully specified in one single diagram. Venn (1881: 510) tackled Euler's dilemma by showing how partial information can be visualized (such as overlaps or intersections among circles). But Peirce pointed out that Venn's system had no way of representing existential statements, disjunctive information, probabilities, and some specific kinds of logical relations. He argued that "All A are B or some A is B" cannot be represented by neither the Euler nor the Venn systems in a single diagram.

Among the first to investigate the relation between imagery and mathematical reasoning was Jean Piaget, who sought to understand the development of number sense in relation to symbolism (summarized in Piaget 1952). In one experiment, he showed a 5-year-old child two matching sets of six eggs placed in six separate egg cups. He then asked the child whether there were as many eggs as egg cups (or not)—the child usually replied in the affirmative. Piaget then took the eggs out of the cups, bunching them together, leaving the egg cups in place. He then asked the child whether or not all the eggs could be put into the cups, one in each cup and none left over. The child answered negatively. Asked to count both eggs and cups, the child would correctly say that there was the same amount. But when asked if there were as many eggs as cups, the child would again answer "no." Piaget concluded that the child had not grasped the relational properties of numeration, which are not affected by changes in the positions of objects. Piaget showed, in effect, that 5-year-old children have not yet established in their minds the symbolic connection between numerals and number sense (Skemp 1971: 154).

A key study by Yancey et al. (1989) has shown that training students how to use visualization (diagrams, charts, etc.) to solve problems results in improved performance. As Musser, Burger, and Peterson (2006: 20) have aptly put it: "All students should represent, analyze, and generalize a variety of patterns with tables, graphs, words, and, when possible, symbolic rules." Another study by Ambrose (2002) suggests, moreover, that students who are taught appropriately with concrete strategies, but not allowed to develop their own abstract representational grasp of arithmetic, are less likely to develop arithmetical fluency.

Is Math Cognition Species Specific?

The study of MC has led to a whole series of existential-philosophical questions. For example: Intuitive number sense may be a cross-species faculty, but the use of symbols to represent numbers is specific human trait. As the philosopher Ernst Cassirer (1944) once put it, we are "a symbolic species," incapable of establishing knowledge without symbols. So is math cognition specific to the human species?

Neuroscientist Brian Butterworth (1999) is well known for his investigation of this question. He starts with the premise that we all possess a fundamental number sense, which he calls "numerosity." Numbers do not exist in the brain in the same way verbal signs such as words do; they constitute a separate kind of intelligence with its own brain module, located in the left parietal lobe. But this does not guarantee that mathematical competence will emerge homogeneously in all individuals. It is a phylogenetic trait that varies ontogenetically. Rather, the reason a person falters at math is not because of a "wrong gene" or "engine part" in the brain, but because the individual has not fully developed numerosity, and the reason is due to environmental and personal psychological factors.

Finding hard evidence to explain why numerosity emerged from the course of human evolution is a difficult venture. Nevertheless, there is a growing body of research that is supportive of Butterworth's basic thesis—that number sense is instinctual and that it may be separate from language. In one study, Izard et al. (2011) looked at Euclidean concepts in an indigenous Amazonian society, called the Munduruku. The team tested the hypothesis that certain aspects of non-perceptible Euclidean geometry map onto intuitions of space that are present in all humans (such as intuitions of points, lines, and surfaces), even in the absence of formal mathematical training. The subjects included adults and age-matched children controls from the United States and France as well as younger American children without training in geometry. The responses of Munduruku adults and children converged with that of mathematically educated adults and children and revealed an intuitive understanding of essential properties of Euclidean geometry. For instance, on a surface described to them as perfectly planar, the Mundurucu's estimations of the internal angles of triangles added up to ~180 degrees, and when asked explicitly they stated that there exists one single parallel line to any given line through a given point. These intuitions were also present in the group of younger American participants. The researchers concluded that, during childhood, humans develop geometrical intuitions that spontaneously accord with the principles of Euclidean geometry, even in the absence of training in such geometry. There is however contradictory evidence that geometric notions are not innate, but subject to cultural influences (Núñez et al. 1999). In one study, Lesh and Harel (2003) got students to develop their own models of a problem space, guided by prompts. Without the latter, they were incapable of coming up with them. It might be that Euclidean notions may be universal and that these are concretized in specific cultural ways. For now, there is no definitive answer to the issue one way or the other.

The emergence of abilities such as speaking and counting are a consequence of four critical evolutionary events—bipedalism, a brain enlargement unparalleled among species, an extraordinary capacity for toolmaking, and the advent of the tribe as a basic form of human collective life (Cartmill et al. 1986). Bipedalism liberated the fingers to count and gesture. Although other species, including some non-primate ones, are capable of tool use, only in the human species did complete bipedalism free the hand sufficiently to allow it to become a supremely sensitive and precise manipulator and grasper, thus permitting proficient toolmaking and tool use in the species. Shortly after becoming bipedal, the neuro-paleontological evidence suggests that the human species underwent rapid brain expansion. The large brain of modern-day *Homo* is more than double that of early toolmakers. This increase was achieved by the process of neoteny, that is, by the prolongation of the juvenile stage of brain and skull development in neonates. Like most other species, humans have always lived in groups. Group life enhances survivability by providing a collective form of life. The early tribal collectivities have left evidence that gesture (as inscribed on surfaces through pictography) and counting skills occurred in tandem. This supports the co-development of language and numerosity that Lakoff and Núñez (2000) suggest is part of brain structure.

Keith Devlin (2000, 2005) entered the debate with the notion of an innate "math instinct." If there is some innate capacity for mathematical thinking, which there must be, otherwise no one could do it, why does it vary so widely, both among individuals in a specific culture and across cultures? Devlin connects the math ability to language, since both are used by humans to model the world symbolically. But this then raises another question: Why, then, can we speak easily, but not do math so easily (in many cases)? The answer, according to Devlin, is that we can and do, but we do not recognize that we are doing math when we do it. As he argues, our prehistoric ancestors' brains were essentially the same as ours, so they must have had the same underlying abilities. But those brains could hardly have imagined how to multiply 15 by 36 or prove Fermat's Last Theorem.

One can argue that there are four orders involved in learning how to go from counting to, say, equations. The first is the instinctive ability itself to count. This is probably innate. Using signs to stand for counting constitutes a second order. It is the level at which counting concepts are represented by numeral symbols. The third order is the level at which numerals are organized into a code of operations based on counting processes (adding, taking away, comparing, dividing, and so on). Finally, a fourth order inheres in the capacity to generalize the features and patterns of counting and numeral representations. This is where representations such as equations come into the developmental-evolutionary picture.

Stanislas Dehaene's (1997) work brings forth experimental evidence to suggest that the human brain and that of some chimps come with a wired-in aptitude for math. The difference in the case of the latter is an inability to formalize this innate knowledge and then use it for invention and discovery. Dehaene has catalogued evidence that rats, pigeons, raccoons, and chimpanzees can perform simple calculations, describing ingenious experiments that show that human infants also show a parallel manifestation of number sense. This rudimentary number sense is as basic to the way the brain understands the world as is the perception of color. But how then did the brain leap from this ability to trigonometry, calculus, and beyond? Dehaene shows that it was the invention of symbolic systems that started us on the climb to higher mathematics. He argues this by tracing the history of numbers, from early times when people indicated a number by pointing to a part of their body (even today, in many societies in New Guinea, the word for six is "wrist"), to early abstract numbers such as Roman numerals (chosen for the ease with which they could be carved into wooden sticks), to modern numerals and number systems. Dehaene argues, finally, that the human brain does not work like a computer, and that the physical world is not based on mathematics—rather, mathematics evolved to explain the physical world the way that the eye evolved to provide sight.

Studies inspired by both Butterworth's and Dehaene's ideas have become widespread in MC circles (for example, Ardila and Rosselli 2002; Dehaene 2004; Isaacs et al. 2001; Dehaene et al. 2003; Butterworth et al. 2011). Dehaene (1997) himself showed that when a rat is trained to press a bar 8 or 16 times to receive a food reward, the number of bar presses will approximate a Gaussian distribution with peak around 8 or 16 bar presses. When rats are more hungry, their bar-pressing behavior is more rapid, so by showing that the peak number of bar presses is the same for either well-fed or hungry rats, it is possible to disentangle time from number of bar presses. Similarly, researchers have set up hidden speakers in the African savannah to test natural (untrained) behavior in lions (McComb et al. 1994). The speakers play a number of lion calls, from 1 to 5. If a single lioness hears, for example, three calls from unknown lions, she will leave, but if she is with four of her sisters, they will go and explore. This suggests that not only can lions tell when they are "outnumbered" but also that they can do this on the basis of signals from different sensory modalities, suggesting that numerosity involves a multisensory neural substratum.

Blending Theory

As mentioned above, the study of MC started proliferating and diversifying after Lakoff and Núñez (2000) claimed that the proofs and theorems of mathematics are arrived at via the same cognitive mechanisms that underlie language—analogy, metaphor, and metonymy. This claim has been largely substantiated with neurological techniques such as fMRI and other scanning devices, which have led to adopting the notion of *blending*, whereby concepts in the brain are sensed as "informing" each other in a common neural substrate (Fauconnier and Turner 2002). Determining the characteristics of this substrate is an ongoing goal of research on MC (Danesi 2016).

Blending can be used, for example, to explain negative numbers. These are derived from two basic metaphors, which Lakoff and Núñez call *grounding* and *linking*. Grounding metaphors encode basic ideas, being directly "grounded" in experience. For example, addition develops from the experience of counting objects and then inserting them in a collection. Linking metaphors connect concepts within mathematics that may or may not be based on physical experiences. Some examples of this are the number line, inequalities, and absolute value properties within an epsilon-delta proof of limit. Linking metaphors are the source of negative numbers, which emerge from a connective form of reasoning within the system of mathematics. They are linkage blends, as Alexander (2012: 28) elaborates:

> Using the natural numbers, we made a much bigger set, way too big in fact. So we judiciously collapsed the bigger set down. In this way, we collapse down to our original set of natural numbers, but we also picked up a whole new set of numbers, which we call the negative numbers, along with arithmetic operations, addition, multiplication, subtraction. And there is our payoff. With negative numbers, subtraction is always possible. This is but one example, but in it we can see a larger, and quite important, issue of cognition. The larger set of numbers, positive and negative, is a cognitive blend in mathematics … The numbers, now enlarged to include negative numbers, become an entity with its own identity. The collapse in notation reflects this. One quickly abandons the (minuend, subtrahend) formulation, so that rather than (6, 8) one uses -2. This is an essential feature of a cognitive blend; something new has emerged.

This kind of metaphorical (connective) thinking occurs because of gaps that are felt to inhere in the system. As Godino, Font, Wilhelmi, and Lurduy (2011: 250) cogently argue, notational systems are practical (experiential) solutions to the problem of counting:

> As we have freedom to invent symbols and objects as a means to express the cardinality of sets, that is to say, to respond to the question, how many are there?, the collection of possible numeral systems is unlimited. In principle, any limitless collection of objects, whatever its nature may be, could be used as a numeral system: diverse cultures have used sets of little stones, or parts of the human body, etc., as numeral systems to solve this problem.

Fauconnier and Turner (2002) have proposed arguments along the same lines, giving substance to the notion that ideas in mathematics are based on blends deriving from experiences and associations within these experiences. Interestingly, the idea that metaphor plays a role in mathematical logic seems to have never been held seriously until very recently, even though, as Marcus (2012: 124) observes, mathematical terms are mainly metaphors:

> For a long time, metaphor was considered incompatible with the requirements of rigor and preciseness of mathematics. This happened because it was seen only as a rhetorical device such as "this girl is a flower." However, the largest part of mathematical terminology is the result of some metaphorical processes, using transfers from ordinary language. Mathematical terms such as *function, union, inclusion, border, frontier, distance, bounded, open, closed, imaginary number, rational/irrational number* are only a few examples in this respect. Similar metaphorical processes take place in the artificial component of the mathematical sign system.

Actually, already in the 1960s, a number of structuralist linguists prefigured blending theory, by suggesting that mathematics and language shared basic structural properties (Hockett 1967; Harris 1968). Their pioneering writings were essentially exploratory investigations of structural analogies between mathematics and language. They argued, for example, that both possessed the feature of double articulation (the use of a limited set of units to make complex forms ad infinitum), ordered rules for interrelating internal structures, among other things. Many interesting comparisons emerged from these writings, which contained an important subtext—by exploring the structures of mathematics and language in correlative ways, we might hit upon deeper points of contact and thus at a common cognitive origin for both. Those points find their articulation in the work of Lakoff and Núñez and others working within the blending paradigm. Mathematics makes sense when it encodes concepts that fit our experiences of the world—experiences of quantity, space, motion, force, change, mass, shape, probability, self-regulating processes, and so on. The inspiration for new mathematics comes from these experiences as it does for new language.

A classic example of this was Gödel's famous proof, which Lakoff has argued (see Bockarova and Danesi 2012: 4–5) was inspired by Cantor's diagonal method. As is well known, Gödel proved that within any formal logical system there are results that can be neither proved nor disproved. Gödel found a statement in a set of statements that could be extracted by going through them in a diagonal fashion—now called Gödel's diagonal lemma. That produced a statement, S, like Cantor's C, that does not exist in the set of statements. Cantor's diagonal and one-to-one matching proofs are mathematical metaphors—associations linking different domains in a specific way (one-to-one correspondences). This insight led Gödel to envision three metaphors of his own: (1) the "Gödel number of a symbol," which is evident in the argument that a symbol in a system is the corresponding number in the Cantorian one-to-one matching system (whereby any two sets of symbols can be put into a one-to-one relation); (2) the "Gödel number of a symbol in a sequence," which is manifest in the demonstration that the n^{th} symbol in a sequence is the n^{th} prime raised to the power of the Gödel number of the symbol; and (3) "Gödel's central metaphor," which was Gödel's proof that a symbol sequence is the product of the Gödel numbers of the symbols in the sequence.

The proof exemplifies how blending works. When the brain identifies two distinct entities in different neural regions as the same entity in a third neural region, they are blended together. Gödel's metaphors come from neural circuits linking a *number* source to a *symbol* target. In each case, there is a blend, with a single entity composed of both a *number* and a *symbol sequence*. When the symbol sequence is a formal proof, a new mathematical entity appears—a "proof number."

It is relevant to turn to the ideas of René Thom (1975, 2010) who called discoveries in mathematics "catastrophes," that is, mental activities that subvert or overturn existing knowledge. He called the process "semiogenesis," which he defined as the emergence of "pregnant" forms within symbol systems themselves, that is, as forms that emerge by happenstance through contemplation and manipulation of the previous forms. As this goes on, every so often, a catastrophe occurs that leads to new insights, disrupting the previous system. Discovery is indeed catastrophic, but why does the brain produces catastrophes in the first place? Perhaps the connection between the brain, the body, and the world is so intrinsic that the brain cannot really understand itself.

Epilogue: Selected Themes

The chapters of this book span the interdisciplinary scope of MC study, from the empirical to the educational and speculative, as well as examining aspects of mathematical method, such as proof, and what this tells us about the nature of MC. The objective has been twofold: to show how this line of inquiry can be enlarged profitably through an expanded pool of participating disciplines and to shed some new light on math cognition itself from within this pool. Only in this way can progress be made in grasping what math cognition truly is. Together, the chapters of this book constitute a mixture of views, findings, and theories that, when collated, do hopefully give us a better sense of what math cognition is.

References

Alexander, J. (2012). On the cognitive and semiotic structure of mathematics. In: M. Bockarova, M. Danesi, and R. Núñez (eds.), *Semiotic and cognitive science essays on the nature of mathematics*, pp. 1–34. Munich: Lincom Europa.

Allwein, G. and Barwise, J. (eds.) (1996). *Logical reasoning with diagrams*. Oxford: Oxford University Press.

Ambrose, C. (2002). Are we overemphasizing manipulatives in the primary grades to the detriment of girls? *Teaching Children Mathematics* 9: 16–21.

Ardila A. and Rosselli M. (2002). Acalculia and dyscalculia. *Neuropsychology Review* 12: 179–231.

Barker-Plummer, D. and Bailin, S. C. (1997). The role of diagrams in mathematical proofs. *Machine Graphics and Vision* 8: 25–58.

Barker-Plummer, D. and Bailin, S. C. (2001). On the practical semantics of mathematical diagrams. In: M. Anderson (ed.), *Reasoning with diagrammatic representations*. New York: Springer.

Barwise, J. and Etchemendy, J. (1994). *Hyperproof*. Stanford: CSLI Publications.

Black, M. (1962). *Models and metaphors*. Ithaca: Cornell University Press.

Bockarova, M., Danesi, M., and Núñez, R. (eds.) (2012). *Semiotic and cognitive science essays on the nature of mathematics*. Munich: Lincom Europa.

Butterworth B., Varma S., Laurillard D. (2011). Dyscalculia: From brain to education. *Science* 332: 1049–1053

Butterworth, B. (1999). *What counts: How every brain is hardwired for math.* Michigan: Free Press.

Cartmill, M., Pilbeam, D., and Isaac, G. (1986). One hundred years of paleoanthropology. *American Scientist* 74: 410–420.

Cassirer, E. (1944). *An essay on man.* New Haven: Yale University Press.

Chandrasekaran, B., Glasgow, J., and Narayanan, N. H. (eds.) (1995). *Diagrammatic reasoning: Cognitive and computational perspectives*. Cambridge: MIT Press.

Cummins, R. (1996). *Representations, targets, and attitudes*. Cambridge: MIT Press.

Danesi, M. (2013). *Discovery in mathematics: An interdisciplinary perspective.* Munich: Lincom Europa.

Danesi, M. (2016). *Language and mathematics: an interdisciplinary guide.* Berlin: Mouton de Gruyter.

Danesi, M. and Bockarova, M. (2013). *Mathematics as a modeling system*. Tartu: University of Tartu Press.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics.* Oxford: Oxford University Press.

Dehaene, S. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology* 14: 218–224.

Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology* 20: 487–506

Devlin, K. J. (2000). *The math gene: How mathematical thinking evolved and why numbers are like gossip*. New York: Basic.

Devlin, K. J. (2005). *The math instinct: Why you're a mathematical genius (along with lobsters, birds, cats and dogs)*. New York: Thunder's Mouth Press.

Fauconnier, G. and Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities.* New York: Basic.

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, Teil I. *Monatshefte für Mathematik und Physik* 38: 173–189.

Godino, J. D., Font, V., Wilhelmi, R. and Lurduy, O. (2011). Why is the learning of elementary arithmetic concepts difficult? Semiotic tools for understanding the nature of mathematical objects. *Educational Studies in Mathematics* 77: 247–265.

Hammer, E. (1995). Reasoning with sentences and diagrams. *Notre Dame Journal of Formal Logic* 35: 73–87.

Hammer, E. and Shin, S. (1996). Euler and the role of visualization in logic. In: J. Seligman and D. Westerståhl, (eds.), *Logic, language and computation,* Volume 1. Stanford: CSLI Publications.

Hammer, E., and Shin, S. (1998). Euler's visual logic. *History and Philosophy of Logic* 19: 1–29.

Harris, Z. (1968). *Mathematical structures of language.* New York: John Wiley.

Hockett, C. F. (1967). *Language, mathematics and linguistics*. The Hague: Mouton.

Isaacs E. B, Edmonds C. J., Lucas A. and Gadian D. G. (2001). Calculation difficulties in children of very low birthweight: A neural correlate. *Brain* 124: 1701–1707.

Izard, V. Pica, P., Pelke, E. S., and Dehaene, S. (2011). Flexible intuitions of Euclidean geometry in an Amazonian indigene group. *PNAS* 108: 9782–9787.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination and reason*. Chicago: University of Chicago Press.

Kant, I. (1790. *Critique of pure reason*, trans. J. M. D. Meiklejohn. CreateSpace Platform.

Kauffman, L. K. (2001). The mathematics of Charles Sanders Peirce. *Cybernetics & Human Knowing* 8: 79–110.

Kiryushchenko, V. (2012). The visual and the virtual in theory, life and scientific practice: The case of Peirce's quincuncial map projection. In: M. Bockarova, M. Danesi, and R. Núñez (eds.), *Semiotic and cognitive science essays on the nature of mathematics.* Munich: Lincom Europa.

Kulpa, Z. (2004). On diagrammatic representation of mathematical knowledge. In: A. Sperti, G. Bancerek, and A. Trybulec (eds.), *Mathematical knowledge management*. New York: Springer.

Lakoff, G. (1987). W*omen, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. (1980). *Metaphors we live by*. Chicago: Chicago University Press.

Lakoff, G. and Johnson, M. (1999). *Philosophy in flesh: The embodied mind and its challenge to western thought*. New York: Basic.

Lakoff, G. and Núñez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Langer, S. K. (1948). *Philosophy in a new key*. New York: Mentor Books.

Lesh, R. and Harel, G. (2003). Problem solving, modeling, and local conceptual development. *Mathematical Thinking and Learning* 5: 157.

Marcus, S. (2012). Mathematics between semiosis and cognition. In: M. Bockarova, M. Danesi, and R. Núñez (eds.), *Semiotic and cognitive science essays on the nature of mathematics*, pp. 98–182. Munich: Lincom Europa.

McComb, K., Packer, C., and Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behavior* 47: 379–387.

Musser, G. L., Burger, W. F., and Peterson, B. E. (2006). *Mathematics for elementary teachers: A contemporary approach.* Hoboken: John Wiley.

Núñez, R., Edwards, L. D., and Matos, F. J. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational Studies in Mathematics* 39: 45–65.

Peirce, C. S. (1931–1958). *Collected papers of Charles Sanders Peirce*, ed. by C. Hartshorne, P. Weiss and A.W. Burks, vols. 1–8. Cambridge: Harvard University Press.

Piaget, J. (1952). *The child's conception of number.* London: Routledge and Kegan Paul.

Radford, L. (2010). Algebraic thinking from a cultural semiotic perspective. *Research in Mathematics Education* 12: 1–19.

Roberts, D. D. (2009). *The existential graphs of Charles S. Peirce*. The Hague: Mouton.

Royer, J. (ed.) (2008). *Mathematical cognition*. Charlotte, NC: Information Age Publishing.

Shin, S. (1994). *The logical status of diagrams*. Cambridge: Cambridge University Press.

Skemp, R. R. (1971). *The psychology of learning mathematics*. Harmondsworth: Penguin.

Stjernfelt, F. (2007). *Diagrammatology: An investigation on the borderlines of phenomenology, ontology, and semiotics*. New York: Springer.

Taylor, R. and Wiles, A. (1995). Ring-theoretic properties of certain Hecke algebras. *Annals of Mathematics* 141: 553–572.

Thom, R. (1975). *Structural stability and morphogenesis: An outline of a general theory of model*s. Reading: John Benjamins.

Thom, R. (2010). Mathematics. In: T. A. Sebeok and M. Danesi (eds.), *Encyclopedic dictionary of semiotics*, 3rd edition. Berlin: Mouton de Gruyter.

Venn, J. (1880). On the employment of geometrical diagrams for the sensible representation of logical propositions. *Proceedings of the Cambridge Philosophical Society* 4: 47–59.

Venn, J. (1881). *Symbolic logic*. London: Macmillan.

Yancey, A., Thompson, C., and Yancey, J. (1989). Children must learn to draw diagrams. *Arithmetic Teacher* 36: 15–19.