INTIMATIONS OF INFINITY

Kirk Weller, Anne Brown, Ed Dubinsky, Michael McDonald, Cynthia Stenger

> *The natural numbers? The natural numbers are* $1, 2, 3 \ldots, \infty$*, but there is no such number as* ∞*, there is nothing you can think of as a concrete value.*

> *There is no actual infinity and when we speak of an infinite collection, we understand a collection to which we can add new elements unceasingly.*

Overture

The comments above (see below for who made them and when) represent one type of thinking about infinity. There are other types, as we will see, and they all create difficulties for students, philosophers and even mathematicians. The purpose of this article is to show how a particular theory about how people come to understand mathematics, APOS Theory, can be helpful in understanding the thinking of both novices and practitioners as they grapple with the notion of infinity. In APOS theory, which will be more fully explained later, an individual develops an understanding of a concept by employing certain mechanisms called *interiorization*, *encapsulation* and *thematization*. These mechanisms are used to build and connect mental structures called *actions, processes, objects* and *schemas*.

To get a feeling for the complexity of how people grapple with infinity, see how you and perhaps some of your colleagues would answer the following questions. How do you think your answers compare with what has been said by western mathematicians and philosophers over the last 3000 years, or by students today?

- If the slow tortoise starts a little ahead of the swift Achilles, how can this demi-god ever catch up? For, Achilles must first advance to where the tortoise started, by which time the plodder has moved on a little, so Achilles must then advance to that spot, and so on, forever.
- How can the quantity dx be treated both as a positive quantity with which calculations can be made, and something that can be ignored as if it were 0?
- Is $0.999\ldots = 1$?
- Suppose you put two tennis balls numbered 1 and 2 in Bin A and then move ball 1 to Bin B, then put balls 3 and 4 in Bin A and move 2 to Bin B, then put balls 5 and 6 into Bin A and move 3 to Bin B, and so on without end. How many balls are in Bin A when you are done? Infinitely many because the number increases by one each time, or none since every ball is eventually removed?
- If you build a set by putting in the integer 1, then 2, then 3 and so on, how do you get from this unending process to a conception of the full set of natural numbers?

- Is the infinite union $\bigcup_{k=1}^{\infty} P(\{1, 2, \ldots, k\})$ equal to the power set of the natural numbers $P(\mathbf{N})$? Here, P stands for the power set operator.
- Is there any sense in which an uncountable set can be the outcome of countable algorithm?

For each of these questions, rigorous formal thought provides answers on which mathematicians can agree. But agreement has not come easily, students often have a hard time accepting the formal solutions, and describing the infinite can be difficult. For example, the first statement at the very beginning of this article is a paraphrase of what a student said recently in describing her conception of the natural numbers. The second statement is a quote from H. Poincaré [10] almost a century ago.

The following examples illustrate the thinking of students, mathematicians and philosophers on the above bulleted questions. They indicate some of the many aspects of the infinity concept, the variety of approaches taken to deal with these aspects, and some of the difficulties encountered by experts and students in their efforts to understand the infinite.

Achilles and the tortoise. All objections to the infinite, Aristotle insisted, are objections to the actual infinite. The potential infinite, on the other hand, is a fundamental feature of reality. Aristotle used this distinction between the two types of infinity to resolve paradoxes like Achilles and the tortoise.[7]

Infinitesimals. "And what are these fluxions? The velocities of evanescent increments. And what are these same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities?" [1, p. 83]

Is 0.999 . . . = 1? Responses to this question from university students included:

"Just less than one, but it is the nearest you can get to one without actually saying it is one."

"It is just less than one, but the difference between it and one is infinitely small."

"The same, because the difference between them is infinitely small." [12]

Tennis balls. During interviews of college students for a current research project, one said: ". . . they're both gonna contain half the balls." Another claimed that Bin A contained "two infinity minus infinity which would be infinity." Yet another said: "A doesn't really have a limit on how big it is. . . so A goes to infinity." One student felt that ". . . you cannot decide what's gonna be in A." Only one student thought that Bin A would be empty. [11]

Getting N from a process. "It [the natural numbers] means this is the collection of things 1,2,3 and . . . then I keep adding 1 . . . and now I'm going take the union over all those sets 1 through n . . . What annoys me about this is that when I take that union . . . somehow you have to know in advance what the integers are before you can take the union over all the finite truncations of the natural numbers." *A research mathematician's response during an interview for a current research project.*

The result of taking an infinite union. During an interview, a college student tried to determine whether the infinite union $\bigcup_{k=1}^{\infty} P(\{1, 2, \ldots, k\})$ is equal to $P(\mathbf{N})$. She noted that the power sets are nested, and thought of the union iteratively in terms of an infinite sequence $P(\{1\})$, $P(\{1, 2, \})$, ..., $P(\{1, 2, \ldots, k\})$, ..., but then remarked, "And then you keep adding one and you'll still have finite sets but eventually you have to, I just *still* want to include that infinity!" [3]

Getting an uncountable set from a countable algorithm. "I don't believe the power set of the natural numbers is countable and you have presented what appears to be a countable process for creating this set." *A college teacher's response during an interview for a current research project.*

"A procedure that purports to construct $P(\mathbf{N})$ only gives an illusion of construction." *From a private conversation with a set theorist.*

The issues raised here are controversial for both mathematicians and philosophers, and have been for centuries. In this article, we describe several of our investigations (some completed, some ongoing, and some only contemplated) into the issues raised by the variety of thinking about infinity that history and current discourse provide. We illustrate how research in mathematics education can contribute to resolving classical issues related to the concept of infinity, and point out how this same research can also explain certain difficulties we see in students who are trying to understand infinite processes and objects in mathematics.

We will explain how the APOS theory of learning helps us analyze the type of thinking exemplified in the quotes given above. Specifically, we use the theory to argue that human beings *can* and *do* conceive of an infinite process as a totality and think about actual infinity, that one can apply certain mental mechanisms to think about the set of natural numbers as a totality, and that there is a sense in which those students who claim that $.999\ldots$ is not the same as 1 are right. And, though we will leave the full details for a future article, we will also suggest that the same mental mechanism that allows an individual's thoughts to shift from enumerating the natural numbers one-by-one to considering the entire set of natural numbers can also be used to make the leap from thinking about a countable process to seeing an uncountable set such as $P(\mathbf{N})$ as a mental object.

Motivation

The concept of mathematical infinity appears throughout the collegiate mathematics curriculum, especially in pre-calculus and calculus courses, where students consider topics such as limits, the asymptotic behavior of rational functions, infinite sequences and series, and improper integrals. However, this represents but a small portion of all the situations where the infinite appears. For instance, many of the mathematical structures studied in linear algebra, abstract algebra, real analysis, and topology are infinite sets, existence proofs frequently require the construction of infinite mental procedures, and problem situations involving collections of mathematical objects indexed by an infinite set occur throughout the undergraduate curriculum.

Although the concept of infinity permeates the undergraduate curriculum, students experience little if any formal instruction on the concept prior to a study of cardinality in a “bridge” course or a formal study of Cantorian theory in an upper division set theory course. Our experiences as teachers and researchers convince us that this is insufficient. As one can see from the quotes given in the Overture, students continue to struggle with the concept of infinity, despite their experiences in lower division courses. Two projects ([3] and [11]) show the degree to which students with fairly strong mathematics backgrounds struggle with infinity, even after having completed at least one course where aspects of mathematical infinity were considered in some depth.

In the study reported in [3], students were asked to prove or disprove the statement

$$\bigcup_{k=1}^{\infty} P(\{1, 2, \ldots, k\}) = P(\mathbf{N}),$$

where $P(\mathbf{N})$ indicates the power set of the set of natural numbers. Of the 13 students interviewed, only one student solved the problem correctly, and even this student needed significant prompting before providing the correct solution. In trying to interpret the meaning of the infinite union, all of the students constructed an infinite iterative process that yielded an infinite sequence of power sets of the form $P(\{1, 2, \ldots, k\})$. Subsequently, every one of them seriously entertained the notion that the sequence would “eventually” yield the power set of an infinite set.

In [11], a different set of students were asked to work on the tennis ball problem mentioned in the Overture. Even though every one of the 13 students could articulate the time at which the nth ball would be dropped into Bin A and then transferred to Bin B, only one student concluded that Bin A would be empty after all of the balls had been dropped. Their conflicting thoughts about how to deal effectively with an unending process appeared to be at the root of their difficulties.

In addition to playing a role in understanding infinite processes (such as those arising in considering an infinite union or the tennis ball problem), students’ conceptions of infinity have a bearing upon their ability to solve other types of problems. In order to act on an infinite set, say to compare the cardinality of two infinite sets or to show that the set of all linear combinations of a set of vectors in a vector space over $\mathbf{R}$ is a subspace, one must be able to think of these infinite sets as mathematical objects, or entities, which can be transformed. Helping students formulate object conceptions of mathematical concepts requires the development and implementation of carefully designed instruction. In the sections that follow, we discuss how APOS Theory can provide an explanation of how human beings conceive of the infinite. This is a first step toward the development of pedagogical strategies intended to help students to understand and apply the kinds of transformations required for the successful solution of various problems involving the infinite.

There are two reasons why we believe APOS Theory might be a useful tool in these endeavors. First, APOS Theory has been used over the past 20 years to analyze student thinking about various concepts in the undergraduate mathematics curriculum. The results of these analyses have guided the development of effective pedagogical strategies ([13]). Second, our initial attempts to apply APOS Theory to understand how individuals think about infinity have been encouraging. Based on an analysis of interview data of students attempting to solve the above infinite union problem, the research study [3] describes how students appear to construct their conceptions of infinite iterative processes. We have recently prepared a comprehensive report to discuss how APOS Theory can be used to explain, and in some cases to propose resolutions of, many of the issues and paradoxes of the infinite that have plagued philosophers and historians of mathematics for centuries ([2]).

Inspiration

Before we describe the investigations that have ensued from thinking about problems like those mentioned above, we discuss the mental mechanisms and structures to which we will be referring. This brief explanation of APOS Theory is meant to familiarize the reader with the terminology used in subsequent sections.

The *interiorization* of actions is an "everyday" activity in the mathematics classroom. For example, an algebra student may wish to describe the behavior of a quadratic function over a given interval to see whether it increases for a while and then decreases. The transformation of calculating functional values over the interval is first conceived as an *action*, in that it requires specific instructions, e.g., a formula. Repeating this action and reflecting on the relationship between functional values as x varies over the interval, the student may begin to *interiorize* the action into a mental structure called a *process*. This is a structure which implements the action, not externally, but internally, in the individual's mind. A process enables the individual to imagine the calculation of several values of the function and to think about these calculations all at the same time. Thus, the individual can observe the behavior of the functional values as x varies over the interval without having to evaluate $f(x)$ for explicit values of x. At this point, if the student becomes aware of the process as a totality, realizes that transformations can act on that totality and can actually construct such transformations explicitly or in her or his imagination (e.g., think about horizontal and vertical shifts or compressions and expansions), then we say that the individual has *encapsulated* the process into a cognitive *object*.

There are two aspects of encapsulation that are important to keep in mind. First, according to APOS Theory, an encapsulation occurs *because* the individual desires to perform an action (or process) on a process. This is not possible because a process is dynamic, something that is in progress, and as such is not susceptible to being acted upon. For example, students who have not encapsulated the process of set formation into an object will think that a set such as $\{2, \{5, 6, 8\}\}$ has cardinality 4 (and not 2). This may be because their thinking about 2, 5, 6, 8 does not go farther than the process of inserting these four numbers into a set. Encapsulating the process of forming $\{5, 6, 8\}$ into an object eliminates this difficulty and allows one to perform the desired action (in this example, determining cardinality). Second, it is often important in a mathematical activity to de-encapsulate an object, that is, to go back to the process from which it came.

The encapsulation and de-encapsulation of processes in order to perform actions is a common experience in mathematical thinking. For example, one might wish to add two functions f and g to obtain a new function $f + g$. Thinking about doing this requires that the two original functions and the resulting function are conceived as objects. The actual transformation is imagined by de-encapsulating back to the two underlying processes and coordinating them by thinking about all of the elements x of the domain and all of the individual transformations $f(x)$ and $g(x)$ at one time so as to obtain, by adding, the new process which consists of transforming each x to $f(x) + g(x)$. This new process is then encapsulated to obtain the new function $f + g$.

The mental mechanisms of interiorization and encapsulation allow one to think about what happens after a process is completed. In many cases, the domains and the ranges of functions are infinite sets so these mechanisms allow an individual to think about infinity in these contexts.

While these mental structures describe how an individual constructs 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*. The mental structures of action, process, object, and schema constitute the acronym APOS. In this article, we use these mental structures and the mental mechanisms of interiorization and encapsulation to analyze from a cognitive perspective various issues raised as a result of careful thinking about the infinite. For more information about APOS Theory and a summary of how it has been used in mathematics education research, see [5].

Application

Our discussion of applications of APOS Theory to the issues listed at the beginning of this essay will be divided into four parts. First, we consider some problems that have been around for a long time. These include the classical paradoxes, disputes about the infinitely small and the value of an infinite, repeating decimal. Next, we consider the tennis ball problem as an example of a situation involving infinite iterative processes in which one's conception of the set of natural numbers plays a role. This is followed by a description of a more general consideration of infinite iterative processes based on a recently completed research project. Finally we describe a new research project that begins to look at mental constructions of uncountable sets.

Classical paradoxes, the infinitely small and repeating decimals

Aristotle's resolution of paradoxes such as Achilles and the tortoise consisted of making a distinction between actual and potential infinity and then rejecting the former. While he believed human beings could conceive of potential infinity, he considered the idea of an actual infinity to be beyond the understanding of mortals ([8, pp. 34 - 44]). Although there were dissenters, Aristotle's idea persisted for millenia and it has been expressed quite explicitly in modern times by various writers, including mathematicians such as Poincaré ([10, pp. 46-47]).

In our view, however, the rejection of actual infinity is unnecessary and ignores important mathematical notions such as actually infinite sets, comparison of infinite cardinalities, mathematical induction, etc. In fact, there are explanations of how an individual might think about infinity that incorporate the mind's ability to contemplate actual infinite objects.

Today's formal definition of limit (in terms of epsilons and deltas) provides a satisfactory *mathematical* explanation of how a symbol such as dx can be used in one part of a calculation as if it were a positive quantity and in another part, ignored (or "neglected") as if it were zero. Extensive research shows, however, that the dispute between Newton's "evanescent quantities" and Berkeley's "ghosts of departed quantities" ([1]) is alive and well in the minds of many (most?) of today's calculus students ([4], [12], [15], [16]). Again, we feel that there is an alternative cognitive explanation that could be more helpful to students struggling to resolve the paradoxes of the "infinitely small".

The importance of the pervasive difficulty in dealing with the relation $0.999\ldots = 1$ and the persistent idea that there is an "intermediate state" between a sequence of values and its limit ([4]) lies not only in the specific mathematical errors that students make, but also in what it tells us about how far so many students are from having a useful intuitive understanding of the concept of limit of a sequence. As nearly as we can tell from our review of the literature ([2]), no explanations, other than those based on APOS Theory, have been offered regarding the cognition of $0.999\ldots = 1$, that is, regarding how an individual might think about this relation in ways that fit with the mathematics. Moreover, the only pedagogical strategy available seems to be a reiteration of the mathematical facts.

We will use a single kind of analysis, based on APOS Theory, to propose resolutions of cognitive issues as disparate as the paradox of Achilles and the tortoise, the question of the existence of infinitely small, non-zero quantities, and the relation between a sequence and its limit. We hope the unity of our explanations makes them clearer and more satisfactory than other explanations that have been offered. But more importantly, we feel that these analyses could lead to pedagogical strategies that will be effective in helping students understand the mathematics of the infinite.

Achilles and the tortoise

An individual can think about an infinite iterative process using the mental structure of process as described in APOS Theory. In terms of that theory, performing a small number of iterations constitutes an action. By interiorizing these actions, an individual can use the resulting process structure to imagine repeating the actions indefinitely, or "forever," so to speak. This corresponds to potential infinity. Using the process mental structure, an individual can see the process as a totality, even if it is inconvenient or impossible to think explicitly about each step in the process, and decide to perform actions on the total process. Here, the mental structure of encapsulation comes into play. Encapsulation consists in transforming the process to an object and applying the desired action.

In the case of Achilles and the tortoise, we have two coordinated processes of Achilles repeatedly covering the previous distance traveled by the tortoise while the tortoise continually moves further along. As processes, we can imagine this going on forever and we encapsulate the completed processes in order to perform on them the action of comparing the total distances covered by the two. With a cognitive grasp of the question, we can then 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.

It is important to note that, given an infinite process, the mental mechanisms of interiorization and encapsulation allow one 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 takes care of since one does not have to actually perform all of the steps, whether there be finitely or infinitely many of them. And, in our view, the ability to encapsulate an infinite iterative process requires thinking about actual infinity. Thus, we claim that human beings can conceive of actual infinity.

Infinitesimals

The dispute about the meaning of dx (Newton actually used o) in the expression

$$\frac{f(x + dx) - f(x)}{dx}$$

can be represented by the ideas of Newton and Berkeley. In ([2]), we consider the ideas of these two thinkers in some detail and try to show how our interpretations are based on their actual writings. Here, we briefly summarize our interpretations and explain, in APOS terms, how an individual can think about infinitesimals in calculus.

In terms of an APOS analysis, it appears that the crux of the issue was that when writing the difference quotient, Newton intended dx to represent a process of approaching 0. That is, the symbol dx in the difference quotient stands for the process of replacing the symbol dx by smaller and smaller positive numbers. For each one of these positive numbers, one can compute the difference quotient. Hence, that expression represents the process of obtaining values by replacing dx with smaller and smaller positive numbers and making the calculations. On the other hand, when Newton wrote

> Quantities, and the ratios of quantities, which in any finite time converge continually to equality, and before the end of that time approach nearer the one to the other than by any given difference, become ultimately equal. ([9, p. 29])

he was seeing those processes as totalities and encapsulating them in order to obtain the ultimate value of dx (which is 0) and of the difference quotient (its limit, in contemporary parlance). In the examples that he considered, Newton was generally able to obtain the latter value by simplifying the difference quotient expression and then replacing dx by 0, that is, “neglecting” it.

Thus, in our interpretation, when dx represents a process, it is a positive number, so the difference quotient makes sense mathematically. But then, with the encapsulation, dx represents an object to which Newton referred as its ultimate value, which is zero. The ultimate value of the difference quotient, again no longer as a process but representing an object, is its limit. This distinction between process and object is our resolution of the “contradiction” of dx being sometimes positive and sometimes 0.

Our interpretation of Newton’s thinking has an additional attribute that we think helps explain why many people had such difficulty in understanding what he was saying. In a finite process, there is always an object produced in the last step of the process. Even though there is no last step in an infinite process, such a process can still result in an object. This requires, however, a much more powerful mental mechanism than imagining a last step. According to APOS Theory that mechanism is encapsulation, and the resulting object is what is called in [3] the *transcendent object* of the process. Our interpretation of Newton’s thinking is that he understood the distinction between process and object in this context and realized that a more powerful mental step was required in order to go from an infinite process to an ultimate value. We conjecture that he had encapsulated the limiting process but had no mathematical tools (such as the formal concept of limit) to express precisely his object conception.

Drawing on an analogy with determining the velocity of a body "at the very instant it arrives", Newton emphasized what happens "ultimately," which we interpret as an attempt to apply an action of evaluation to a completed process, leading to an encapsulation. He clearly distinguished the objects produced by the process from the transcendent object produced by encapsulating the process.

Newton's critics, however, insisted that dx and the difference quotient itself must always be viewed as static objects. Thus, when Berkeley insisted that Newton's evanescent increments were "neither finite quantities, nor quantities infinitely small, nor yet nothing" [1], our view is that he was not distinguishing between an object produced by a process and an object that is brought into being by encapsulating the process as a result of applying the action "What is the ultimate value of the process?"

0.999... and 1

Maybe the students are right, maybe $0.999\ldots$ is *not* the same as 1, at least not cognitively. APOS Theory can offer an explanation of such thinking. Mathematicians consider $0.999\ldots$ to stand for the limit of an infinite sequence. However, the ability to think of this expression in that way requires certain mental constructions that some students may not yet have made. For such individuals, the symbol $0.999\ldots$ appears to represent a process (that is the only possible explanation of the $\ldots$ or phrases such as "and so on"). It is an infinite process, and so there is no object produced by a last step. The symbol 1, however, refers to an object. Since a process is something different from an object, it makes sense to say that the process $0.999\ldots$ cannot be the same as the object 1. What makes this particularly difficult is that the number 1 is not an object produced by any step in the process but is the result of encapsulating it, and so transcends the process.

In fact, we conjecture that certain mental constructions are necessary before it is possible to even think about the mathematical solution, much less understand it. Specifically, to understand that $0.999\ldots = 1$, the individual must first realize that $0.999\ldots$ is an infinite process and, as such, does not produce an object (a numerical value in this case) directly. Rather, the process must be encapsulated to an object in order to find this value. Once this encapsulation has been made, one can then do some mathematics to determine the value. For example, one might argue that there are two numbers L and 1, L being the object produced by the encapsulation. Since L is understood to be a value that is determined after the process is finished, there might be less of a tendency to try to force it to come from the process and be very close to 1 or "the number just before 1". One might then calculate that $|L-1|$ is smaller than any positive real number, so it must be 0, and so $L=1$.

We have no data to support the idea that students are aware of the subtle distinction between infinite processes and their transcendent objects. We plan to design experiments to investigate this question as well as the conjectures above. This might lead to the design of new pedagogy, focused on the process/object distinction, that is intended to help students overcome difficulties in understanding infinite decimals, such as those reported in [4] and [12].

Construction of the natural numbers and student thinking on the tennis ball problem

College students' thinking on the tennis ball problem is the focus of a current study ([11]). During interviews, students were asked to imagine three bins of unlimited capacity: a holding bin, where balls numbered $k = 1, 2, 3 \dots$ would originate; bin A, where each ball would be held temporarily; and bin B, the final destination for each ball. Their task was to determine the contents of bins A and B at noon, if at $\frac{1}{2^k}$ seconds before noon, balls numbered $2k - 1$ and $2k$ are placed in bin A, while the ball numbered k is moved from bin A into bin B.

The data in [11] is part of an analysis of students' thinking on this problem. The aim is to investigate how they think about the set of natural numbers and to test the generalizability of the description of the mental construction of infinite iterative processes developed in [3]. This latter study, which we describe later in this article, looks at a more complex mathematical situation.

A theoretical analysis of the tennis ball problem

There are two competing conceptions of the problem which make it paradoxical. On one hand, the number of balls in bins A and B increases by one at each step, suggesting that both bins have infinitely many balls at noon. However, the kth ball is moved from bin A to bin B at the kth step, from which it follows that bin A will be empty at noon.

This paradox is resolved by comparing the contents of the bins. The mental mechanisms of interiorization and encapsulation allow one to think about what happens after the process is completed. In the tennis ball problem, there are two coordinated processes, one in which Bin A receives the next two balls and, in the other, the lowest numbered ball in Bin A drops into bin B. One imagines these coordinated processes continuing forever and then encapsulates the completed processes in order to perform the action of comparing the number of balls in Bin A and Bin B. At this point, with a cognitive understanding of the question of the contents of Bin A after the process is completed, one can see that at noon Bin A will be empty. Even though there is no last step in the infinite process of filling the bins, one can use encapsulation to imagine a resulting transcendent object, and then determine mathematically that it is the empty set, e.g. by checking that Bin A and the empty set contain exactly the same elements.

Preliminary results of student interviews

To unravel the basic notions undergraduates hold concerning infinite processes, students (with majors in mathematics, mathematics education or computer science) were asked to solve the tennis ball problem. A preliminary analysis of the data suggests that students who had difficulty solving the problem did not see the underlying infinite iterative processes as completed totalities, which is a necessary precursor to encapsulation. Thus, they did not see the "ultimate" contents of each bin as an object, and so the question of how many balls each contains was meaningless to them.

Nearly every student could articulate that the kth ball would be moved from bin A to bin B at the kth step. However, only one student argued that bin A is empty. This student seemed to grasp that the movement of the kth ball not only describes what happens to a single, randomly selected ball, but to every ball in the holding bin. Several of the students who had difficulty with the problem believed bin A would never be empty because the iterative process would continue beyond the kth step. Although they understood that every ball would "eventually move to B," they believed that "there will always be more to come."

Other students argued on the basis of cardinality. At each succeeding step, the number of balls in each bin increases by one. At the kth step, balls numbered 1 through k are in B, and balls numbered $k+1$ through $2k$ are in A. Some of the students generalized the finite case to assert that each bin would contain "half of the balls," the "upper half" in A and the "lower half" in B. Others substituted ∞ for k to conclude that B would contain 1 through ∞ and A would contain $\infty + 1$ to 2∞. When prompted, these students generally acknowledged, as did those who did not focus on cardinality, that each ball would "eventually" be moved to bin B. However, they did not find this information useful. In their view, the problem could not be solved. Because the procedure would continue indefinitely, one could never identify a particular number (other than ∞) that could be substituted for k to identify the precise "halfway" point that would denote which numbered balls would be in A and which would be in B at noon.

A preliminary analysis of the data suggests that the students who had difficulty may not have made certain mental constructions. Those who simply substituted ∞ for k and used this to conclude that bin B contains 1 through ∞ and A contains $\infty + 1$ through 2∞ may not have constructed a useful infinite iterative process from their conceptions of finite iterative processes. They generalized from the finite case to the infinite in a manner that did not account for the way in which the infinite case transcends the finite.[1] Those students who did not substitute ∞ for k, but asserted that the "ones that are higher" remain in bin A because "there's an infinite number of natural numbers," likely did not see the process as being complete. Because the process "does not stop," or "noon can never be reached," they could not imagine that all the steps of the process could be carried out and finished.

The student who correctly proved the result appeared to make this latter construction. In his proof, he argued that if the nth ball were contained in A at noon, a contradiction would result, because the ball would appear in bin B at any time within $1/2^n$ seconds before noon. In making this argument, the student realized that all of the balls would be in bin B at noon. This required him to see the process as a totality or a single operation. The student in question gave evidence of having made this construction when he said: "So any time I choose an n, like say I choose an n out of the holding thing. Well after the nth time, it's going to be in B, and so that for me was like saying, okay if I have an n, n has to be in B, so that means all the holding bin will end up in B." Because he could see the process as a completed totality, he was able to encapsulate it and argue that bin A is empty. In making this construction, he understood that the situation at noon transcends the process, in the sense that it differs from and is not produced by any step of the process.

These preliminary findings suggest that a correct solution is dependent upon the student's ability to see the underlying infinite iterative process as a completed totality. Without making this mental

[1] One possible explanation is that some students inappropriately felt that this process was "continuous" in the sense that the number of balls at the end would be the limit, as k goes to infinity, of the number of balls at the kth step. Such overgeneralization is common. For example, some calculus students will overgeneralize the "zero product property" and reason that $\lim_{x\to 0} x\cot(x) = 0$, even though $\lim_{x\to 0}\cot(x)$ is not finite.

construction, the student finds the problem difficult or impossible to solve because the process "does not stop," from which it follows that "there are always more balls to come."

Conceptions of infinite iterative processes

The ongoing study concerning the natural numbers and the tennis ball problem is closely related to the research study reported in [3]. In this investigation, the authors interviewed students solving a particular elementary set theory problem and developed a description of the mental constructions an individual might make and use to understand infinite iterative processes. As stated earlier, students were asked to prove or disprove the following equality:

$$\bigcup_{k=1}^{\infty} P(\{1, 2, \ldots, k\}) = P(\mathbf{N}).$$

Although one can resolve the question by noting that the set on the left side is countable and the set on the right is not, in the context of the course, the authors expected students to compare the two sets on the basis of set inclusion. In particular, they expected students to note that the union on the left side contains only finite sets as elements, while $P(\mathbf{N})$ contains infinite sets as elements.

While mathematicians may see both sides of this proposed equality as static objects, the students interviewed saw these as processes they had constructed or needed to construct. Thus, the authors decided to look carefully at the actions and processes that individuals might construct in order to understand the formal notation represented in this problem. The central role of one's conceptions of infinite iterative processes was brought to light by this analysis. We give a short description of the mental construction of infinite iterative processes proposed by the authors and then make some brief observations.

The construction of a mathematically useful conception of infinite iterative processes appears to be based on one's process conception of finite iteration. One must be able to apply the relevant finite process to an initial object and understand generally how an object is produced via the process from the preceding object or objects. A process conception of infinite iteration develops as the individual becomes able to coordinate multiple instantiations of this finite process. A successful coordination leads to the individual becoming able to conceive of this infinite process as being complete, even though there is no final step of the process and no last object. Once the process can be imagined as being complete, the individual may reflect upon it, and begin to see it as a totality, in the sense of seeing it as a single operation that can be carried out and finished. Depending on the situation, the individual might attempt to construct an action of evaluation on the process, typically with the goal of determining the state at infinity for the process. A successful application of an action of evaluation happens in tandem with the encapsulation of the process into an object, called its transcendent object. This object is understood to be related to, but beyond, the objects produced by the process, in the sense that it cannot be produced by applying the iterative process to any of the previously produced objects. The objects produced by the process, followed by the transcendent object, are then conceived of as forming an extended sequence (i.e., an ordered set indexed by $\mathbf{N} \cup \{\infty\}$).

To consider the infinite union above, many students began by constructing a process of the following form, noting that successive power sets are nested:

$$P(\{1\}) = P(\{1\})$$
$$P(\{1\}) \cup P(\{1,2\}) = P(\{1,2\})$$
$$\vdots$$
$$P(\{1\}) \cup P(\{1,2\}) \cup \ldots \cup P(\{1,2,\ldots,k\}) = P(\{1,2,\ldots,k\}).$$

Students who were successful realized that they could continue this process and see it as being complete, even though there is no final step to the process and no final object produced. Many of the students' difficulties were associated with this issue, in that it was hard for them to consider the infinite process without imagining that the set $P(\mathbf{N})$ or an entity that they referred to as $P(\{1,2,3,\ldots,\infty\})$ was produced by the process. Only a few students could see the process as complete, and in order to be successful with the problem, they also had to be able to see the completed process as a totality. This means that they had to understand that every set constructed within the process contains only finite sets. With that understanding, they could see that the transcendent object for this process is the set of all finite subsets of $\mathbf{N}$, an object that is not constructed anywhere in the process, but rather is constructed only as the process is encapsulated.

The above analysis focused on the infinite union that appears in the problem. Students' struggles in determining what that union is equal to tell us a great deal about their construction of this set. Because this set is *not* equal to the uncountable set $P(\mathbf{N})$, however, this investigation tells us little about the mental construction of $P(\mathbf{N})$. This is a completely different matter to which we now turn.

Conceptions of $P(\mathbf{N})$

While one may come to understand $P(\mathbf{N})$ through its formal definition as the set of all subsets of $\mathbf{N}$, we would argue that the definition alone is not sufficient, at least not for some undergraduate mathematics students. To apply actions or processes to this set in certain problem situations, APOS Theory posits that an individual needs to be able to access a rich process conception through de-encapsulation of an object conception. The set $P(\mathbf{N})$ cannot, of course, be constructed by encapsulating the infinite iterative process of taking the union of the power sets of initial segments of $\mathbf{N}$. Nevertheless, we may ask if it is possible to conceive of this set as arising from *any* infinite iterative process, given that the set itself is uncountable. Our investigation of mental constructions of uncountable sets begins with the question: How do experts get a rich conception of $P(\mathbf{N})$? Is it simply through mathematical formalism, or are there underlying mental constructions that can be revealed through careful research? These questions, as well as our interest in studying conceptions of uncountable infinity, led us to investigate how mathematicians come to understand $P(\mathbf{N})$ and whether their construction of this understanding admits an APOS analysis. This is a project that we are currently in the midst of conducting.

Is the mental construction of $P(\mathbf{N})$ made by a sophisticated mathematical thinker, for example a research mathematician, derived mainly from the formal definition? Certainly this definition states completely what the set is, and one could imagine various infinite processes for building a given set in $P(\mathbf{N})$. But the question still remains as to how one would think of the set "all at once," as a cognitive object. According to APOS Theory, when one uses this object in a problem situation, a de-encapsulation to a process is needed. But, what is the process that was encapsulated to give the set in the first place? Could this process be an infinite iterative process, even though the set itself is uncountable? In our current research, we have in fact developed a countable, iterative process along with multiple encapsulations which we believe yields the uncountable set $P(\mathbf{N})$. This process seems to be related to the binary tree construction of $P(\mathbf{N})$, where $P(\mathbf{N})$ is the set of all branches of an infinite binary tree.[2] Our concern, however, is with the cognitive meaning of "the set of all branches" and how an individual might construct this set in her or his mind.

As of this writing, we are conducting interviews with mathematicians who do have rich conceptions of uncountable sets to compare the mental mechanisms they appear to use in developing these conceptions with the process we have constructed and its encapsulation. Preliminary results are encouraging and we hope to produce reports on this work in the near future.

A particularly interesting aspect of this current work has to do with our suggestion that a countable process can yield an uncountable set. Mathematically, this is not possible, but it may be that in our minds we can make such a leap. We propose that what makes the difference cognitively is encapsulation, which is not a mathematical tool but a mental mechanism. We believe that encapsulation may just be a sufficiently powerful mechanism to allow the mind to make such a leap from the countable to the uncountable. In future reports, we hope to provide data and analyses that tend to support (or, as the case may be, not support) this conjecture.

Finale

Some of the explanations we have given in this article are at variance with those of other commentators on infinity in mathematics, both past and present. There is, of course, no issue here of determining who is correct (whatever that may mean). Rather, we hope that our explanations exhibit a coherence, unity and simplicity that may render them worth thinking about when trying to understand how human beings can and do think about infinity.

Perhaps more important, certainly from a practical point of view, is the hope that our explanations of student difficulties with infinity will point to pedagogical strategies that can lead to improvement in learning. The reason for our optimism is that explanations of other mathematical concepts using these mechanisms, and the totality of APOS Theory of which they are a part, have led to effective pedagogy that has been reported in the literature. Whether a similar outcome will occur for infinity, only time and future research and development can tell.

References

[1] C. B. Boyer, The history of the calculus, *The Two-Year College Mathematics Journal* **1** (1970),

[2] This construction is familiar to set theorists. See, for example, the proof that the Cantor set is homeomorphic to $2^{\aleph_0}$ in [14], the identification of the Baire space $\mathbf{N}^{\mathbf{N}}$ with the irrationals in [6], etc.

60-86.

[2] A. Brown, E. Dubinsky, M. A. McDonald, and K. Weller, An APOS perspective on historical paradoxes and dichotomies of the infinite (under review).

[3] A. Brown, M. A. McDonald, and K. Weller, Students' conceptions of infinite iterative processes (under review).

[4] B. Cornu, Limits, *Advanced mathematical thinking* (D. O. Tall, ed.), Kluwer, Dordrecht, pp. 153-166.

[5] E. Dubinsky and M. A. McDonald, APOS: A constructivist theory of learning in undergraduate mathematics education research. *The Teaching and Learning of Mathematics at University Level: An ICMI Study* (D. Holton et al., eds.), Kluwer, Dordrecht, pp. 273-280.

[6] K. Hrbacek and T. Jech, *Introduction to Set Theory*, M. Dekker, New York, 1978.

[7] A. W. Moore, A brief history of infinity, *Scientific American* **272** (1995), 112-116.

[8] A. W. Moore, *The Infinite* (2nd edition), Routledge, London and New York, 1999.

[9] I. Newton, *The Principia: Vol. 1. The Motion of Bodies* (Translated into English by A. Motte in 1729. Translations revised and supplied with an historical and explanatory index by F. Cajori), University of California Press, Berkeley, 1934.

[10] H. Poincaré, *Mathematics and Science: Last Essays* (J. W. Bolduc, trans.), Dover, New York, 1963.

[11] C. Stenger, D. Vidakovic, and K. Weller, Students' interview responses to a problem involving an infinite iterative process (data for a study in progress), 2003.

[12] D. Tall and R. L. E. Schwarzenberger, Conflicts in the learning of real numbers and limits, *Mathematics Teaching* **82** (1978), 44-49.

[13] K. Weller, J. Clark, E. Dubinsky, S. Loch, M. A. McDonald, and R. Merkovsky, Performance and Attitudes in Courses Based on APOS Theory, *Research in Collegiate Mathematics V*, CBMS Issues in Mathematics Education, Volume 12, A. Selden et al., editors. (2003) 97-131.

[14] S. Willard, *General Topology*, Dover, Mineola, NY, 1970.

[15] S. R. Williams, Models of limit held by college calculus students, *Journal for Research in Mathematics Education* **22** (1991), pp. 219-236.

[16] S. R. Williams, Predications of the limit concept: An application of repertory grids, *Journal for Research in Mathematics Education* **32** (2001).

Acknowledgement. The authors would like to thank Judy Roitman for many helpful comments about the work described in this article and about the article itself.