Meaning and Formalism in Mathematics

Ed Dubinsky

Georgia State University

The purpose of this essay is to explore the possible sources of mathematical concepts in the minds of people who understand them and those who are trying to learn them. Thus I am not concerned with *mathematical* sources but rather in the *psychological* question of how an individual can develop her or his understanding of mathematical ideas. The obvious relation of such an inquiry to teaching and learning is an important issue that, although not a primary focus of this discussion, will not be ignored.

Quite a bit has been written on the psychological sources of mathematical concepts and I will comment on some of this literature. My focus, however, will be on two themes: the relations between meaning and formalism in mathematics and the transition between elementary and advanced mathematical concepts. Considering the role of formalism in mathematics can contribute to our understanding of the gap between school mathematics and collegiate mathematics. The role of formalism may be quite different in these two domains, and I suggest that the nature of this role could go a long way towards characterizing and explaining the gap between them.

It will surprise no one who has thought about these things that the relation between meaning and formalism is far from simple. A part of knowing something is knowing what it is not. Therefore, we need to think about meaning in mathematics in the absence of formalism as well as any role that formalism can play with a (relative) absence of meaning. Of course, the heart of the matter is what is going on when both are present. It seems to me that in this case, we must ask the following questions: what formalism can arise from meaning and how does that come about; and what meaning could come from formalism and how does that come about?

I would like to begin with a discussion of some of the literature about the psychological sources of mathematical concepts. It can be seen how at least part of the difference between elementary and advanced mathematical concepts lies in the differences between various sources that have been proposed. I will then discuss meaning and formalism, beginning with a brief, relatively unsophisticated explanation of what I am talking about when I use these two terms. I will proceed then to a systematic discussion of possible relationships between these two concepts. Finally, I will close things up by exploring what implications my comments might have for research in mathematics education and for teaching mathematics.

1 The sources of mathematical ideas

There seems to be an almost universal agreement that mathematical concepts arise, initially, out of human experience. As Piaget points out, for example, the idea that the cardinality of a finite set is independent of the order of its elements comes for young children from counting a set of similar objects (e.g., stones, marbles, coins, cousins, etc.), rearranging, counting, and so on, always getting the same count (Beth &

Piaget, 1966, p. 231.) Other authors agree with this initial source (e.g., Davis & Maher, 1997) for very elementary concepts. But what, then, are the sources of more sophisticated ideas that are not so directly connected with normal human experience. Consider, for example, the convergence of a sequence (f_n) of (say, real-valued) functions, defined on a common domain D, to a function f. There are at least two relevant mathematical concepts: pointwise convergence, defined as,

$$\forall x \in D,\ \forall \epsilon > 0,\ \exists \text{ a positive integer } N \ni \text{ if } n > N, \text{ then } |f(x) - f_n(x)| < \epsilon,$$

and uniform convergence defined as

$$\forall \epsilon > 0,\ \exists \text{ a positive integer } N \ni \forall x \in D, \text{ if } n > N, \text{ then } |f(x) - f_n(x)| < \epsilon.$$

Clearly one can point out that the idea of these two definitions arose out of considerations of continuity of the limit of a sequence of continuous functions — which Cauchy apparently got wrong (Lakatos, 1976). I am not asking about source in the sense of how this idea entered the body of mathematical knowledge, what we might call the *mathematical source*. I am concerned with the *psychological source*[1]. If you ask how a person comes to understand the concept of convergence of functions, then there is a big jump from normal human experience to these two apparently similar, but very different definitions. How does a person make, in her or his mind, such a jump?

I think this relates to the nature of "advanced mathematical thinking" as opposed to "elementary mathematical thinking" (the careful reader will note that the ambiguity of meaning of these two triplets is resolved by imagining parentheses around the first two words and not the last two.) It might even be reasonable to define "elementary" and "advanced" in this context. The former might be taken to refer to mathematics that arises out of reflection on normal human experience, whereas some sort of cognitive jump might be required in order to make the transition to the advanced level. I would go so far as to suggest that school mathematics (levels K-12 in the US) should focus on what I am calling here elementary mathematics, at most working towards a transition to what I am calling advanced mathematics which should be the main concern of university studies.

Let me turn now to a consideration of what some other people have said about the transition.

1.1 Platonism vs. constructivism

As Hersh, 1986 suggests, investigating the source of mathematical ideas is one reason why we should pay serious attention to the philosophy of mathematics. For example, if we accepted a Platonist philosophy then we would consider that mathematical ideas, whether elementary or advanced, have their own independent existence and it is a question of an individual absorbing them into her or his mind. For some people,

[1] I am aware that the mathematical and psychological sources can be the same for some concepts, but that they can be very different for others. See Piaget & Garcia, 1989 for a deep investigation of this question.

this means that the psychological transition can be made by listening to or reading carefully crafted explanations, or looking at appropriate pictures. Others, who reject the Platonist view, believe that mathematical ideas exist only in the minds of individuals, perhaps as shared, social knowledge and that a person has to construct these ideas (or at least an understanding of them) for her or himself.

Let me digress briefly and consider this notion of constructivism and the tension between it and Platonism. The constructivist idea does not mean that a person is free to make up her or his ideas about mathematical concepts in any manner and with any results whatsoever. Not only must an individual's understanding of a particular concept fit with that of the community of mathematicians (von Glasersfeld, 1987), but it must also be compatible with mathematical phenomena, experiences and methods such as proof and counterexample.

Nor does a constructivist position entail the requirement that an individual discover all mathematical knowledge for her or himself, without relation to others. From Piaget to Vygotsky, constructivist epistemologists have insisted that the mental constructions that a person uses to understand a mathematical concept are made in a social context and with considerable intervention from teachers and fellow students. It is true that constructivists generally consider the opportunity to discover a mathematical idea to be a powerful stimulus for making mental constructions, but the overall set of mechanisms for doing this are far more complex (see Piaget & Garcia, 1989, Asiala et al., 1996 for further discussion of these mechanisms) then discovery alone.

Platonism and constructivism are two competing philosophical positions and one may ask which is "correct". Piaget (Beth & Piaget, 1966) suggests, however, that this is an unreasonable question. He argues that if we observe a person moving from not understanding a concept to understanding it, we might be able to see what mental mechanisms are employed, but we cannot tell if these mechanisms are being used to construct the concept or to gain access to it. For example, even when a person is learning an idea by listening to a lecture about it (which certainly happens on occasion), we cannot tell if the speaker is in some sense taking this idea from an ideal Platonic world and transferring it through her or his mind and mouth to the ears and mind of the listeners, or if the speaker is creating a situation to which some listeners may react by making appropriate mental constructions. Thus Piaget concludes that it is just as unreasonable to deny Platonism as it is to assert it.

My personal conclusion from this is reminiscent of Hersh's dichotomy (Hersh, 1986) for Platonism vs. formalism in which he asserts that "the typical working mathematician is a Platonist on weekdays and a formalist on Sundays." Without equating formalism with constructivism, I would suggest the parallel dichotomy that when you are doing mathematical research and trying to obtain new ideas, or solutions to problems, you almost have to think (Platonically) that you are looking for something that is "there". By the same token, if you are trying to understand an idea that is new for you, or if you are working with students to help them learn mathematics, experience strongly suggests to me that nothing is learned by individuals unless they, ultimately, construct the ideas in their minds and "make them their own".

In a sense, then, a resolution of the Platonism/Constructivism dichotomy amounts to a personal choice, perhaps depending on the situation. This does not help very much. If we decide to adopt a Platonist point of view, then we still must determine the mechanisms by which an individual gains access to the mathematical ideas that exist in some ideal world, and if we are thinking in constructivist terms, we still must work out the mechanisms by which the individual constructs those ideas. As Piaget & Garcia, 1989 emphasize, it is the mechanisms that are important and must be studied, regardless of our choice of philosophical position.

1.2 From human activity to mathematical concepts

As I have already noted, there seems to be a considerably widespread agreement that mathematical ideas begin with human activity and move from there to abstract concepts. Thus, the mechanisms we are searching for are those by which this movement takes place. Here there is a variety of suggestions that have been made. I will describe some of them from the literature and relate what has been suggested to the example of convergence of a sequence of functions. In particular, I am interested in how effective the suggested mechanisms might be for moving from a first intuition of convergence to an understanding that there is an issue connected to the dependence of convergence on the domain point and that two important and different ways of resolving this issue are pointwise and uniform convergence.

1.2.1 Mental representations.

According to Davis and Maher, 1997, the key mechanism for an individual to obtain new mathematical meaning is for her or him to construct mental representations of direct experiences. For example, they report on an experiment by Machtinger in teaching kindergarten children to understand the concept of even and odd numbers. The teacher had the children walk out of the classroom and along corridors in pairs and reflect on the phenomenon that sometimes there was a child left over and sometimes not. The pupils were assumed to have represented those experiences in their minds and Machtinger was able to use this approach to get the children to conjecture and (it is claimed) even prove properties about such matters as the parity of the sum of two even numbers or of an even plus and odd number.

The mechanism for moving to more sophisticated mathematical concepts in this point of view is to develop more sophisticated human experiences that can be represented. Included among these experiences are notational systems used as representations of mathematical operations.

Although one might imagine such experiences and notations appropriate for the concept of convergence of a sequence of numbers, it is hard to see how to do this for pointwise vs. uniform convergence. It has been suggested that approximation, for example of a continuous function by polynomials, might be one approach, but I am unaware of any attempt to investigate how this might be done and what its effects could be.

Davis and Maher seem to focus on elementary mathematics at the pre-college level. For them, the passage to more sophisticated mathematical ideas together with formalism is not an important issue and is relegated to "those who continue their studies sufficiently far".

1.2.2 Deductive reasoning.

MacLane (1981) may agree with Davis and Maher for the most elementary mathematical concepts, but he focuses on more advanced mathematical topics and introduces formalism as a mechanism. According to MacLane, the mathematician extracts from human activities certain notions essential to them, formalizes these notions and embeds them in an axiomatic system, which is then studied by deductive methods including rigor. The formalization and deduction can reveal deep and non obvious properties of the original activities. He applies this idea to argue that human activities such as measuring, shaping, estimating and proving are, respectively, the sources of meaning for mathematical topics such as analysis, topology, probability, and logic.

One consequence of this point of view is that rigor is not just a formal exercise that mathematicians like to insist on, but is an essential aspect of *understanding* mathematics. Another component is the level of sophistication for human activities required by this point of view. MacLane makes a short list of these activities which he feels are sufficient to lead to the full range of mathematical concepts. He includes in this list relatively simple activities such as counting, measuring, shaping, moving and grouping. But he also includes forming as in architecture, estimating, calculating, proving and puzzling.

Thus MacLane might suggest that a student trying to understand convergence of a sequence of functions would engage in human activities such as estimating, calculating and proving, in order to make sense of the convergence, for example, of the sequence (f_n) of functions defined on $[0, b]$ where $f_n(x) = x^n$ and b might be taken variously as 0.5 or 1. In "puzzling" about the difference in the two limit functions (e.g., one is continuous and one is not) the student might realize that the "rate of convergence" depends in some sense on the value of x and this could lead to formalizations such as the two definitions for pointwise and uniform convergence given above. At this point deduction could produce a full mathematical treatment of these concepts.

MacLane focuses on more advanced mathematical concepts than do Davis and Maher. Thus, if Davis and Maher are talking mainly about school mathematics, MacLane appears to be concerned primarily with university mathematics.

Of course, all of this leaves open the question of how an individual moves from the one to the other. How can he or she develop the ability to engage in the more sophisticated human activities listed by MacLane? This is important because, as every teacher knows, many students are not able, at least initially, to engage in all of MacLane's activities. What are the mechanisms by which these abilities develop and why does it go further for some students than for others?

Some of the suggestions which follow attempt to describe such mechanisms.

1.2.3 Grounding and linking metaphors.

Lakoff and Núñez, 1997 present an entirely different point of view in which they suggest that all mathematical concepts can be formed through a combination of two kinds of metaphors: grounding and linking. The general notion of mental representation of human activity considered by Davis and Maher is replaced by that of grounding metaphor. The mechanisms of MacLane's formalism for thinking about advanced concepts is, in some cases, also replaced by these metaphors. In other situations, when the advance involves relating one branch of mathematics to another, use is made of what are called linking metaphors.

The claim that all mathematics can be obtained by their metaphors is illustrated by Lakoff and Núñez with a vast array of examples. In their view, "grounding metaphors allow us to ground our understanding of mathematics in familiar domains of experience." Thus, the human activity of grouping objects into conceptual containers becomes a metaphor for sets as in "A Set Is A Container", for elements of sets as in "An Object In A Container" and for subsets as in "A Container Within A Container". An important conceptual advance that allows sets to be considered as elements of other sets is the metaphor that "Sets Are Objects".

Metaphors such as "Numbers are Points on a Line", "Zero is the Origin", "Quantities are Distances", "Greater Is Above" are taken to map "truths of Euclidean Geometry onto Arithmetic" and therefore serve to link the domains of Geometry and Arithmetic.

More advanced mathematical concepts are included in this metaphor analysis as well. For example the limit of a function at a point is expressed metaphorically as, "Approaching A Limit Is Preservation Of Closeness Near a Point." Essentially the same metaphor is used to express continuity of a function at a point and continuity throughout an interval. Because of the claim of completeness, one assumes that similar metaphors can be used for (pointwise and uniform) convergence of sequence of functions.

Although they give them a far less central role than does MacLane, Lakoff and Núñez also include formalizations of metaphors, such as the ϵ-δ definition of limit. As opposed to MacLane, they deny that such a definition (or, indeed, any formalization) adds rigor to the arithmetization of calculus. They consider it rather to be no more than a precise characterization of certain components of the metaphorical definition of this concept.

Appealing to the way mathematicians talk about various ideas, Lakoff and Núñez argue that their metaphors form the source of mathematical ideas. Hence, their answer to our question of moving from elementary to advanced mathematical concepts, from direct physical experience to more abstract notions, is to express the latter as metaphors and to think in such terms. To the extent that they are correct, it would provide a pedagogical strategy for helping students understand these ideas — explain the metaphors.

In this point of view there is really no jump from elementary to advanced mathematics. Since all of the metaphors are taken from human experience, the metaphors needed for university mathematics are not different in kind from those required for school mathematics.

1.2.4 Natural language.

Sfard, 1997 is even more explicit about the role (and power) of metaphor to "create for us the world we live in—including the most remote and esoteric regions of abstract mathematics." Thus for example, "the notion of rational number results from an interaction between several 'concrete' metaphors such as *fraction as portitioning*, *fraction as piece*, and *fraction as number*.

This is not very different from the view of Lakoff and Núñez, but what is new is the close connection Sfard makes between metaphor and natural language as mechanisms for creating meaning. She makes the overall assertion that "the attention to metaphor is an attention to language" and that "It is the way we speak, the way we transplant linguistic expressions from one context to another, which shapes our way of looking at the world." In the case of fraction, for example, "Incorporation of fractions into the discourse on numbers... helped to determine the way they were subsequently used and operated." On the specific question of how one can come to understand a process as an object, for example in going from a function conception of an input/output machine to functions as objects that can be acted upon, she claims that "Introduction of nouns into those places where till now people have only been talking about processes refocuses the discourse." In fact, she writes, "If we write a new name or a new symbol in the slot reserved for objects—the new signifier will eventually bring about an emergence of a new mathematical object."

It is perhaps not an accident that Sfard, unlike Lakoff and Núñez, restricts her examples to relatively elementary mathematical concepts, not very far from notions arising directly from physical experience, such as dividing things up, or taking a part of something. Although I would agree with Sfard that constructing mental processes and objects is a key to developing an understanding of mathematical concepts, I am not convinced that talking about them, even in very sophisticated terms, is the essential step in actually making these constructions, at least not at the university level. As Hersh, 1986 suggests, we would not expect someone to develop a feel for music by reading the notes. In any case, I certainly don't think Sfard's approach works very well with a concept such as convergence of a sequence of functions. Indeed, it is not hard to talk about the distinction between pointwise and uniform convergence and even to express this point metaphorically. Why then, do so many people, even some of our brightest students, have so much difficulty with it? I think that something more is needed.

1.2.5 Mediation by a computer

It could be that representing a situation on a computer and using software that supports sophisticated operations helps students expand their repertoire of the human activities as called for by Davis and Maher and required by MacLane. Moreover, the syntax of a programming language might mediate the students' transition from mental representation to axiomatization and formalization that is essential to MacLane's point of view.

According to Noss, 1997, mathematical meaning can come from an individual's awareness that a particular expression can be recognized by a computer. Thus, entering formal statements on the computer, testing and running them induces the student to construct meaning for these statements.

Noss gives examples in which students use the software *Cabri Geomètre* to investigate the composition of two reflections of a geometric shape and a Logo Microworld to explore the concept of ratio in mixing (paint) and scaling (rug sizes) problems.

To do this for our example of convergence of a sequence of functions, one would have to work out ways of representing such a sequence on a computer. One might, for example, use a programming language such as ISETL which supports representing a function with a procedure that depends on a parameter. For example, one might write,

```
F := func(n);
        return func(x);
                    return x**n;
              end;
     end;
```

Once this has been entered and run, commands such as

```
f3 := F(3);
plot(f3);
```

can be used to construct the third function in the sequence, name, and then graph it. One might do something like the following to graph the first four of the functions in the sequence on the domain $[0, 2]$:

```
plot([F(1), F(2), F(3), F(4)], 0, 2);
```

New experiences for students can arise from working with a (finite) sequence of these functions such as,

```
[F(n) :  n in [1..100]];
```

together with the sequence of numbers,

```
[F(n)(a) :  n in [1..100]];
```

where the code can be edited to use different values of a and reflect on the various numerical sequences so obtained and how they seem to be converging.

Considering, in this approach, the transition from school mathematics to university mathematics, there seems to be a continuity rather than a jump. At both levels, computer tasks are used and the only thing that changes are the particular programs the students are asked to write. A closer look at such programs, however, reveals a more abrupt change. At the school level, for example, it would be reasonable to ask students to write programs which accept numerical parameters and produce numbers or geometrical figures. It is perhaps only at the university level that one can expect the majority of students to be comfortable with a program that accepts a positive integer and produces a function. I base this suggestion first on the experiences I (and others) have had in working with students who are trying to write such programs and on the fact that writing a computer program that returns a function involves an understanding of a function as an object which can be the result of an operation (computer program any other kind of operation). Researchers seem to agree that object conceptions in general and of functions in particular can be very hard for students to develop. (See, for example Sfard, 1992, 1994.)

An area of research that is currently very active is the study of the mental representations students make after working with such computer constructions. These investigations also try to relate these mental constructions to success in learning mathematical concepts. Much of this work is being conducted by the members of a group known as RUMEC (Research in Undergraduate Mathematics Education Community) and the interested reader can consult their Web Page at

http:
homer.cs.gsu.edu/ rumec/Papers/index.htm.

1.2.6 APOS Theory.

In my other paper for these proceedings, *A Theory-Based Approach to Help Students Learn Post Secondary Mathematics*, I proposed one set of mechanisms for constructing mathematical concepts. These mechanisms involve mental steps such as interiorization to reinterpret an action as a process, encapsulation to understand a process as an object, and encapsulation/de-encapsulation for the process/object dialectic. These ideas are part of what we have called *APOS Theory* (see Asiala et al., 1996 for more details.) The paper also describes some pedagogical strategies based on this theory, using computer programming and cooperative learning, for helping students make these mental constructions.

One can certainly make an APOS analysis of convergence of a sequence of functions. An object conception of function, referred to above would be key here as would the mental construction of processes that transformed these function objects. The general pedagogical approach based on this theoretical perspective and described in my other paper could certainly be applied to this topic and research to study its effect would be very interesting.

In this paper, however, I would like to consider an alternative for this specific concept and in general for the transition from elementary mathematics at the school level to advanced mathematics at the university level—the use of formalism to construct mathematical meaning. Before turning to that discussion, let me hasten to point out that I have not yet developed an explicit pedagogical approach to implement the ideas I am about to discuss, although some of my work may be so interpreted. Nevertheless, I think that at this point, my remarks are by way of a conjecture on how mathematicians construct meaning. Hopefully, later work will apply these ideas to teaching practice.

2 Meaning and Formalism

In this section, I will begin by explaining what I mean by these two terms. Then I will consider the four possible relations between them.

2.1 Clarification of terms

Both of the terms, meaning and formalism, are complex and can have different interpretations from various points of view. Therefore, it is important to clarify the sense in which I will be using these terms throughout this section.

2.1.1 Meaning

The idea of meaning is a deep philosophical concept about which many things have been written by many people. For the purposes of discussion in this paper, I would like to take a fairly unsophisticated point of view and describe a few categories which I think most people would agree to include in an interpretation of the term.

The physical world. The term "real world" has been used but I think that it can be confusing. Constructivists and Platonists will argue about what is "real" and others might point out that one person's reality might be completely alien to another. Although these issues are important, I don't think they relate very much to the questions I am considering in this essay and I would like to avoid them. Instead, let me be fairly narrow and suggest that meaning includes those phenomena to which we have access through our five senses.

Familiar experiences. Another part of meaning for an individual is the set of personal experiences he or she has had and can call up to consciousness through memory. In this category, an individual makes sense out of a situation by relating it in some way to one or more of these experiences.

Connections. Both Vygotsky, 1978 and Skemp, 1976 have suggested that the meaning an individual gives to a concept consists in the relationships he or she makes between that concept and other concepts. In mathematics then, the meaning of the concept of derivative might consist in its relation to the concept of function, to various situations such as rates of change or slopes of tangents, and to the concept of integral or area.

Calculations. Calculations form an essential part of mathematics and we can include in our understanding of meaning the calculations one might make in a particular situation. Thus, for example, a part of the meaning of the concept of derivative resides in the calculations one might make in determining an expression for the derivative of a function given by a polynomial, or some other expression.

Mental images. Although he or she may not always be aware of it, a part of a person's thinking about a concept is based on mental images. Most people have a picture of the set of positive integers, for example, and a part of understanding the concept of limit is a mental process involving motion of a domain value, motion of a range value, and the connection between the two by means of a function (Cottrill et al., 1986).

There is one attribute that is generally assigned to meaning and that is its dynamic nature. Most of the above categories can involve some sort of motion. Our sensual access to phenomena includes—indeed, according to Piaget, consists in—transformation. We sense the weight of an object by hefting it, or trying to move it. Our memories of familiar experiences and many of our mental images are often movies, rather than photographs. It is certainly the case that calculation is active, not static.

2.1.2 Formalism

It is perhaps a little easier to say what formalism is. The key, of course, is notation. By formalism, I am referring to strings of symbols representing mathematical objects and operations and put together according to certain rules of syntax. The objects include characters and words referring to number, sets, functions, booleans, etc. The operations include not only standard operations on numbers and functions such as arithmetic, derivative, integral, but also the logical operations of first and second order propositional calculus. These latter include not only the standard boolean operations of disjunction, conjunction, implication, negation, but also the existential and universal quantifiers.

Unlike meaning, formalism is considered by many to be static and not dynamic. If a mental image is a motion picture, an expression is a string of symbols which sits, doing nothing, on a piece of paper or the blackboard. Actually, there is less than universal agreement on this point and I will have more to say about that below.

2.2 The relationships between meaning and formalism.

Combinatorially, there are four mathematical situations we might consider: meaning alone, the effect that meaning can have on formalism, the effect that formalism can have on meaning, and formalism alone. It is interesting that all four occur and are important. I will discuss each of them separately.

2.2.1 Meaning alone

As Lakoff & Núñez point out, mathematics begins with direct human experience. For some people, unfortunately, it stops there.

One of the problems of mathematics education is that many people almost exclusively use meaning alone in their thinking. That is, a situation will be understood entirely in terms of remembered or imagined personal experience. For example, in a recent study, (Dubinsky & Yiparaki, in preparation) we asked a number of students to interpret certain statements (involving two quantifiers) about everyday situations and about mathematical situations. Following is one of the statements.

> There is a fertilizer for all plants.

In explaining what the statement meant and deciding if it was true or false, many students, on both a written questionnaire and subsequent interviews, focused on their experience and the fact that “you can’t grow plants without fertilizer". The issue of whether this statement asserts that given any plant there is a (possibly varying) fertilizer, or that there is a (single) fertilizer that can be used with all plants was very far from their consciousness. In some cases, they did not seem to be able think about anything other than their remembered experiences and did not address this point, even when the interviewer asked them about it explicitly.

The limitations of focusing only on the meaning were indicated in the same study. When these students were given statements with a similar logical structure, but in a mathematical context that was less familiar to them, most were unable to make any reasonable interpretations at all. For Davis and Maher, one way to learn mathematical concepts is to enhance human experiences, at least at the elementary, pre-college levels. Except for a brief mention of notation such as using superscripts for exponentiation, they say very little about experiences for more advanced mathematical concepts.

2.2.2 Meaning drives formalism

If meaning alone is not enough to understand mathematical situations, we can also consider formalism as well. One way to introduce formal notation is to use it to express the meaning one sees in a situation, to capture an idea in a formal expression. Thus, for example, children who can perform arithmetic might think of a calculation in which one of the numbers involved is not known, but could be determined by the result of the calculation. With this kind of conception, it is possible to work on problems such as,

> If 5 times a number is 30, what is the number?

A more difficult version of essentially the same kind of thinking would be a problem like,

> If 5 times a number is 12 less than 3 times that number, what is the number?

These two problems can be more tractable if one uses a letter, say x for "the number" and writes,

$$5x = 30$$

or

$$5x = 3x - 12.$$

One can then use formal manipulations to find the value of x. Full understanding of this situation might be considered to consist in using the approach of formal manipulation, and simultaneously think of the situations in the verbal terms with which they were originally expressed.

Consider now the more difficult problem that is one of the classical "applications" of calculus.

> A woman is in a boat 3000 meters from her house on the shore along a line perpendicular to the (straight) shoreline. A lighthouse is 2000 meters down shore from her house. She intends to go to the lighthouse by rowing to a point on the shore between her house and the lighthouse and

walking the rest of the way. If she rows at 2 kilometers per hour and walks at 4 kilometers per hour, where should she land so as to minimize the time it takes her to get to the lighthouse?

For this problem, one might begin to think in the same way as in the previous problem. That is, given a point on the shore at which she will land, the time it takes her to row is a certain distance divided by 2 and the time it takes her to walk is a certain distance divided by 4. Using a letter at first in the same way as before, one might write x for the distance from the house at which she lands and express the total time it takes the woman as,

$$\frac{\sqrt{x^2+9\times 10^6}}{2}+\frac{2000-x}{4}.$$

At this point, the issue is no longer to use this expression directly to find the value of x and this change can present serious difficulties for students. It is not uncommon for students to reach this point in the problem, set the expression equal to 0 and try to solve for x. They are trying to apply the methods they are familiar with and have used successfully to solve this new problem. But those methods do not work here and a different conception is needed. Returning to the problem to review this new conception, one can think of the situation as, given a position of the landing point, find the distance traveled by sea and the distance traveled by land, divide each by its appropriate speed and add the two times to obtain the total time of travel. Given the formal expression, one must think now about x not so much as standing for an unknown number to be found, but as a place holder for a given position. One might then express this more complicated conception as a function T given by,

$$T(x)=\frac{\sqrt{x^2+9\times 10^6}}{2}+\frac{2000-x}{4}$$

Graphing techniques can then be applied to obtain an approximation for the landing position which minimizes time or calculus can be used to obtain it exactly.

Thus we see how the use of formalism to express meaning can be an aid to solving algebraic equations. In more complex situations, one can use formalism to define a function whose properties can be studied using calculus. In either case, it is true here that mathematics serves as a language. If a situation is understood, then mathematical formality serves as a language to express that understanding. This is entirely analogous to the early language learning. Once the child understands a chair, say, as a permanent object with various properties and uses, the parent can point to this object and say “chair”. Because the chair is something that exists for the child and is understood by her or him, the parent is able to transmit her or his knowledge that the symbol “chair” expresses this understanding.

In mathematics, this use of formality as language does even more. It is not so easy to perform manipulations on a dynamic situation. A formal expression can capture this dynamism, “tie it down” so to speak, so that manipulations such as arithmetic, or operations such as differentiation, can be applied to it.

Asking a student to express a situation formally for such purposes can provide a problem situation to influence a learner in moving from elementary to more sophisticated mathematical concepts. One example of this is a function defined in parts. Consider the following problem.

> A rock is sliding down a slope. The distance d (in feet) traveled by the rock t seconds after it started sliding is 6 times t for the first 3 seconds and the square of t plus 9 thereafter. Find the velocity of the rock 3 seconds after it starts sliding.

This problem can be solved by expressing the situation as a function d which gives distance as a function of time, and calculating the derivative of d at $t = 3$. The function can be expressed as,

$$d(t) = \begin{cases} 6t & \text{if } t \leq 3 \\ t^2 + 9 & \text{if } t > 3 \end{cases}$$

Using either the definition of derivative at a point, or a theorem about one sided derivatives, one can find the desired velocity fairly easily.

MacLane takes the position that all advanced mathematical concepts can be understood by expressing complex human actions formally and then arguing logically and rigorously about the formal statements. In our examples, the formal expressions consisted in an algebraic equation or a function and the rigorous arguments involved manipulations to solve the equation or taking derivatives of functions given by expressions. In general, including the most sophisticated situations, MacLane, 1981 asserts that "mathematics started from various human activities which suggest objects and operations (addition, multiplication, comparison of size) and thus lead to concepts (prime number, transformation) which are then embedded in formal axiomatic systems (Peano arithmetic, Euclidean geometry, the real number system, field theory, etc.)"

The difficulty with this idea of trying to move forward in mathematical understanding by expressing a situation or human activity formally and reasoning about or manipulating the formal expression to reveal deep properties of the original situation is the danger of losing sight of the relation between the formal expression and the situation it expresses. It is certainly true that the situation is dynamic, involving transformations and motions, while the formal expression, once written down is entirely static, consisting of pencil marks on paper or chalk marks on a blackboard. It is indeed, as I indicated above, this static nature that allows manipulating the situation and reasoning about it. But a price must paid and that is that care is necessary to maintain the connection between a situation and the formalism that expresses it. When educators object to the formal aspects of mathematics as consisting in meaningless strings of symbols that make no sense to students, I do not think the difficulty lies in the nature of the formal expressions, but rather in the loss of the connections between them and the situations.

Let me pursue this point in terms of the example of the sliding rock. It is clear what I mean by the formalism in this case — it is the expression used to define the function in parts. By the dynamic situation, however, I am not referring to the motion of the sliding rock. Rather, I mean the mental dynamics of taking a particular time, determining whether it is less than or equal to 3 or if it is greater than 3, and using this dichotomy to choose one of the two expressions in which the value of t is substituted and the calculation made. For an individual who connects, in her or his mind, the formal expression with the process of evaluation, the expression is not a static string of meaningless symbols, but rather an expression of the full dynamics of the situation.

Unfortunately, this connection can be difficult for students to make. An intermediate step in going from the situation to the expression and maintaining an awareness of the connection between the two can be to express the situation as a computer program. Thus the distance function for the sliding rock can be embodied in a computer program as follows,

```
d := func(t);
       if t <= 3 then return 6*t;
       else return t**2 + 9;
       end;
     end;
```

It seems that writing such a program and thinking about what the computer does when asked to evaluate, say, `d(2.5)` helps students understand and maintain an awareness of the connection between the formal expression and the process it embodies. There is evidence that such an approach can lead to significant enhancements of mathematical understanding. Some of the results can be found in the RUMEC Web mentioned above and there is a work in progress that will summarize all that is presently in the literature.

2.2.3 Formalism drives meaning

In the previous section I considered how formalism can be used to capture the meaning that arises out of experience. In that point of view, formalism acts like a language to describe the meaning that already exists for a person, and then the manipulations that can be made with formal expressions are tools that can be used to deepen understanding of the mathematical experiences. It is possible to go in the opposite direction. In this section, I would like to introduce and begin to explore an alternative means of creating meaning. It is possible to use formalism to create meaning.

One of the tentative conclusions reached in the study Dubinsky & Yiparaki (in preparation), was that many students dealt with complex statements (involving quantifications, for example) by calling up familiar experiences and relating them, not always in a completely accurate manner, to the statement. Then, all of the subsequent thinking by these students was about the situation and not about the statement. For example, two of the statements we asked the students to interpret were the following.

For every positive number a there exists a positive number b such that $b < a$.

and

There exists a positive number b such that for every positive number a $b < a$.

A number of students in this study and in other studies (e.g., Dubinsky et al., 1989) were able to interpret the first statement reasonably accurately, and then insisted that the second statement was the same as the first. When an interviewer pointed out the difference between what the two statements actually said (e.g., that the order of quantifiers was reversed), some students reported explicitly what appeared to be true of most students—they did not pay attention to the syntax of the statement.

Paying attention to the syntax is what seems to be missing for the students who equated the two statements. I want to suggest here that some people who avoid this error do it by working with the formal syntax of an expression in order to make sense out of complicated mathematical statements. Moreover, with appropriate pedagogy, other students can learn to do this. They can learn to make and use a close and detailed analysis of the formal expressions in a statement to create the meaning which the formal expression describes. This is, of course, exactly the opposite direction of what was described in the previous section.

As an example, let me return to the concepts of pointwise and uniform convergence of a sequence of functions and describe how I think such an analysis can be made. Here, again, are the formal definitions of the concepts. Recall that we are considering a sequence (f_n) of real-valued functions all defined on a common domain D.

Pointwise convergence to a function f on D:

$$\forall x \in D,\ \forall \epsilon > 0,\ \exists \text{ a positive integer } N\ \ni\ \text{if } n > N,\ \text{then } |f(x) - f_n(x)| < \epsilon,$$

Uniform convergence to a function f on D:

$$\forall \epsilon > 0,\ \exists \text{ a positive integer } N\ \ni \forall x \in D,\ \text{if } n > N,\ \text{then } |f(x) - f_n(x)| < \epsilon,$$

Although this may not be the only way to do it, my conjectured analysis will be in the framework of the APOS theory. Thus, the key abilities I am attributing to the individual trying to make sense out of this statements is to understand functions as objects; to be able to interiorize a process by reading a description of it — provided that those concepts the process transforms do exist for the individual as objects; to be able to interiorize a process the individual has constructed so that it becomes an object; and, in the latter case, to be able to de-encapsulate that object back to the process from which it came.

For the first statement, describing pointwise convergence, the individual would think of x, ϵ, N as being fixed and would construct a process of n visiting every positive integer greater than N, checking the inequality $|f(x) - f_n(x)| < \epsilon$. He or she would then encapsulate that process to an object and apply a universal quantification to get the overall value, $true$ or $false$. Now, thinking now about varying x, ϵ, N and obtaining, for each value of the triple, a boolean value, the individual creates a process which is a function—of three variables. There is then a general mechanism for a boolean-valued function of several variables in which all but one of the variables are kept fixed and the resulting function of the remaining variable, considered as a process is encapsulated to obtain an object to which a quantification is applied, resulting in a boolean value. Now, allowing the other variables to vary, one obtains a new boolean-valued function, but of one less variable. This total process, analogous to the iterated integral of a real-valued function of several variables is iterated until one obtains a single, boolean, value.

Of course, no actual value is obtained unless the investigation is being made of a particular example. The general situation is a process in which the individual can imagine doing all this if a sequence of functions and a candidate for the limit function were in hand.

My assertion is that it is possible (see, for example, Dubinsky, 1997 for data that supports this contention) for students to learn to make such mental constructions and when they do, they can reflect on this overall process they have constructed and see that if they rearrange things, for example instead of iterating in the order x, ϵ, N, they could use the sequence ϵ, N, x, then the mental process *they construct* is different from the original. A discussion of the distinction between pointwise and uniform convergence is then more likely to be meaningful than if these mental constructions had not been made.

Incidentally, this "inside-out" construction is not the only way to apply the construction of processes and objects and the formalism of quantification to such a problem. One can also posit an "outside-in" approach. In this case, for pointwise convergence, the individual would construct a process P_1 of successively assigning to a variable x each value in the domain D. Every time an assignment is made, a process P_2 is constructed, using that value of x. This process assigns to the variable ϵ each positive number and again, with each assignment a process P_3 is constructed consisting of assigning to the variable N each positive integer. Now, with this value of N (and the same values of x and ϵ) the individual constructs yet another process P_4 consisting of assigning to the variable n every positive integer larger than N and checking whether the absolute value of the difference $f(x) - f_n(x)$ is less than ϵ.

Now, the individual reviews this collection of nested processes. The process P_4 is encapsulated and a universal quantification is applied to obtain result, true or false. This allows the encapsulation of the process P_3 and the application of an existential quantifier to obtain a boolean value, and so on, to obtain a single boolean value for the entire expression.

As in the first case, a similar operation can be constructed for the second statement and the student has a chance of seeing the difference between the two.

The research quoted above uses a mixture of the "inside-out" and "outside-in" approaches and it would be interesting to investigate the relative value of these two alternative approaches.

I acknowledge that what I have just described is extremely difficult to comprehend and requires very powerful mental activity. I suggest, however, that for many people, such thinking is exactly what is required to understand these two kinds of convergence and the distinction between them. The fact that this thinking is difficult is my explanation for why convergence of functions is a difficult mathematical concept.

Let me note that the kind of thinking required to use formalism to construct meaning as suggested here may be one characteristic of mathematical thinking that separates advanced from elementary, school (or street) mathematics from advanced mathematics. Note also that this is one position relative to the question of whether a formal statement is static or dynamic. It is not possible to disagree with the observation that a string of symbols written on a piece of paper is static. But the nature of anything is never in the thing itself, but in the relation between it and an individual who observes it. If what I am describing does actually occur, then whatever the nature of the formal statement in the absence of human thought, when an individual is using it to construct meaning then the formal expression is full of the dynamism of the several processes the individual constructs. My conclusion is that those who speak of meaningless formalism are not talking about the concept, but of the people who are failing to use it to construct meaning.

The two relations between meaning and formalism I have discussed in this and the preceeding section can help put into focus the relation between language and meaning. My view is that once meaning has been constructed, language can be used to describe it and this forms a powerful tool for thinking about that meaning. On the other hand, although the necessity to express an idea linguistically can be a motivation and provide a drive for constructing meaning, the use of language is not a mechanism for making that construction. This must be done some other way, for example constructing processes and objects and using them to analyze the formal expression directly as I have described here. Or to use the words of L. Tolstoy, 1903 to express what I am saying succinctly, "There is a word available nearly always when the concept has matured".

Finally, let me say that although I have described the bi-directional relationship between meaning and formalism as two separate relations, this is, so to speak "the laboratory description". What probably happens "in nature" is that there is a dialectic in which meaning drives formalism and formalism drives meaning at the same time, or in an oscillation, each affecting the other.

2.2.4 Formalism alone

Finally, completeness requires that we consider the case of formal statements alone with little or no meaning attached to them—strings of symbols with rules of manipulations but with no interpretations for the symbols or the manipulations. Anyone who is not aware that such exists in the minds of many people has not tried to teach mathematics. We all have seen calculus students whose total understanding of differentiation consists in pieces of "knowledge" such as,

> "The meaning of the derivative of x^3 is that you move the 3 down and put it before the x and reduce the exponent by 1 to get $3x^2$.

At first thought, we reject this level of understanding and devote all of our pedagogical efforts to helping students go beyond such a primitive level of understanding. But it is perhaps a truism in mathematics that *any* idea which is persistent in the minds of individuals, even if they are only our weakest students, is never completely useless. We can think of at least one context in which it is useful to think about formal manipulations without paying attention, at least for a time, to any meaning.

Consider for example the general question: Does every function come from an expression? In calculus, we are fully aware that not only is this statement false, but also that getting our students to reject it and expand their view of function is a major pedagogical challenge. But in other contexts, the situation can be different. Suppose we are considering function in the sense of a boolean-valued function of n boolean variables. Does every such function come from an expression?

It is interesting that not only is the answer to this question yes, but the proof consists in thinking about expressions and manipulations without paying much attention to meaning. Roughly speaking, the argument begins with forming all expressions in the n variables $p_1, p_2, \ldots, p_n$ of the form $x_1 \wedge x_2 \wedge \ldots \wedge x_n$ where each x_i is either p_i or not p_i. There are 2^n such expressions. Using them, one can form a total of 2^{2^n} expressions by taking a subset of these "monomials" and connecting them with conjunctions $\vee$. They all form different functions. On the other hand, by representing a boolean valued function of n boolean variables with a table, it is easy to see that there are exactly 2^{2^n} such functions.

This proof of the assertion that can be paraphrased as saying that in this context, all meaning (functions) comes from formalism (boolean expressions), is made entirely with meaningless (but far from mindless) manipulation (including counting) of symbols. Hence, this form of thinking so dear to our students cannot be totally rejected.

Al Cuoco has referred to thinking like this in suggesting that one of the characterizations of abstract algebra is as a study of the forms of manipulations of formal expressions. With his permission, I quote his comment from a private communication.

> Gradually, over the past 150 years, algebraists have developed a "systems approach" to the analysis of operations. The things studied in algebra have evolved from *techniques* for solving equations, to a study of the *properties of operations* that allow one to develop techniques for solving equations, to a study of entire *systems* in which one can calculate (and hence solve equations). These *algebraic structures* have become the primary focus of modern algebra. An algebraic structure is a collection of objects together with one or more operations that can be used to calculate with the objects. . . . The structural approach to algebra has enormously widened the kinds of systems in which algebraists work, and hence has changed the face of what's considered "algebra."

3 Implications for research and pedagogy

What does all of this say for research and its applications to pedagogy, especially as regards advanced mathematical thinking and the transition from learning school mathematics to learning university level mathematics? What we have here are competing philosophical, epistemological, and pedagogical positions, and although for each of these positions there is evidence that tends to support it, there is nothing like decisive results. Therefore, a choice among these positions must be a matter of personal opinion, informed as much as possible by existing research. What follows are my own opinions on these matters.

Because we are talking about learning mathematics (as opposed to working on unsolved problems or creating new mathematical concepts) I take the constructivist position that an individual must, in a social context, construct her or his understanding of mathematics and it must fit with generally accepted understandings. I do not feel that human experiences in the physical world can be sufficient to form the bases of representations that lead to understanding advanced mathematical concepts.

Although I find the ideas of grounding and linking metaphors very interesting from a theoretical point of view, I am seriously disturbed that these ideas cannot lead to any pedagogy other than using such metaphors in discourse with students and suggesting they think in such ways. This is not really different from what many college mathematics instructors do and I think the vast experience over the last few decades shows clearly that this will not be very helpful for many students other than those who would succeed in learning more or less under any pedagogical system.

Although the idea of using metaphors to create mathematical concepts in the minds of individuals is very attractive in that it can give mathematics an esthetic literary flavor, I remain convinced that language is a tool whose real value is the expression of ideas which an individual already has constructed. I do not feel there is any reason to believe that it is powerful enough to be useful in creating more than the simplest mathematical concepts.

What remains are the ideas of MacLane, APOS Theory and the use of formalism to create meaning. I would propose a combination of these three mechanisms. It could look something like this. The starting point for getting students to move from elementary to advanced mathematical thinking would be to consider that the population we are talking about is sufficiently able to create actions relative to even fairly complex mathematical situations. The first step then is to help them develop the ability to interiorize actions to processes, encapsulate the resulting processes to objects, de-encapsulate the objects back to processes, and organize all these constructions in coherent schemas. There exist methods for doing this using computers and cooperative learning and these are described in the literature. There may be other methods that can be developed.

The next step would be for students to learn to analyze complex formal statements in mathematics and construct meaning based on these formalisms. The APOS activities would be a key mechanism for making the required analyzes.

At the same time, a parallel development would take place in which students would learn to express in formal language meaning they have already constructed. Writing, running, and reflecting on computer programs to express mathematical concepts can be an important mechanism for achieving this.

The two relationships between meaning and formalism would then be coordinated in a manner that points, as a goal, to the use of axiomatization and rigorous deductive reasoning about formal statements (which, at this point, would all have meaning for the students) to reveal — and prove — deep properties of the concepts expressed formally.

Finally, I would like to recommend the continuation of research that develops this approach and applies it to specific material in the university mathematics curriculum using pedagogy based on this development. Large-scale research studies could then investigate the extent to which this theoretical approach resonates with what students need to help them learn advanced mathematics and realities of teaching practice in trying to meet those needs.

4 References

M. Asiala, A. Brown, D. DeVries, E. Dubinsky, D. Mathews & K. Thomas, 1996. A Framework for Research and Curriculum Development in Undergraduate Mathematics Education. In Research in Collegiate Mathematics Education II, pp 1-32.

E. W. Beth & J. Piaget, 1965. Mathematical Epistemology and Psychology (W. Mays, trans.), Dordrecht: Reidel. (Original published 1965)

J. Cottrill, E. Dubinsky, D. Nichols, K. Schwingendorf, K. Thomas and D. Vidakovic, 1986. Understanding the Limit Concept: Beginning with a Coordinated Process Schema, Journal of Mathematical Behavior, 15, 2 pp. 167-192.

R. B. Davis & C. A. Maher, How Students Think: The Role of Representations, in L.D. English (Ed.), Mathematical Reasoning: Analogies, Metaphors and Images, Mahwah, NJ: Erlbaum, 1997, pp. 93-15.

E. Dubinsky, 1995. ISETL: A Programming Language for Learning Mathematics, Comm. in Pure and Appl. Math., 48, pp. 1-25.

E. Dubinsky, 1997. On Learning Quantification, Journal of Computers in Mathematics and Science Teaching, Journal of Computers in Mathematics and Science Teaching 16(2/3) 335-362.

Ed Dubinsky, Flor Elterman, and Cathy Gong, 1989. *The Student's Construction of Quantification,* For the Learning of Mathematics, (2), 44-51.

E. Dubinsky & O. Yiparaki, in preparation. On student Understanding of AE and EA quantifications.

E. von Glasersfeld, Learning as a constructive activity. In C. Janvier (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics,* Hillsdale, NJ: Erlbaum, 1987, pp. 41–69.

R. Hersh, Some Proposals for Reviving the Philosophy of Mathematics, In T. Tymoczko (Ed.), New Directions in the Philosophy of Mathematics, Boston: Birkhauser, 1986, pp. 9-28.

I. Lakotos, 1976. Proofs and Refutations, London/New York: Cambridge University Press.

G. Lakoff & R. E. Nuñez, The Metaphorical Structure of Mathematics: Sketching out Cognitive Foundations for a Mind-Based Mathematics, in L.D. English, (Ed.), Mathematical Reasoning: Analogies, Metaphors and Images, Mahwah, NJ: Erlbaum, 1997, pp. 21-89.

S. MacLane, 1981. Mathematical Models: A Sketch for the Philosophy of Mathematics, American Mathematical Monthly, Aug.-Sep., pp. 462-472.

R. Noss, Meaning Mathematically with Computers. In: Bryant P. & Nunes T. (eds) *Learning and Teaching mathematics: an International Perspective.* pp. 289-314. Hove, UK: Psychology Press. 1997.

J. Piaget, & R. Garcia, 1989. Psychology and the History of Science, New York: Columbia University Press.

A. Sfard, 1992. Operational origins of mathematical notions and the quandary of reification—the case of function. In (G. Harel and E. Dubinsky, ed.) The Concept of Functions: Aspects of Epistomology and Pedagogy, MAA Notes, 25, pp. 59-84.

A. Sfard, 1994. Reification as the birth of metaphor. For the Learning of Mathematics, 14, 1, pp. 44-55.

A. Sfard, Commentary: On metaphorical roots of conceptual growth, in L.D. English, (Ed.), Mathematical Reasoning: Analogies, Metaphors and Images, Mahwah, NJ: Erlbaum, 1997, pp. 339-371.

R. Skemp, 1976. Relational understanding and instrumental understanding, Mathematics Teaching, 77, 20-26.

L. Tolstoy, 1903. *Pedagogicheskie statli* [Pedagogical Writings], Moscow: Kushnerev.

L. Vygotsky, 1978. Mind and Society: The development of higher psychological processes, Cambridge, MA: Harvard University PRess.