Running Title: An APOS analysis of infinity issues

Some historical issues and paradoxes regarding the concept of infinity: An APOS based analysis, Part 1

Ed Dubinsky, Kent State University
Kirk Weller, University of North Texas
Michael A. McDonald, Occidental College
Anne Brown, Indiana University South Bend

Correspondence:
Ed Dubinsky
265 Northwoods Rd.
Hermon, NY 13652
(315) 386-2787
Email: edd@math.kent.edu
Some historical issues and paradoxes regarding the concept of infinity: An APOS based analysis, Part 1

Ed Dubinsky  Kirk Weller  Michael A. McDonald  Anne Brown
August 13, 2004

Abstract

This paper applies APOS Theory to suggest a new explanation of how people might think about the concept of infinity. We propose cognitive explanations, and in some cases resolutions, of various dichotomies, paradoxes, and mathematical problems involving the concept of infinity. These explanations are expressed in terms of the mental mechanisms of interiorization and encapsulation. Our purpose for providing a cognitive perspective is that issues involving the infinite have been and continue to be a source of interest, of controversy, and of student difficulty. We provide a cognitive analysis of these issues as a contribution to the discussion. In this article, Part 1, we focus on dichotomies and paradoxes and, in Part 2, we will discuss the notion of an infinite process and certain mathematical issues related to the concept of infinity.

KEYWORDS: Actual and potential infinity, APOS Theory, Classical paradoxes of the infinite, Encapsulation, History of Mathematics, Human conceptions of the infinite, Large finite sets

1 Introduction

From time immemorial, men and women have asked questions about the infinite. A wonderment that is the stuff of which our philosophy is made. And our religion. Even our mathematics. As mathematicians, most of us have little difficulty believing that we can think of the natural numbers all at once, as an entity. Yet as very young children, none of us could count more than a few numbers. How did we manage to develop our mental ability to imagine that all of the elements of the set $\mathbb{N}$ of natural numbers have been counted? Indeed, we even seem to be able to consider that this infinite task has been “finished” with enough time left over in the existence of our universe to do other things, like apply transformations to the mathematical object $\mathbb{N}$.

More specifically, by what mental mechanisms do we determine that, while the set $\mathbb{N} = \{1, 2, 3, \ldots\}$ has infinitely many elements, the set $S = \{-3, -2, -1, 0, 1, 2, 3, \ldots\}$ has only 5? And if you think that question is trivial, try convincing some of our undergraduate students.
In addition to wrestling with the issue of whether an infinity of things can be conceived as a single entity, people have also struggled to resolve various paradoxes. Consider the infamous race between Achilles and the tortoise. The slow tortoise is given a head start, but then, although Achilles moves much faster, he must first reach the place where the tortoise has arrived. By then, the tortoise has moved on a little so Achilles must get to this new spot. Still not having caught the tortoise, he must get to the next place, and so on. Since Achilles must accomplish an infinite number of steps, he can never finish, so he never catches the tortoise.

Philosophers, mathematicians, mathematical historians, students, and mathematics education researchers, among others, have struggled to resolve various paradoxes, dichotomies and issues regarding conceptions of infinity. Although in just about every case, there is a clear, rigorous mathematical resolution (for instance, one can see how Achilles wins the race by summing an infinite series), many students have considerable difficulty in understanding these solutions. Our research has led us to the hypothesis that, before being able to understand the mathematical resolutions, an individual must develop certain mental mechanisms and apply them to build mental structures that he or she can use to understand the mathematics. Therefore we have taken for investigation the following theme:

A particular theory of how mathematics is learned, APOS Theory, can be used to analyze, from a cognitive perspective, classical issues related to the concept of mathematical infinity. APOS Theory involves the mental mechanisms of encapsulation and interiorization and it can be used to formulate a new explanation of how people might think about the mathematical concept of infinity.

Our approach is motivated by the following considerations:

1. The question of thinking about infinity has been of interest to mathematicians, philosophers of mathematics, and mathematical historians for at least 3000 years. We believe an APOS analysis contributes something new to the discussion.

2. We choose to use APOS Theory for two reasons. First, the mental mechanisms posited by the theory, in particular interiorization and encapsulation, seem to be part of the “everyday” thinking of mathematicians and many students. We provide analyses that show how these mechanisms can be used to explain, in similar ways across a variety of situations, how individuals may think about various aspects of infinity. Second, the kind of theoretical analyses we suggest for infinity have been made for many topics in the undergraduate mathematics curriculum, and have provided the basis for effective pedagogy (Weller et al., 2003), resulting in improved student learning. The extent to which APOS Theory turns out to be useful in studying the concept of infinity will provide another reason for using this theoretical perspective.

We would like to emphasize that the study on which we report in this paper is neither an historical analysis of the development of the concept of infinity nor a comprehensive study of all of the mathematical, epistemological and pedagogical issues that are related to this concept. Rather, we are looking at historical texts (not always in the original, but as reported by various scholars), the experiences of mathematicians (including ourselves) in teaching the
concept of infinity and preliminary information from ongoing empirical studies as sources for how mathematicians, philosophers and students may be thinking about the infinite. Then, we select particular topics related to infinity which are amenable to an analysis based on APOS Theory, and we make that analysis.

In keeping with our emphasis on selecting topics that appear to be amenable to an APOS analysis, we have not considered all possible infinity-related issues in mathematics and thus, we do not describe the development of the concept of infinity in a chronological manner. In particular, we do not discuss issues such as part-whole vs. bijection in comparing the size of infinite sets. We list below the specific problems that are considered. Although this list is not exhaustive, we feel that it covers a fairly large portion of the conceptual domain of the infinite.

This paper is in two parts. In Part 1, which is this article, we analyze certain dichotomies and paradoxes that have generated controversy among scholars and caused difficulty for students:

- The distinction between potential infinity and actual infinity: Can a human being conceive of infinity as an actuality as opposed to a potentiality?

- The difference between actual infinity and the notion of attainability: Speaking cognitively, when is infinity attainable and when is it unattainable?

- Paradoxes related to infinity: There is a whole genre of paradoxes related to infinite processes (for example, Achilles and the tortoise and what we will call the Tennis Ball Problem). We will propose a general approach for resolving the cognitive conflict in such situations.

- The relationship between finite and infinite phenomena: What is the relation between how humans understand the finite and how they understand the infinite? In particular, what properties of finite phenomena carry over to the infinite? We will also discuss the ways in which very large finite sets are like infinite sets and in what ways they are different.

Our analyses will be used to propose cognitive resolutions of these issues as well as to serve as a basis for the analyses in Part 2.

These problems present potential stumbling blocks in an individual’s understanding of infinity. APOS Theory proposes reasonable cognitive explanations and resolutions. The purpose of this paper will be to describe each issue, show evidence both past and present as to why the issue is or has been problematic, use the theory to analyze the nature of the conflict and suggest a resolution. We use these analyses to show the applicability of APOS Theory in understanding conceptions of the infinite and the power of the mechanisms of interiorization and encapsulation to provide plausible explanations that may have implications for student learning. In fact, the eventual goal of this program of study is to design pedagogy that may help students develop mathematically useful conceptions of various aspects of infinity.

In Part 2, which will appear shortly, the results of the analyses in Part 1 will be combined with preliminary findings from an ongoing study of students’ conceptions of infinite iterative processes to discuss several issues related to infinity that are important in mathematics:
nature of infinite processes and how an individual can conceive of an infinite process as a totality; the cognitive relation between an infinite process and any object(s) it may produce; conceptions of the set of natural numbers; the cognitive relation between an infinite decimal and the real number it represents; and infinitesimals.

The remainder of this introductory section contains an explanation about APOS Theory and its terminology. Section 2 is the main body of the paper in which we give a very brief historical background followed by our cognitive analyses, based on APOS Theory, of the issues we are considering. Finally, in our conclusion we summarize our results and point to future efforts in this research and curriculum development project.

1.1 APOS Theory

APOS Theory adheres to the principle that there is a close relationship between the nature of a mathematical concept and its development in the mind of an individual (Piaget, 1970, p. 13). Thus, the explanations proposed by this theory are both epistemological and psychological.

According to APOS Theory, an individual deals with a mathematical situation by using certain mental mechanisms to build cognitive structures that are applied to the situation. The main mechanisms are called interiorization and encapsulation and the related structures are actions, processes, objects, and schemas. The theory postulates that a mathematical concept begins to be formed as one applies a transformation on objects to obtain other objects. A transformation is first conceived as an action, in that it requires specific instruction as well as the ability to perform each step of the transformation explicitly. For example, an individual who requires an explicit expression in order to think about the concept of function and can do little more than substitute for the variable in the expression and manipulate it is considered to have an action understanding of functions.

As an individual repeats and reflects on an action, it may be interiorized into a mental process. A process is a mental structure that performs the same operation as the action being interiorized, but wholly in the mind of the individual, thus enabling her or him to imagine performing the transformation without having to execute each step explicitly. Thus, for example, an individual with a process understanding of function will construct a mental process for a given function and think in terms of inputs, possibly unspecified, and transformations of those inputs to produce outputs.

If one becomes aware of a process as a totality, realizes that transformations can act on that totality and can actually construct such transformations (explicitly or in one’s imagination), then we say the individual has encapsulated the process into a cognitive object. For the function concept, encapsulation allows one to apply transformations of functions such as forming a set of functions, defining arithmetic operations on such a set, equipping it with a topology, etc.

While these structures describe how an individual may construct a single transformation, a mathematical topic often involves many actions, processes, and objects that need to be organized and linked into a coherent framework, which is called a schema. It is coherent in the sense that it provides an individual with a way of deciding which mental structures to use in dealing with a mathematical problem situation. In the case of functions, it is the schema structure that is used, according to APOS Theory, to see a function in a given
mathematical or “real-world” situation. The mental structures of action, process, object, and schema constitute the acronym APOS.

It is important to note that in dealing with standard mathematical concepts such as function, the mental mechanisms of interiorization and encapsulation and the mental structures of action, process, object, and schema can be used to describe the “everyday” activities of people engaged in mathematical activity. In that sense, we can say that it is reasonable to postulate that such structures are available to many people in mathematical work not explicitly related to the concept of infinity. Thus, our use of these structures to explain how people may think about infinity does not represent the introduction of totally new tools for thinking, but the extension to (not very) different ways of using these tools.

Explanations offered by an APOS analysis are limited to descriptions of the thinking of which an individual might be capable. We do not assert that such analyses describe what “really” happens in an individual’s mind since this is probably unknowable. Moreover, the fact that an individual possesses a certain mental structure does not mean that he or she will necessarily apply it in a given situation. This depends on other factors regarding managerial strategies, prompts, emotional state, etc. In other words, APOS explanations form only part of the story, although a part that has proven useful in various studies. For more information about APOS Theory and a summary of how it has been used in mathematics education research, see Dubinsky and McDonald (2001).

2 Historical and Philosophical Issues

From Aristotle (384-322 BC) on, a key aspect of the concept of infinity has been the distinction between potential infinity, an ongoing activity that never ends, and actual infinity, a definite entity encompassing what was potential. In this section we discuss the history of this distinction and how it can be seen from an APOS point of view. Then we consider certain classical paradoxes related to infinity, Aristotle’s resolution of the paradoxes by rejecting actual infinity, and an alternative solution based on APOS Theory. Finally we consider the relation between finite and infinite phenomena, and observe that many objections to the idea that human beings can conceive of actual infinity also apply, almost verbatim, to certain finite situations such as extremely large finite sets. If one accepts the existence of all finite sets, then this observation would render invalid some of the historical objections to the concept of actual infinity.

It may appear that all of these problems, while sharing some aspects, are quite different and it might not be best to lump them together. What we feel unites and justifies considering them in one study is that it is possible to analyze all of these issues in much the same way and to propose the same kinds of resolutions for each using the APOS mechanisms of interiorization and encapsulation. One of the major results of this study is that these problems, although perhaps different in their formulation, are very similar in their possible resolutions.
2.1 Potential vs. Actual Infinity

Aristotle introduced this dichotomy as a means of dealing with paradoxes of the infinite that he believed could be resolved by refuting the existence of actual infinity. We briefly describe Aristotle’s thinking on this issue and point to its influence on mathematical thought for many centuries. We will also mention historical writings opposing Aristotle’s ideas. Then we explain the dichotomy from the point of view of APOS Theory and propose a resolution based on that theory. Finally, we reflect briefly on why it took so many centuries for the notion of actual infinity to be accepted as being possible both mathematically and cognitively.

2.1.1 Aristotle introduces the dichotomy

Unlike his predecessors, many of whom were interested in thinking about the infinite in metaphysical terms, Aristotle wanted to determine whether anything in space and time is infinite. This led him to define the infinite as that which is untraversable. His definition applied to anything that could be described in terms of an endless process, that is, an infinite sequence of steps in which each succeeding step differs from all of its predecessors. Using this definition, a circle, despite having no beginning and no ending points, was not considered to be infinite, because each succeeding traversal of its circumference is like the first. On the other hand, the natural numbers are considered to be infinite, because in constructing them, the process of adding one always produces a successor which differs from all of its predecessors. Thus, although Aristotle accepted the existence of each natural number, he argued that the totality of all natural numbers was not traversable and so could not be conceived of by human beings. In fact, Aristotle did not consider the natural numbers to be an actual infinite collection. Rather, they represented a potential infinity, in that the process of counting—the act of traversing—could not be completed, because it would require the whole of time. Because our existence is constrained by time, Aristotle believed humans to be incapable of thinking of an infinite process in its entirety. This is why he found the notion of an infinite quantity to be incomprehensible. To him, a quantity was a number, a number was something arrived at by counting, and, given the untraversability of the process of counting there could be no such thing as infinite quantity (Moore, 1999)\footnote{Because this is not an historical study, we are not attempting to consult original sources in all cases. When, as here, we refer to a secondary source, the reader should understand that we are examining this source’s interpretation of the original.}.

However, Aristotle could not reject the infinite completely, as its existence was indicated by a number of considerations: time, which appears to be infinite both by division and by addition; matter, which seems to be infinitely divisible; and space, whose expanse appears to be endless. To reconcile the apparent human inability to conceive of an infinite entity with these aspects of reality that suggest its existence, Aristotle defined two different notions of infinity, potential and actual. This allowed him to acknowledge the existence of the infinite, provided that it was not present “all at once” (Moore, 1999, p. 39).

Aristotle defined actual infinity to be the infinite present at a moment in time. He considered it to be incomprehensible, because the underlying process of such an actuality would require the whole of time. He argued that if the infinite were to be grasped at all, it could only be understood as being presented over time, that is, as being a potential
infinite. As far as Aristotle was concerned, all objections to the infinite were objections to actual infinity, and therefore valid. On the other hand, potential infinity was seen as being a “fundamental feature of reality” and thus acceptable (Moore, 1995, p. 114). Aristotle believed this distinction could be used to resolve various paradoxes, some of which will be considered in Section 2.2.

2.1.2 The dichotomy maintains a stronghold

Aristotle’s potential/actual dichotomy dominated conceptions of the infinite for centuries. Kant (1724-1804), for example, believed that we are finite beings in an infinite world. He viewed the world, which he thought of in metaphysical terms, as being independent of humans, “an absolute complete unified whole” (Moore, 1999, p. 86). Because of our inherent finitude, he argued, we cannot conceive of the whole, but can only receive that which is partial and finite. Each finite part we receive is conditioned, in that there is some condition preceding, or causing it. The condition itself also represents a finite part, which itself is conditioned. This sets up an infinite regress, an endless series of conditions each itself conditioned by some further condition in the series. Aristotle expressed a similar thought: “Something is infinite if, taking it quantity by quantity, we can always take something outside” (Aristotle, as quoted in Moore, 1999, p. 43). It appears that Kant, and possibly Aristotle, may have believed that an infinite process can exist as a totality (the metaphysical whole), but it cannot be conceived by human beings as such (Moore, 1999).

For many philosophers, the possibility of an infinity of past time, together with the actuality of the present, only revealed the paradoxical nature of the infinite and thus did not suggest the existence of actual infinity.

Even more contemporary thinkers such as Poincaré (1854-1912) held largely Aristotelian views. For instance, Poincaré, as quoted in Balduc’s 1963 translation, wrote “there is no actual infinity, when we speak of an infinite collection, we understand a collection to which we can add new elements unceasingly” (p. 47). In his view, an infinite collection entails an indefinite introduction of new elements, each of which may change the way the collection is defined. This makes it impossible to define the collection in its entirety. As a result, the collection cannot exist actually, but only potentially.

Many modern researchers, while acknowledging Cantor’s work in developing transfinite numbers, continue to be strongly influenced by Aristotle’s dichotomy. Fischbein (1987, p. 52) argues that referring to actual infinite sets is “contradictory in natural intuitive terms.” He leaves the door open to the concept of actual infinity as a non-intuitive mental construction when he says that “an actual infinity is a pure logical, conceptual construct, not intuitively acceptable.” Several years later he writes that “the moment we start dealing with infinity — actual infinity — we seem to run into contradictions. . . . [Although we cannot] conceive the entire set of natural numbers, we can conceive the idea that after every natural number, no matter how big, there is another natural number” (Fischbein, 2001, p. 310).

We can see reflections of a development of a concept of actual infinity in students of today. Núñez (1993) reports on a study of the construction of infinite processes (a topic we will return to in Part 2 of this study) by children aged 9-14. He notes that none of his subjects showed any signs of thinking about an actual infinity; all of their comments were in terms of potential infinity. He suggests that the reason is that the concept of actual
infinity does not arise before the age of 15. His view is supported by results of Hauchart and Rouche (1987) who found that some students aged 12-18 did seem to have a concept of actual infinity. For these authors, the notion of actual infinity is tied to the concept of limit. They consider mainly infinite sequences and series, transforming issues such as those that arise in the paradox of Achilles and the tortoise to issues of infinite series. They discuss the relation between potential and actual infinity in terms of the movement from an infinite process to its limit (op. cit., p. 91), but do not explore any cognitive mechanisms for making this transition. They do not say whether there is an actual infinity related to the potential infinity in a sequence or series that does not converge. In Part 2 of this study, we will compare this view to our general explanation using the mechanism of encapsulation, which we present in Section 2.1.5.

2.1.3 Support for the actual infinite

Despite past and current favor toward Aristotle’s views, there were some dissenters, such as the rationalists, who encouraged acceptance of actual infinity. The rationalists believed we could invoke pure reason to understand the world around us. In applying reason, they believed it was possible to transcend one’s finitude, because we are not bound solely by our physical limitations.

An early advocate of the actual infinite was Rabbi Hasdai Crescas (1340-1410) whose views may have influenced later thinkers like Galileo (1564-1642) and Bolzano (1741-1848). In the following excerpt, we see that his notion of “an infinity of things known” was motivated in part by religious belief.

It may be suggested that [God’s] knowledge does not extend to details as such, which is to say this set of three [things] or four, or any specific set of three or four [things]. But, since there is no escaping the fact that His knowledge must include some number, would that I knew whether He knows all the other numbers as well or not. Now, if He knows them, since number can be added to without end, then His knowledge extends to an infinity of numbers. If He does not know all of them, there must of necessity be a bound beyond which He does not know. But then the question remains — why is it that He knows the numbers up to that bound but does not know greater ones. Have weariness and fatigue befallen His knowledge? There is no avoiding the conclusion that there is an infinity of things known (Crescas, as quoted by Rabinovitch, 1970, p. 228).

Many rationalists who followed Crescas also appealed to their religious beliefs to justify the existence of actual infinity. According to Descartes (1596-1650), we could invoke reason to touch the idea of the infinite which God had imparted to us, “even though our inherent finitude precluded us from experiencing or imagining it” (Moore, 1999, p. 76). In other words, our a priori capacity to reason is a gift from God that allows us to transcend our inherent limitations. The infinite, inherent in God’s nature, was, for Descartes, something real which transcended the merely indefinite.

I never use the word ‘infinite’ to mean only what has no end, something which is negative and to which I have applied the word ‘indefinite’, but to mean something
real, which is incomparably bigger than whatever has an end (Descartes, as quoted by Moore, 1999, p. 77).

Bolzano played a major role in advancing the notion of an infinite totality when he rejected Aristotle’s assertion that a collection does not exist as a completed whole unless one forms an image of every item, or reflects on every step of the process that generates it.

‘Nowhere can an infinite set exist,’ they say, ‘for the simple reason that infinite sets can never be united to form a whole, never collected together in thought.’ I must stigmatisé this assertion a mistake, and a mistake engendered by the false opinion that a whole consisting of certain objects \(a, b, c, d, \ldots\) cannot be constructed in thought without forming separate mental representations of its separate components (Bolzano, as quoted in the 1950 translation by Prikhonsky).

In other words, one can think of a set simply by describing its elements; we can use our minds to conceive of an infinite collection as being complete without having to think of each element individually. This led Bolzano to see an infinite collection as a totality.

I can think of a set, of the aggregate, or if you prefer it, the totality of the inhabitants of Prague or of Pekin without forming a separate representation of each inhabitant (Bolzano, op. cit.).

One concern with Bolzano’s defense of actual infinity is that the example he used is a finite set! We will return to this issue below in Section 2.3.

2.1.4 Actual infinity and the notion of attainability

In Moore’s view, Cantor’s (1845-1918) theory of transfinite arithmetic perpetuated, in some ways, Aristotle’s dichotomy between potential and actual infinity. He reasons that because sets come into being after their members, their collective infinitude can only be considered potential, not actual. Furthermore, Moore (1999, p. 198) points to Cantor’s belief that, because infinite sets are subject to mathematical investigation, those sets whose cardinality or order type can be assigned “enjoy a kind of finitude”, or are “really finite.” In this way, Moore argues that Cantor distinguished them from collections that are genuinely infinite, those that are “endless, unlimited, and unsurveyable” (Moore, 1995, p. 116).

Cantor’s view that collections to which a value or order cannot be assigned are immeasurably large, or unattainable, is in evidence in this passage from one of his letters to Dedekind (1831-1916):

A multiplicity can be such that the assumption that all its elements ‘are together’ leads to a contradiction, so that it is impossible to conceive of the multiplicity as a unit, as ‘one finished thing.’ Such multiplicities I call absolutely infinite or inconsistent multiplicities. As we can readily see, the ‘totality of everything thinkable,’ for example, is such a multiplicity (Cantor, as quoted in Rucker, 1982, pp. 48-9).
Cantor’s work thus suggests three cognitive categories: the finite, the attainably infinite, and the unattainably infinite. Actual infinite entities are considered to be attainably infinite, while potentially infinite collections which cannot be actualized are considered to be unattainably infinite. The unattainably infinite is what Moore describes as “really infinite.” In this paper we subscribe to Cantor’s point of view and assert that there exist infinite entities which are attainable, at least cognitively speaking.

2.1.5 APOS explanations

In this section, we describe how APOS Theory can be used to support Cantor’s point of view. We describe our hypotheses concerning how the mechanism of interiorization is involved in the mental construction of infinite processes, and how encapsulation relates to the issues of conceiving of the actually infinite, distinguishing between the attainable and the unattainable, and to the potential/actual dichotomy.

In defining the infinite as that which is untraversable, Aristotle viewed the infinite as a process that could not be completed. By asserting that one cannot think of a collection in its entirety without reflecting on each of its elements, or a process without physically or mentally performing each step, Aristotle expressed his ideas in temporal terms, thereby arguing that it is impossible to conceive of the infinite as a completed totality. Bolzano asserted that he was free of such temporal constraints and could conceive of a process, such as the formation of a set, as a completed totality without considering each element. Unfortunately, his main example was a finite set. Nevertheless, he appeared to be thinking of the process of constructing a set as a completed totality. Our APOS analysis is that the individual steps are actions and the ability to imagine an infinite set of steps as having been completed is possible by means of the mechanisms of coordinating these actions and interiorizing them into a process. This process can then be seen as a completed totality which is a crucial aspect of the mental mechanism of encapsulation. In this sense, we might argue that Bolzano’s thinking is evidence in support of our analysis.

Cantor extended Bolzano’s thinking. In establishing the existence of infinitely many transfinite cardinal and ordinal numbers, and defining arithmetic on cardinal and ordinal numbers, he was performing mathematical operations. Our APOS analysis is that Cantor had encapsulated the infinite processes involved in constructing cardinal and ordinal numbers and was able to think of them as objects to which actions and processes (e.g., arithmetic operations, comparison of sets) could be applied.

APOS Theory can help us to understand the distinction between the potential and the actual, and between the attainable and the unattainable. Potential infinity is the conception of the infinite as a process. This process is constructed by beginning with the first few steps (e.g., 1, 2, 3 in constructing the set $\mathbb{N}$ of natural numbers), which is an action conception. Repeating these steps (e.g., by adding 1 repeatedly) ad infinitum requires the interiorization of that action to a process. Actual infinity is the mental object obtained through encapsulation of that process. The underlying infinite process that led to the mental object is still available and many mathematical situations require one to de-encapsulate an object back to the process that led to it. Thus we would say that, through encapsulation, the infinite becomes cognitively attainable. On the other hand, the unattainable is an instance of the infinite, in the form of a process, that has not been encapsulated. This may happen because
the process is not yet seen as a totality, either because it cannot be seen as a totality, or because no encapsulation has taken place.

Our APOS analysis indicates that once the individual can see all of the steps of an infinite process as a single operation that can be carried out and finished, and present at a moment in time, he or she can conceive of the infinite as a completed totality. Once this totality is encapsulated, the notion of potentiality is transformed into an instance of actual infinity, a mathematical entity to which actions can be applied. Hence, the existence of the one does not negate that of the other, nor is either a misconception with respect to the other. Instead, the potential and actual represent two different cognitive conceptions that are related by the mental mechanism of encapsulation. These conceptions and their relationship becomes part of the individual’s infinity schema.

The existence of processes that cannot be conceived as totalities and encapsulated is either a mathematical issue (as in the set of all sets) or a theological matter (only God can attain the unattainable). From the cognitive point of view, it is not unreasonable in either case to accept the possibility of thinking about such unmanageable sets, as long as it is understood that the ability to see something as a totality and the cognitive mechanism of encapsulation may not always be available.

We see that the “truly infinite” to which Moore alludes really refers to the unattainable, both cognitively and mathematically. The collection of transfinite cardinal and ordinal numbers is cognitively unattainable to the degree that it cannot be thought of as a completed totality. Nor is it mathematically attainable: assigning a cardinal to the collection of all cardinals or an ordinal to the sequence of all ordinals leads to a contradiction. On the other hand, using our APOS analysis we argue that the set of natural numbers is cognitively attainable by means of encapsulation. Yet, this does not make the set of natural numbers any less infinite. The only difference is that the set of natural numbers can be conceived as an entity to which actions can be applied. The unattainable cannot.

2.1.6 The significance of encapsulation

If the mental construction of actual infinity and other mathematical objects such as functions simply require one to encapsulate, one may wonder what all the fuss is about. Specifically, why did issues surrounding the dichotomy ever surface in the first place? Although encapsulation refers to but a single mental mechanism, it entails a radical shift in the nature of one’s conceptualization. Rather than thinking in terms of an interiorized series of actions that are applied to an existing object or objects, encapsulation signifies the ability to think of the same concept as a mathematical entity to which new, higher level transformations can be applied. Although a process is transformed into an object by encapsulation, this is often neither easy nor immediate. We can see this difficulty throughout history, as indicated by the following examples in which we claim that the difficulty of encapsulation appears in some form:

- Kleiner, 2001: pp. 153-157 (Newton (1642-1727), Berkeley (1685-1753), and Leibniz (1646-1716) on the issue of infinitesimals)
- Knobloch, 1999: pp. 91-92 (Galileo on the transition from finite to infinite in the context of the paradox of concentric circles marking off the same distance when each
is rolled through one revolution); p. 97 (Leibniz on the need for a leap in going from the finite to the infinite)

- Moore, 1999: p. 10 (the question of whether infinity is too big to count as one); pp. 75-83 (opposing perspectives of the rationalists and the empiricists); pp. 86-87 (Kant’s explanation of why human beings cannot conceive of the infinite); pp. 90-100 (Hegel’s (1770-1831) discussion of the mathematically infinite as a succession of finite elements, as well as a comparison/contrast between his views and those of Kant and Aristotle); p. 102 (Taylor’s (1869-1945) contention that that the truly infinite is indicated by a self-consistent internal structure); pp. 110-117 (the impact of Bolzano and Cantor); p. 132 (a discussion of Cantor and Brouwer’s (1881-1966) views on the issue of whether the sets \( \mathbb{R} \) and \( \mathbb{N} \) be conceived as wholes); pp. 135-136 (Weyl (1885-1955) and Hilbert’s (1862-1943) ideas about about actual infinity)

- Sierksma and Sierksma, 1999: pp. 444-448 (Bernoulli (1667-1748) and Leibniz’s correspondence regarding the question of whether an infinite sequence has an infinitesimal element as one of its terms)

- Fischbein, 2001: p. 310 (issue about the equality of .333\ldots and 1/3 and whether one can conceive of \( \mathbb{N} \) as a totality)

- Richman, 1999: p. 397 (actual and potential infinity and the relation of these notions to students’ difficulties in determining whether .999\ldots = 1)

The delay often seen in the development of encapsulation may explain why the idea of function as an object and the field of functional analysis did not really develop until the 20th century, why it took centuries for the notion of actual infinity to be widely accepted, and why many still find the dichotomy between potential and actual infinity to be perplexing.

### 2.2 Paradoxes related to infinity

Throughout the history of mathematics a multitude of paradoxes of many different types has been discussed in the literature (see Moore, 1995 and Moore, 1999, pp. 1-13 for a partial listing). These include paradoxes of self-reference (Can a set be a member of itself?), paradoxes of time (If time goes on indefinitely has it always gone on?), mechanical paradoxes (If two concentric circles are rolled through one revolution, how can they cover the same linear distances since their circumferences are different?), and many other types. We discuss here the notion of an infinite iterative process paradox, of which there are many examples. We will then see how one might resolve such paradoxes using exactly the same mental mechanisms (interiorization and encapsulation) as were used in considering how an individual can think about actual infinity.

We begin the APOS analysis of some of the classical paradoxes by explaining what we mean by an infinite iterative process\textsuperscript{2} paradox and then focus on two specific examples: Achilles and the tortoise and the placement of infinitely many tennis balls on a table.

\textsuperscript{2}This term is introduced by A. Brown, M. McDonald and K. Weller in an as yet unpublished work.
2.2.1 Infinite iterative process paradoxes

Generally speaking, an infinite iterative process paradox involves an iterative process which consists of an action repeated without end. The paradox deals with the conception of what happens when the process is complete. Two examples are the famous race between Achilles and the tortoise (Moore, 1995) and the placement of infinitely many tennis balls on a table (Falk, 1994).

Achilles and the tortoise. According to Moore (1995), “In this conundrum the swift demigod challenges the slow tortoise to a race and grants her a head start. Before he can overtake her, he must reach the point at which she began, by which time she will have advanced a little. Achilles must now make up the new distance separating them, but by the time he does so, she will have advanced again. And so on, ad infinitum. It seems that Achilles can never overtake the tortoise” (p. 112).

Tennis balls. Suppose that an infinite set of numbered tennis balls and a large table are available. Place balls numbered 1 and 2 on the table and remove number 1. Next, place balls 3 and 4 on the table and remove number 2. Place balls 5 and 6 on the table and remove number 3. And so on, ad infinitum. It seems that one cannot determine how many balls are on the table at the end of the activity because, on the one hand, the number increases by one at each step so there are an infinite number, but on the other hand, given any tennis ball, one can say exactly when that ball was removed so that none is left (Falk, 1994, p. 44).

We can see infinite iterative processes in each of these situations. In the case of Achilles and the tortoise, we have two such processes. In one process, each iteration consists of Achilles traveling the distance between his position and the position of the tortoise at the start of the iteration. In the other process, each iteration consists of the tortoise traveling a certain distance during the time Achilles is traversing the distance between them at the start of the iteration. The iterations continue indefinitely. Note that the two processes are coordinated in that each iteration of Achilles’ process makes use of what happened in the previous iteration of the tortoise’s process. The paradox consists in the idea that since the two processes never end, that is, cannot be completed, Achilles does not catch up with the tortoise.

In the case of the tennis balls, an iteration consists of placing the next two balls on the table and removing the ball on the table with the lowest number. Again, the iteration continues indefinitely. Here the paradox arises in trying to determine how many tennis balls are on the table when the process is completed.

These two situations differ in one respect. In the case of Achilles and the tortoise, one can imagine in the “real” world an actual race and, clearly, if Achilles runs faster than the turtle, he will catch up eventually. The placement of the tennis balls, however, has no relation to any process that could actually go on (in toto) in the “real” world. Thus, one process (actually a coordinated pair of processes) is a particular way of thinking about a situation in the “real” world whereas the other has to do with a situation that can exist only in one’s mind.
The classical resolution of paradoxes such as that of Achilles and the tortoise is based on the distinction between potential and actual infinity. According to Aristotle, what we have called an infinite process represents potential infinity, whereas thinking about the completed process requires the notion of actual infinity, which Aristotle and his followers rejected. Thus, according to this point of view, although we can describe the process of Achilles covering the distance, the tortoise moving on a little further, Achilles covering again the distance, and so on, we cannot think about the total process or ask about its effects. A similar analysis would conclude that we can think about the repetition without end of the process of putting two tennis balls on the table and removing one, but we cannot think about what happens when the process does end, i.e., is completed.

2.2.2 APOS analyses of infinite iterative process paradoxes

We recall that according to APOS Theory, a process is a mental structure that enables an individual to imagine transforming an object, or objects, without having to perform explicitly all of the steps of the transformation. As we have seen, the theory postulates that an individual who can think of a function as a process can imagine operating on an element of the domain to obtain an element of the range without actually needing to apply a formula to carry out a specific rule of assignment. In fact, an individual with a process conception of function can think of a function even in a situation where a formula is not available or does not exist. Rather, using the process structure, the individual can imagine the transformation of domain objects to range objects, to a greater or lesser degree of specificity, without actually performing the transformation.

It is in this sense that an individual could use the infinite iterative process mental structure to think about a concept. Performing a small number of iterations constitutes an action. Thus, Achilles covering the distance previously covered by the tortoise as the tortoise moves forward is considered to be an action. Placing two tennis balls on a table and removing a tennis ball would also be actions. An individual could think explicitly of performing these actions once or a few times, but, at some point, this would not be possible because each repetition must be thought through explicitly. Only by interiorizing these actions and using the resulting process structure could the individual imagine repeating the actions indefinitely, or “forever,” so to speak. This corresponds to potential infinity.

According to APOS Theory, when an individual has constructed a process in her or his mind and wishes to perform an action on this process the mental structure of encapsulation comes into play. Encapsulation consists in transforming the process to an object so as to apply the desired action. Before using this in offering explanations of our paradoxes we again wish to emphasize that encapsulation of processes in order to perform actions is an “everyday” activity in mathematics. For instance, we return to the example of functions and consider adding two functions, say $f$ and $g$, to obtain a new function. Thinking about doing this requires that the two functions be conceived as objects. The actual transformation is imagined by de-encapsulating back to the processes and thinking about all of the elements $x$ of the domain and all of the transformations $f(x)$ and $g(x)$ at the same time so as to obtain the new process which consists of transforming each $x$ to $f(x) + g(x)$ (a coordination of the processes for $f, g$). This new process is then encapsulated to obtain the new function $f + g$. In many cases, the domains and ranges of functions are infinite sets. Hence, we can see how
the mechanisms of interiorization and encapsulation describe how an individual might be able to think about infinite sets in these contexts.

Given an infinite process, moreover, the mental mechanisms of interiorization and encapsulation describe how one may come to think about what happens after the process is completed. The objection that this cannot be done, since one can never actually perform an infinite number of steps, is precisely what the structure of process deals with since, assuming an individual uses such a mechanism, he or she does not have to actually perform all of the steps, whether there be finitely or infinitely many of them.

In the case of Achilles and the tortoise, we have the two processes of Achilles repeatedly covering the previous distance traveled by the tortoise while the tortoise continually moves further along. The structure of process explains how we could imagine this going on forever. The individual processes would be coordinated and the resulting infinite process encapsulated so that the action of comparing the distances covered by the Achilles and the tortoise could be applied. At this point, with a cognitive understanding of the question, one could do the calculations (which amount to summing infinite series) and see that in a finite time, the total distance covered by Achilles exceeds that covered by the tortoise.

One can describe the paradox of the tennis balls in a similar manner. According to APOS Theory, one can imagine the completed process of placing balls on the table and removing balls from the table. This process could be encapsulated so as to compare the resulting object, the set of all tennis balls remaining on the table after the process is completed, with the empty set. It is then easy to see that given tennis ball $n$, it is placed on the table at the step which is the smallest integer greater than $\frac{n}{2}$ and removed from the table at step $n$. Thus there are no tennis balls remaining on the table and the object in question is equal to the empty set. At any given step of the process, however, the number of balls on the table exceeds that of the previous step. This represents an example in which the completion of a process leads to an object which is not produced by the process itself but rather by encapsulating the process.

Of the many classical paradoxes that can be interpreted as infinite iterative process paradoxes and therefore treated in the same way, we would include Zeno’s paradox of the race course, the Staccato run, Hilbert’s hotel with infinitely many rooms, the paradox of the arrow, the paradox of the lamp, and the paradox of the spaceship (see Moore, 1999, pp. 1-13 for detailed statements of these paradoxes). The similarities of these treatments, together with the treatment of the actual vs. potential dichotomy, help the individual to organize the mental structures discussed in this section into an infinite iterative process paradox schema that can be used in dealing with certain situations involving infinity.

### 2.3 Finite vs. Infinite

There is no doubt in mathematics about what is finite and what is infinite. Since Dedekind, mathematicians have agreed that an infinite set is one which can be put in 1-1 correspondence with a proper subset; a finite set is a set which is in 1-1 correspondence with a set \{1, 2, \ldots, n\} for some $n$; and every set is either infinite or finite. However, Moore’s discussion (Moore, 1999, pp. 1-13 for detailed statements of these paradoxes).

---

3 Of course, once the encapsulation is made, some students may still have difficulties with understanding the empty set. This is an additional issue not related to the concept of infinity.

4 A more detailed analysis of such situations is made in a current study by Brown, McDonald and Weller.
1995) of the distinction between what is “really” finite and what is “really” infinite, as well as his comment that sets like \( \mathbb{N} \), because of Cantor’s theory, “enjoy a kind of finitude,” suggests that philosophically, and perhaps cognitively, the distinction may not be so clear. Indeed, in considering continued subdivision of a segment into smaller and smaller pieces, Galileo, through the words of Salviati, states that the number of quantified parts in the bounded continuum are “neither finite nor infinite.” His opponent, Simplicio, is surprised at the idea that there could be a middle ground between finite and infinite, which suggests to him that “the dichotomy or distinction that makes a thing finite or else infinite were somehow wanting and defective.” To this, Salviati replies:

It seems to me to be so. Speaking of discrete quantity, it appears to me that between the finite and the infinite there is a third, or middle term; it is that of answering to every designated number. Thus in the present case, if asked whether the quantified parts in the continuum are finite or infinitely many, the most suitable reply is to say “neither finite nor infinitely many, but so many as to correspond to every specified number.” To do that it is necessary that these be not included within a limited number, because then they would not answer to a greater [number]; yet it is not necessary that they be infinitely many, since no specified number is infinite. (Galilei, as quoted in the 1974 translation by Drake.)

In this section, we introduce another candidate for quantities that, cognitively, seem to fall between finite and infinite, what we will call \emph{large finite numbers}. We propose that the cognitive difference between large finite numbers and the infinite relates to the issue of “final objects.” For the former, the last number enumerated indicates its completion. For the latter, there is no final step, or number enumerated, so the ability to think of an infinite process as being complete seems more difficult to develop. We also consider two other issues regarding the relationship between the finite and the infinite: how an individual might make the mental transition from the one to the other, and the question of which properties carry over when such a transition is made.

\subsection{Large finite numbers}

As we have seen, many of the objections throughout history to the ability of human beings to think about the steps of an infinite process being present all at once, or about an actual infinity, rest on the fact that no human being can actually enumerate an infinite collection of steps. Infinite sets, however, are not the only ones which human beings cannot enumerate. For example, no individual can actually construct, either physically or by reflecting on each individual element, a set with \( 10^{10^{10}} \) elements, nor can an individual count the grains of sand in a desert. If we take the empiricist view that a process exists in totality only if one can physically or mentally reflect upon and complete each step, then we must deny the existence of such finite numbers.

Since it is very difficult for any but the most radical empiricist to deny the existence of large finite numbers, one might argue for their existence by saying that an individual can \emph{imagine} counting all of the grains of sand in the desert. However, if a process is said to exist because it can be imagined, then one must also accept the existence of the natural
numbers, because an individual can imagine it as a completed totality. Therefore, it seems that the non-existence of an infinite collection cannot rest on either the physical impossibility of enumerating all its elements or the inability to imagine doing so, because in either case, one would then have to deny the existence of large finite numbers.

So we must turn to another criterion for distinguishing, cognitively, between the finite and the infinite. In enumerating $10^{10^{10}}$ objects, there is a final object that indicates completion of the process. This is not so with the natural numbers. Our APOS analysis is that in order to conceive of an infinite process as being complete, which is a necessary precursor to encapsulation and the application of actions, an individual would have to realize that enumerating the entire set can be described as the repeated application of a single action even though that individual recognizes that such a process does not produce a final object.

There is also a difference in an individual’s conception of “all at once.” In the case of enumerating the first $10^{10^{10}}$ natural numbers, an individual can reflect on the last counting number as indicating the cardinality of the set. This cannot occur with an infinite set, because there is no last counting number. Rather, according to our APOS analysis, an individual would have to be able to conceive of an infinite process as a totality, a single operation that can be thought of at a moment in time.

In an APOS analysis, conceiving of a process as complete and as a totality is required to encapsulate the process into a cognitive object, as is the desire to apply an action to the process. In the case of an infinite process, the object that results from encapsulation transcends the process, in the sense that it is not associated with nor is it produced by any step of the process. This is a major finding of a recently submitted study of infinite iterative processes conducted by Brown, McDonald and Weller. The authors have termed such an object the transcendent object of the process.

In general, a process may or may not produce an object, although this is not necessarily the object obtained by encapsulating the process. Consider, for example, the process of adding the first $n$ natural numbers. This process produces the number $\frac{n(n+1)}{2}$ and if we think of this process as an iterated process of adding the first natural number, the first two natural numbers and so on, then $\frac{n(n+1)}{2}$ is produced by this process as its $n^{th}$ term. But $\frac{n(n+1)}{2}$ is not the object obtained from encapsulating the process. Suppose, for example, we wish to compare this process to, say the process of adding the squares of the first $n$ natural numbers. This would be an action on the two processes and so each process would first have to be encapsulated so that a comparison can be made. The objects that result from the two encapsulations are not the values of the sums but the series themselves to which the action of comparison is applied. In the case of infinite processes, this distinction between final and transcendent objects becomes very important and it is expected that the work of Brown, McDonald and Weller will shed more light on it.

This distinction between final and transcendent objects raises the question of the adequacy of Lakoff and Núñez’ (2000) Basic Metaphor of Infinity as an explanation of how people conceptualize the difference between the finite (even large finite) and the infinite. The subtle differences between the “last object” and the “transcendent object” may explain why it appears to be easier for an individual to think of or accept the existence of a large finite number and why her or his ability to see or to accept the existence of an actually infi-
nite set is apparently more difficult. Whether this is actually the case has not been verified empirically. We indicate the possible existence of three conceptions: small finite, that which can be physically and cognitively completed; large finite, that which cannot be physically completed, but whose cognitive completion is similar to small finite; and the infinite, that which is physically uncompletable and cognitively completable only when certain mental constructions not required in the finite case are made. Whether these three distinctions actually occur in an individual’s mind is a topic for future research.

2.3.2 Do properties carry over from finite to infinite?

There appears to be a strong human desire, when extending a concept, to carry over properties with which one previously felt comfortable. This seems to be particularly natural when the transition from finite to infinite is thought of as being gradual and not requiring major effort. Thus, Leibniz extended arithmetic operations with real numbers to include calculations involving infinitely small quantities, which he referred to as infinitesimals. Cantor extended the notion of order in defining the transfinite ordinals, and he generalized the concept of quantity in developing transfinite cardinals. He even devised rules for combining transfinite ordinal and cardinal numbers—an arithmetic of the infinite.

Although it is very natural to think this way, it can lead to errors. For instance, in a study of elementary school children, Falk (1994) found that while children age 8 and older understood the endlessness of the natural numbers, they did not grasp the difference in magnitude between a large infinite set and \( \mathbb{N} \). When asked to locate sets whose magnitudes they had compared on a long linear strip, many children older than age 8, and some even as old as 12 or 13, placed \( \mathbb{N} \) adjacent to grains of sand in a desert, because they did not believe the quantity of natural numbers to exceed the number of grains of sand by much. One study reported that 31% of 190 school aged students, when asked whether \( \infty \) is an enormous number, responded yes, and in interviews often made statements such as “we think of it as a number to simplify things” (Monaghan, 2001, p. 246).

Tsamir (1999) found that prospective teachers erroneously attribute properties of finite sets to infinite sets. Other studies show that many students consider the limit of a sequence to be the last term of the sequence and, given the sequence \( (a_n)_{n=1}^{\infty} \), will write \( a_{\infty} \) for its limit (Mamona-Downs, 2001).

Also, in a study of first-year university students at Warwick, Tall and Schwarzenberger (1978) found that some students think of infinity as a last number. Their responses indicated that some of their ideas may be described as:

- infinity is a concept invented in order to give an endpoint to the real numbers, beyond which there are no more real numbers;
- infinity is the biggest possible number that exists;
- infinity is a number which does not exist, but is the largest value for any number to have;
- infinity is the last number in a never ending chain of numbers.
Similarly, in a study of 31 sixteen year-old precalculus students, Sierpińska (1987) found that some students think of the infinite as being a large finite number.

In trying to understand and possibly help students overcome these difficulties, we must realize that encapsulation is a substantial mental mechanism requiring time and effort on the part of the individual. When it occurs, it can feel like a somewhat abrupt step to a new level of thought. We do not think it is unreasonable to suggest that focusing on such a mechanism might make it less likely that a student will fall into the seemingly more natural and almost effortless, albeit erroneous, notion that properties of the finite carry over, with necessary changes, to the infinite.

3 Conclusion

In this paper we have considered several issues related to infinity that have been of concern to mathematicians and philosophers for the last 3000 years: the potential versus actual infinity dichotomy and its effect on views of the infinite; analyses of several infinite iterative process paradoxes; the relationship between large finite numbers and infinite quantities; and the transition in thought from the finite to the infinite. Our goal has been to use APOS theory to analyze many of these difficulties and, in some cases, to suggest possible resolutions of them for which a cognitive explanation could be applied. In each case, we ascribed a central role to the mental mechanisms of interiorization and encapsulation as a means of explaining how the human mind can conceive of the infinite and deal effectively with mathematical problem situations involving the infinite.

The main contribution that we obtain from an APOS analysis is an increased understanding of an important aspect of human thought. We feel that the value of this understanding is enhanced by the fact that it is based on a small number of mental mechanisms, mainly interiorization and encapsulation. These two mechanisms have been crucial in the study of many other mathematical concepts and have pointed to effective pedagogy. Now we see that they can also provide plausible explanations of a wide variety of infinity-related issues that are of concern to mathematicians and philosophers and which cause difficulties for students of mathematics. Surely a first step in helping students overcome these difficulties is to understand their nature.

As we expand our analyses to other mathematical topics related to infinity, the next major step in this research program is to develop pedagogical strategies based on these analyses. Such strategies should focus on helping students interiorize actions repeated without end, to construct infinite iterative mental processes and to encapsulate these processes into infinite objects. What APOS Theory suggests is that interiorization comes from repetition and reflection whereas encapsulation results from a desire to apply transformations to these processes. The design of activities which include opportunities to reflect on repeated actions and that require transformations of infinite processes will be our starting point in future work to improve student understanding of the infinite.
4 References


August 13, 2004