Building Formal Mathematics on Visual Imagery: a Case Study and a Theory

MÁRCIA PINTO, DAVID TALL

The transition to formal mathematical thinking involves the use of quantified statements as definitions from which further properties are constructed by formal deduction. Our quest in this article is to consider how students construct meaning for these quantified statements. Dubinsky and his colleagues (Dubinsky *et al*, 1988) suggest that the process occurs through reflective abstraction, in which a predicate with one or more variables is conceived of as a mental process that is encapsulated into a statement (a mental object) by the process of quantification.

In this article, we report on a case study of a university student who constructs the formalism not from processes of quantification, but from his own visuo-spatial imagery. Rather than construct new objects from cognitive processes, he reflects on mental objects already in his mind and refines them to build his own interpretation of the formal theory. This example leads us to consider the development of theory in the literature, in particular Piaget's notions of pseudo-empirical abstraction (focusing on processes encapsulated as mental entities) and empirical abstraction (focusing on the properties of the objects themselves).

It has often been noted that highly successful mathematics researchers show strong preferences for different kinds of approach (e.g. Poincaré, 1913; Hadamard, 1945; Kuyk, 1982; Mac Lane, 1994; Sfard, 1994). Some have a broad problem-solving strategy, developing new concepts that may be useful before making appropriate definitions to form the basis for a formal theory. Others are more formal from the beginning, working with definitions, carefully extracting meaning from them and gaining a symbolic intuition for theorems that may be true and can be proved.

In a recent research study of novice mathematicians' styles of doing mathematics (Pinto, 1998; Pinto and Tall, 1999), we found analogous differences among students' strategies for learning mathematics. Some worked by *extracting meaning* – beginning with the formal definition and constructing properties by logical deduction. This strategy is consonant with the theory of Dubinsky and his colleagues, in which multi-quantified statements are grasped by working from the inner quantifier outwards, converting a predicate (as a process) into a statement (as a mental object).

However, our research shows that there are students who use an entirely different strategy. In this article, we focus on one successful student who builds from his imagery, *giving meaning* to the definition by producing a highly refined image that supports his formal arguments. We suggest he is not encapsulating a cognitive process into a mental object; he progresses by refining and reconstructing his existing imagery, until it is in a form that he can use to construct the formal theory. He uses both visual and symbolic coding of ideas in a complementary manner that is characteristic of the dual-coding theory of Paivio (1971, 1986). His thinking processes resemble those of mathematicians using broad problem-solving strategies, rather than those who focus more particularly on purely formal deduction.

The data presented here form part of a larger qualitative research study of students' understanding of real analysis (Pinto, 1998). A group of U.K. mathematics undergraduate students was followed during the students' first two terms studying analysis, with each participating in an hour's individual interview every two weeks. Interviews were recorded on tape and fully transcribed for detailed analysis.

Procedures of data collection were compatible with those used in Strauss's method of building up a theory from the data (see Strauss, 1987; Strauss and Corbin, 1990). Pinto followed a plan where each set of questions formulated for successive interviews was built upon results of the analysis of data from previous interviews, enabling a gradual transformation and enrichment of the theoretical viewpoint.

The case study

Chris is a U.K. mathematics student (and a native English speaker) who appears to have a consistent background in school mathematics, with well-formed concepts and an ability to construct formal arguments. He is one of a small minority of first-year mathematics students who obtained full marks on a questionnaire designed for selection of a spectrum of students for the full study. He was considered a gifted novice mathematician who later fulfilled his promise by subsequently obtaining a first-class honours degree.

His initial difficulty with the definition of a limit of a sequence is formulated as follows in his first interview in the fourth week of the course:

> after the first time, I mean, in the first lecture on limits, I didn't quite get it [the definition of limit of a sequence].
>
> (Chris, first interview)

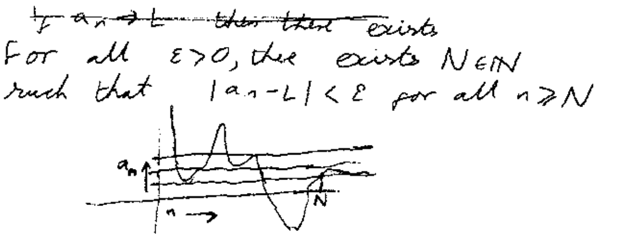
He then comments that he searched for an explanation in sources other than his lectures and lecture notes:

> But then I looked it up [the definition of limit of a sequence] in a book and understood . . and then ... I don't know umm it seems now okay.
>
> (Chris, first interview)

He appears to be looking for other presentations of the concept definition, attempting to make it more *concrete* for himself in the sense of Wilensky (1991), where 'concreteness' involves a richer quality of relationship between the individual and the concept.

For the Learning of Mathematics **22**, 1 (March, 2002)
FLM Publishing Association, Kingston, Ontario, Canada

In the first interview, Chris is able to write down a correct definition of the limit of a sequence:



(Chris, first interview)

His definition is verbalised without symbolic quantifiers and is already accompanied by a visual representation. This visual representation is not, however, a simple picture of an 'increasing sequence' bounded above nor a 'decreasing sequence' bounded below nor even an alternating sequence oscillating above and below the limit. It moves up and down in a more general manner. He uses it not as a *specific* picture, but as a *generic* picture, one that represents the sequence in a manner that is as general as he possibly can make it.

His picture is not perfect (for instance, he uses a continuous curve to represent the wanderings up and down, rather than a sequence of distinct values). Nor is it possible for this picture to cover *all* possible cases in his imagination. For example, it shows a sequence wandering all over the place, but seems inappropriate for the case of a constant sequence.

When Chris is presented with the constant sequence 1, 1, 1, ..., he laughs nervously, apparently realising that a constant sequence clearly satisfies the definition (so it is convergent) and yet feeling it clashes with his current imagery:

> [Laughter] I don't know really. It definitely it will .. it will always be one .. so I am not really sure [laughter] .. umm it's strange, because when something tends to a limit, you think of it as never reaching it .. so if it's .. one .. then, by definition, it has a limit, but .. you don't really think of it as a limit [laughter], but just as a constant value.

(Chris, first interview)

Such clashes, and their ultimate resolutions, are an inevitable part of the strategy of 'giving meaning' that Chris uses to construct meaning for the limit concept.

When he works with the definition of a convergent sequence, his generic picture of a wandering sequence gives him a visual image sufficiently general for him to construct the definition and build the theory, even though he also needs to consider variants such as constant sequences. It allows him to reproduce the definition in a correct formulation that is not a direct reproduction of the definition given in class. For instance, he used the terminology "for all" instead of the term "for any" given in the formal definition during the lectures. He also later prefers to use "if and only if" in his definitions, rather than just "if", as was generally written by the lecturer.

The verbal definition comes *from the picture*, not the picture from the definition

> I don't memorise that [the definition of limit], I just ... think of this [picture] every time .. I work it out .. and then you just get used to it, so . I'm very much getting used to it now, I can .. near enough write that straight down.

(Chris, first interview)

He sees "the general in the specific" (Mason and Pimm, 1984) as a thought experiment. His combination of mental imagery, its verbal equivalent and its ensuing properties all fit together as a powerful cognitive unit (in the sense of Barnard and Tall, 1997).

He has control over the language he uses to express his meaning. As he draws the graph, he explains:

> I think of it .. graphically .. I think of it . so like you've got the graph there .. and you've got like the function there, and I think that .. ... it's got the limit there .. and then epsilon, once like that ... and you can draw along and then . all the ... points after N there .. are inside of those bounds. .... It's just . err when I first . thought of this, it was hard to understand, so I thought of it like this, like . that's the n going across there and that's a_n

(Chris, first interview)

Notice that he refers to "all the .. points after N there .. are inside of those bounds". It is as if he is looking at the points with ordinates in the range $L \pm \varepsilon$ and is scanning from left to right to find the value N for which every point afterwards is in the desired range. He is so focused on the imagistic behaviour of the sequence going up and down, that he draws a continuous graph representing the movement, not a sequence of distinct points.

When asked about the domain, he suddenly corrects his error:

> Err, this shouldn't really be a graph, it should be points

(Chris, first interview)

Chris's mistake in the midst of perceptive observation and analysis is typical of the way in which the human brain works. He is, at the time, focusing on certain vital elements in his argument and, as he does so, is less aware of peripheral elements. In thinking about the movement up and down, and homing in on the limit value, he is not concentrating on the means by which he achieves this in his physical drawing. To be able to realise his error and to self-correct requires him to build a stable schema, one in which dissonant elements can be perceived and corrected whilst other parts remain coherently in place. It is precisely such a mental schema that Chris builds by manipulating his imagery and adapting it as his ideas mature. He has a broad enough grasp of the principles to be able to deal with errors and misconceptions, and to reconstruct them in his quest for a coherent whole.

Notice that his visual representation is both *analogical* (in the sense of Eysenck and Keane, 2000) and also *enactive*. As he draws the picture, he goes through a specific sequence of construction, first drawing the axes and the graph, then labelling it. The graph represents the given sequence. Its wandering path represents a general sequence rather than just an increasing, decreasing or alternating sequence. He then puts in the horizontal line representing the limit L and, for given ε, the lines at distance ε above and below L to represent the range in which the sequence is desired to lie. He then marks the value N, gesturing at the "points after N", which lie in the required range between $L-\varepsilon$ and $L+\varepsilon$ He therefore begins with the whole gestalt in mind, reproducing it sequentially in the order given in the definition.

He does *not* concentrate on the symbols alone to build successively from the inner quantifier out, as postulated in the genetic decomposition suggested by Dubinsky *et al.* (1988). To build the definition, he follows a sequence of actions that is the same as the presentation of the written definition from left-to-right.

The picture already includes a representation both of the limiting *process* and of the limit *value*. The picture and his actions in using it therefore already display the duality of process and concept contained in the procept notion of Gray and Tall (1994) For Chris, it is not a matter of *first* the process, *then* the object; rather, he has a dynamic gestalt which comprises them both.

His approach can be framed in the embodied sense formulated by Lakoff and Johnson (1999), relating the limiting process to the limit concept in what Mason (1989) terms "a delicate shift of attention" This is suggested not only by his visual representation of a convergent sequence, but also by his explanation (given above) which includes the remarks "you've got like the function there" and "it's got the limit there".

In describing his initial struggle (which referred to a 'null-sequence' tending to zero), Chris explains:

> I didn't realise that you had to ... just find an N ... like such that ... modulus of a_n is less than epsilon I didn't quite ... grasp that last bit
>
> [Interviewer: How do you get the N, or the meaning?]
>
> The fact that actually you get one. [laughter] I didn't quite understand it. . I looked it up and then once I realised .. that ... you have to just find ... a value that might depend on epsilon ... then ... [I could] see what the definition meant.
>
> (Chris, second interview)

Chris's focus on "the fact that actually you [need to] get one" value of N, and his earlier mentioned analogical handling of the "points after N", shows him focusing first on the finding of N with the required property for a given value of ε, and only then realising it involved "a value [of N] that might depend on epsilon" His strategy therefore has a subtle complexity. It does not always build strictly left to right, as he did in his explanation of the definition from a dynamic interpretation of the picture. Here, he temporarily fixes ε in the diagram, looks along to see where the points cease to wander outside the range $L \pm \varepsilon$, and then finds the value N Only then does he allow ε to vary and notes that the value of N may then depend on ε. One thing is certain: his sequence of operation is *not* to start at the inner quantifier and move outwards in a symbolic, propositional manner.

As late as his seventh interview, three months later, he still speaks of convergence of a sequence in the same terms:

> Umm ... the thing is ... when you think about why ... why you are actually doing it, .. then ... that's when it becomes clear. You find out why you are choosing the N so they lie all there in, so .. it gradually tends towards that limit.
>
> (Chris, seventh interview)

He explains that, at the time, he was dealing with only one value of ε:

> I think it was that .. I wasn't thinking ... generally about that ... I wasn't thinking that generally it works for any epsilon ... I was just thinking ... of one case.
>
> [Interviewer: Yes, just fixing one.]
>
> Yes. (Chris, seventh interview)

He completes his argument by remarking that although you only need one N at a time, (and so this may not force the terms to get any closer to the limit), the closeness can be achieved by varying the value of ε:

> But ... it's where ... you are just choosing one value of N ... so all the points after that could do whatever they like inside, and when you actually ... think that you can, .. you make epsilon small
>
> (Chris, seventh interview)

His building of meaning involves a considerable struggle In his first interview, he explains that he experimented with the definition by first giving N and then attempting to find a related ε:

> you decide how far out ... and you can work out an epsilon from that ... or if you choose an epsilon you can work how far out.
>
> (Chris, first interview)

Moving N to the right and determining ε allows a dynamical feeling that the sequence is tending to a limit. However, he realised that just finding ε does not ensure that the terms get as close as is desired, for ε need not be forced to be small. Such an experiment may have helped him focus on the correct definition and why ε must be given first in order to find a corresponding value of N.

At the end of the course in the second term, Chris reiterates that he has not just sought to learn the definition by heart, but to give it meaning in his own terms. When asked once more to give the definition of a limit of a sequence, he says:

> I can write down that definition ... without making it formal.

He then writes:

> A sequence has a limit if and only if as the sequence progresses, eventually, all values of the sequence gather around a certain value.
>
> (Chris, seventh interview)

As it stands, this statement could intimate the epistemological obstacle that a convergent sequence does not ever 'reach' the limit, as was present in his earlier discussion of convergence before he met the constant sequence However, he is also able to reproduce the formal definition and, in doing so, attempts a more complex task than simply reproducing a remembered formulation. He uses colloquial language, and metaphors, to explain his visual imagery in a way he can translate into formal terms His link between visual and verbal involves active construction rather than just reproduction of a formal definition We suggest that Chris has built a powerful structure of knowledge, constantly being reconstructed, which will surely be of assistance to him in other contexts.

Chris showed a powerful ability to deal with formalism For instance, when he is asked what it means to say that a sequence does not have a limit, he writes the following straight down:

A sequence (a_n) does not tend to a limit if
for any L, there exists $\varepsilon > 0$ such that
$|a_n - L| \geq \varepsilon$ ~~whenever~~ ~~[illegible]~~
for some $n \geq N$ for all $N \in \mathbb{N}$

(Chris, second interview)

Note his minor alterations at the end, replacing part of the phrase beginning "whenever $n \geq N$ ... " (from the definition of convergence) by "for some $n \geq N$ for all $N \in N$" (required for the negation of the definition). What is impressive is that this struggle for meaning was not performed by writing the quantified definition and then negating it by swapping quantifiers. He seemed genuinely preoccupied with the *meaning* of non-convergence, rather than manipulating the formal syntax of the definition. This focus on meaning involves extremely subtle thinking processes. Chris's construction of a *prototypical* mental picture of convergence, together with his sense of its limitations, afforded him a powerful interplay between thought experiment and formal proof, which *for him* provides a way to translate imagery into formal linguistic terms.

Case study summary

This case study reveals a student grappling with imagistic ideas to translate into a formal definition He constructs the concept of convergence through thought experiments that respond not only to the syntax of the definition but also that attempt to give an imagined *meaning* for the definition He first attempts to understand the statement as a property, one satisfied by his mental image of the object to be defined. He then gives meaning to the statement from his image, by exploring and verifying how it works The crucial idea is to understand how his image characterises the mental concept of convergence that he is attempting to construct. This mental construction involves playing with the image in various ways.

At one stage, he conceives a thought experiment, moving N to the right to force ε to decrease This gives him a dynamical feeling that the sequence tends to the limit. However, making N bigger does not necessarily force ε to become smaller, as the generic sequence he imagines goes all over the place and does not move successively closer to the limit. He therefore confirms that he must *first* select $\varepsilon > 0$ and only *then* find the N for which the following terms of the sequence are within ε of the limit As he can then make ε small, the graphical image reveals terms ultimately within any prescribed value ε of the limit

This guarantees for him that the sequence converges to the limit It suggests a *reconstruction* of his understanding of the formal definition, now acquired as a criterion that characterises the concept of limit. The sequence of actions performed by Chris in his exploration appears to involve mental actions *with* a mental object that is successively refined rather than actions *on* encapsulated processes.

To summarise, Chris *interprets* the definition in terms of his old knowledge, *explores* the concept through thought experiment and *reconstructs* his understanding of the concept definition. He compresses information in a picture, which he evokes when writing down the definition. He is operating in a context that has both limit processes and limit objects and he explores and refines his *existing prototypical image* of limit, rather than encapsulating the limiting process into a limit concept.

Alternative frameworks for analysing the data

In developing a wider theory, one that takes into account the data in this case study in addition to other approaches, we refer to various theoretical frameworks that have been formulated in the literature of advanced mathematical thinking (e.g. Tall, 1991, and subsequent developments).

In particular, Dubinsky and his colleagues (Dubinsky, 1986; Dubinsky *et al*, 1988; Cottrill *et al.*, 1996) suggest a *genetic decomposition* for definitions having several quantifiers, which we understand as follows.

The definition of convergence of a sequence (a_n) to a limit L might be given formally as a three-level quantification, either as:

$$\forall\ \varepsilon > 0 \exists N \forall n{:}(n \geq N \Rightarrow |a_n - L| < \varepsilon)$$

or:

$$\forall\ \varepsilon > 0 \exists N \forall n \geq N : |a_n - L| < \varepsilon$$

Dubinsky *et al.* (1988) suggest that a student copes with such quantified statements by building up from the inner, single-level quantification to successive, higher-level quantifications

Their theory is closely related to the internal structure of the programming language ISETL, which offers a metaphor for how we may think of quantifiers. ISETL deals with *finite* sets and so can test quantified statements by considering all the elements in turn. A single quantified statement is of the form:

$\forall\ x \in S{:}P(x)$, written in ISETL as "for all x in S | $P(x)$"

or:

$\exists\ x \in S{:}P(x)$, written in ISETL as "exists x in S | $P(x)$", where $P(x)$ is a predicate that is either true or false for each x in the finite set S. In ISETL, the truth of the quantified statement:

for all x in S | $P(x)$

is found internally, by running successively through the elements $x_1, \ldots, x_n$ of S and testing each statement $P(x_r)$. If one of the $P(x_r)$ is found to be false, the value 'false' is returned at once. On the other hand, if the tests are completed and every $P(x_r)$ is true, the value is returned as 'true'. In a similar way:

exists x in S | $P(x)$

tests each statement $P(x_r)$ in turn, returning 'true' if one is found to be true and 'false' if all are found to be false.

A predicate $P(x, y)$ with two variables can be handled by a successive application of this principle. The statement:

$\forall x \in S \; \forall y \in T: P(x, y)$

is handled in two steps. First, fix x and consider the inner statement:

$\forall y \in T: P(x, y)$

By iterating through all values of y, the truth of this statement can be tested, to give a statement in the single variable x:

$Q(x) = [\forall y \in T: P(x, y)]$

which is found to be true or false for each value of x.

By iterating through all the values of x, the truth of the statement:

$\forall x \in S: Q(x)$

can be tested which gives the truth value of:

$\forall x \in S \; \forall y \in T: P(x, y)$

The method extends to a statement with several quantifiers, by working successively from the innermost quantifier outwards.

Dubinsky and his colleagues theorise that students may similarly handle a multi-quantified statement by working from the inner statement outwards. Each application of a quantifier turns a predicate $P(x)$ into a statement. Dubinsky *et al* (1988) regard a predicate $P(x)$ as a process (for variable x) and a quantified statement – $\forall x : P(x)$ or $\exists x: P(x)$ – as an object. This relates the quantification of a predicate to the cognitive notion of encapsulating a process as an object, which they regard as of fundamental importance in cognitive development:

> A major cognitive skill (or act of intelligence) that we feel is required here is the ability to move back and forth between an internal process and its encapsulation as an object. (Dubinsky *et al.*, 1988, p. 48)

Successive application of quantifiers from the innermost quantifier outwards, in order to determine the truth of a quantified statement, gives an inductive method of steady reduction of complexity. However, even though the logic is evident, the cognitive complexity of this process of encapsulation for a multi-quantified statement is enormous. We should therefore ask whether students do follow this attractive method of working from the innermost quantifier out, at each stage going through 'all' the values of a variable to test the truth of a multi-quantified statement.

Mathematicians try various ways to enable students to grasp the multi-quantified definition of convergence of a sequence. For instance, instead of attempting to mentally go through 'all' the values of the variables, a more usual alternative is to consider each variable as having a given generic property (such as $n \geq N$) and using this property to make sense of the full statement.

Another approach, used in both ISETL and in non-computer courses, is to begin with numerical examples to give meaning to the definition. For instance, one might consider a given sequence, (say $a_n = 1/n$) and a specific value of ε (say ε = 1/1000), and then seek a value of N (say N = 1001) such that if $n \geq N$, then $a_n \leq \varepsilon$. Such an approach often gives the student a 'sense' of the definition in numerical terms, but it may not give a full picture of the meaning of the definition itself.

In our research, we find students who can write down the definition and perform numerical computations of the type given above, but cannot cope when there is a mixture of specific and general values. For example, given that a (general) sequence (a_n) converges to the (specific) value 1, to prove that after a certain value of N, the terms satisfy $a_n > 3/4$ for $n \geq N$. The problem here is that there are some numerical values given, but not enough to perform an actual calculation. The student may be able to handle the definition with all-numerical examples, but not be able to cope with the general definition.

Furthermore, the sequence in which attention is focused on the quantifiers is affected by the way in which it is presented in the curriculum. A common approach to dealing with the definition of limit is first to fix ε > 0, and then to focus on the two inner quantifiers:

$$\exists N \, \forall n \geq N: |a_n - L| < \varepsilon$$

This reduces the complexity, so that the student is seeking (for a specific value of ε), a value of N for which $|a_n - L| < \varepsilon$ whenever $n \geq N$. When this has been achieved, the initial ε is then allowed to vary. In our case study, Chris used such a broad general procedure to grasp the meaning of the definition. In this approach, the order of consideration of the quantifiers is N, n, ε, which is neither 'inside-to outside' nor is it the left-to-right order in which the quantifiers are read.

Individuals simply do not always read and write quantifiers in the order that seems evident in the written statement. In an earlier study (Bills and Tall, 1998), an able student Lucy wrote the definition of continuity of a function f in an interval $[a, b]$ in the following order (1), (2), (3):

$$\underbrace{\forall x, x_0 \in [a,b]}_{(2)} \;\; \underbrace{\forall \varepsilon > 0 \; \exists \delta > 0 \text{ s.t.}}_{(1)} \underbrace{|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon}_{(3)}$$

She explains that she does not remember the definition by rote, but reconstructs it from its constituent parts:

> I write down everything and say, "no, that's wrong", and then I work backwards. (p. 107)

Our studies of students coming to term with proof (Pinto, 1998; Pinto and Tall, 1999) reveal a range of very different approaches. Some appear to construct meaning by successive quantification from the inner quantifier outwards, as suggested by Dubinsky and colleagues. Others, such as Chris, start from mental images and refine them to provide a sufficiently generic image that can be used to build up the formal statement of the definition.

Thought experiment and propositional proof

Given a range of different ways of building up the notion of proof, our major challenge is to build a theoretical framework that can encompass the range of student approaches. We draw from several distinct sources of which two are seminal. One is rooted in the classical difference between *analogical* and *propositional* mental representations (see, for example, Eysenck and Keane, 2000). *Analogical representations* are typically considered as representing things simultaneously, in an implicit manner, without separate symbols to relate various aspects. They have loose rules of combination of information and are tied to specific instances that are being represented. They may involve sensory images or mental thought experiments using visual or spatial configurations. *Propositional representations* are verbal or symbolic, with relationships established by explicit rules of combination.

We contend that these distinctions can lead to different forms of proof: *thought experiment* and *propositional proof*. A thought experiment is a natural way of thinking for *homo sapiens* and does not require any formal knowledge. To satisfy ourselves of the consequences of taking a certain position, we simply imagine a situation in which the assumptions hold and think through possibilities.

For instance, in arithmetic, to 'see' that multiplication of whole numbers is commutative, one might draw an array of objects, say 3 rows of 4, and switch to seeing it as 4 columns of 3 objects. The total number of objects is 3 x 4 in one case and 4 x 3 in the other, different ways of seeing the same objects whose number 3 x 4 is therefore equal to 4 x 3.

This thought experiment has a vital ingredient. Although it uses specific numbers (3 and 4), it would be just as easy to perform the same experiment with any other pair of numbers (say 2 and 5, or 7 and 3). The experiment is a *prototype* for *any* pair of whole numbers. The individual sees the general argument embodied in the specific case. Such an example is called *generic* (Mason and Pimm, 1984); it uses a specific example to represent a whole class.

Classical Euclidean geometry is another theory that builds using thought experiments. Consider, for example, the theorem that an isosceles triangle has two equal angles. Any picture drawn of an isosceles triangle standing on its third side makes it seem patently obvious that the base angles are equal. But can we be sure that this will always be so? We might try various thought experiments. For instance, what happens if we vary the position of the vertex A of an isosceles triangle ABC where $AB=AC$? We may see that, as one side becomes longer than the other, the angle opposite the longer side becomes larger than the angle opposite the shorter side. At the balance point, when the two sides are equal, we can then 'see' that the angles are equal (Figure 1).

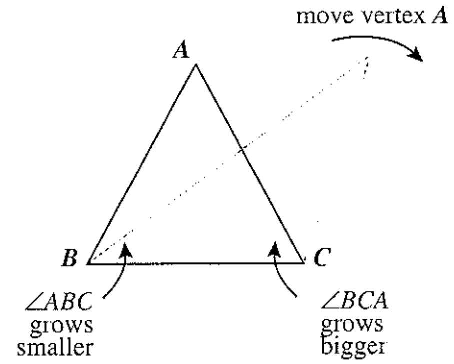


Figure 1 A thought experiment moving the vertex of an isosceles triangle

Another method might involve imagining the isosceles triangle being folded over the axis joining the vertex to the mid-point of the base. Because the two triangular halves have corresponding sides equal, the two halves will exactly match, lying one on the other, and so the angles B and C *must* be equal.

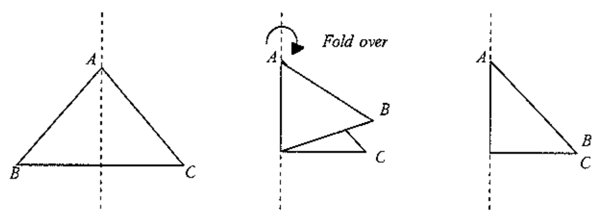


Figure 2 A thought experiment folding an isosceles triangle

Both of these thought experiments give profound intuition that the theorem is true. The first is an experiment that can be performed with the software *Cabri-Géomètre*, enabling the individual to *sense* that the theorem is true. The second can be performed physically by folding a piece of paper in the shape of an isosceles triangle, allowing the individual to *see* and feel that the theorem is true.

Mathematicians, however, feel more secure with an argument if it is translated into a verbal-symbolic form. Both of these thought experiments can be translated into Euclidean proofs. The second happens to be easier. It suggests a construction joining A to the mid-point M of BC and then proving that triangles ABM and ACM are congruent (three corresponding sides). Hence, by the theory of congruence, the angles B and C are equal.

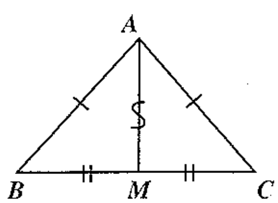


Figure 3 Two congruent triangles

This reveals Euclidean geometry as *a verbalised form of a thought experiment*. It involves imagining and drawing constructions by adding lines where necessary to reveal congruent triangles The constructions are described verbally and the congruency specified verbally, revealing equalities between other corresponding sides and angles. In this way, geometric constructions can be built from concept images of geometrical figures and translated into a *deductive verbal form*. In Euclidean geometry, we move from *thought experiment* to *verbal proof*

This translation from visual to verbal suggests a possible method of moving from visual mathematics to formal mathematics. What is required is the ability to see the general in the particular images, in order to *give meaning* to the corresponding formal definition, and to use the resulting links between imagery and formalism to formulate and prove theorems The more difficult part of this translation is the transfer from intuitive imagery to formal deduction.

'Natural' and 'alien' thinkers

Our second major theoretical influence is the difference between contrasting ways in which individuals react to a given context. Duffin and Simpson (1993, 1994, 1995, 2000) found that their own individual thinking styles proceeded in quite different ways, one of which they termed *natural* and the other, *alien*

> A 'natural' learner always attempts to 'make sense' of experiences by connecting them immediately to existing mental structures, looking for explanations and reasons based on those connections. An 'alien' learner, on the other hand, is willing to accept new experiences in an 'alien' way, building up isolated structures which deal with just those experiences, only constructing understanding out of connection made much later between these often mature, but isolated, mental schemas as a result of conflict. (Simpson, 1995, p 42)

The transition to formal mathematical thinking is accomplished in different ways by natural and alien learners The natural learner is more likely to use current concept images to attempt to build on them or modify them through *thought experiment*. The alien learner accepts the 'rules of the game' and plays them for what they are, with little attempt to connect them to previous experience.

For the alien learner, success follows through building up a coherent theory of proven theorems and techniques for handling the formal concepts. Regrettably, many students following this direction fail to cope with the complexity of the quantified statements and are unsuccessful in building the required formal theory.

The natural learner, on the other hand, builds on a variety of previous experiences - including enactive, visual and symbolic representations - to attempt to give meaning to the formalism. For some, this may be successful but others may find that their thought experiments are so powerful that they 'believe' the theorems to be true without any need for a formal proof. For instance, mental imagery of the limit concept may make it seem absolutely clear that "if a_n gets close to a and b_n gets close to b, then $a_n + b_n$ must get close to $a + b$". Such a 'truth', for them, requires no further proof, especially when that proof involves a complex machinery of definition and deduction that they fail to understand.

The road to success from analogical thinking to deductive formalism is sometimes a difficult one The imagery may occur as a simultaneous *gestalt* that is difficult to translate into a logical sequence. There are also subtle differences between intuitive imagery and formal deductions, differences that can prove very difficult to identify and rationalise. However, in the case of Chris, we have a student who thinks in a natural way but is able to transform his natural thinking into a formal presentation.

Object-based and process-based cognitive construction

In our discussion so far, we are beginning to distinguish between object-focused theory building - where the student physically and mentally manipulates objects to build up an understanding of their properties and the relationships between them - and process-based constructions involving the encapsulation of processes as objects. Such a distinction has a long pedigree in cognitive psychology Piaget distinguished clearly between the child's construction of meaning through *empirical abstraction* (focusing on objects and their properties) and *pseudo-empirical abstraction* (focusing on actions on objects) Later *reflective abstraction* occurs through mental actions on mental concepts This allows the mental operations themselves to become new objects of thought that may then be acted upon in more sophisticated theories (Piaget, 1972, p. 70).

Dubinsky (1991) took this view of reflective abstraction as a basic tenet of the APOS theory that he developed with his colleagues. This hypothesises that *actions* (consisting of a sequence of successive activities) become interiorised as *processes* conceived as a whole without performing them step-by-step, to be encapsulated as *objects*. These are then manipulated mentally as entities in a newly-developing mental *schema* (Asiala *et al.*, 1997)

Confrey and Costa (1996) criticised the single focus on process-object construction in advanced mathematical thinking, claiming a method of construction where:

> rather than demanding departure from activity, the act of seeing similarities in structure across different contexts would be the basis for abstraction. (p. 163)

This abstraction of similarities from a range of contexts has a well-established pedigree. Skemp (1971) proposed it to be the essential method of construction of higher-order concepts

These two views of mental growth focus on two distinct ways of building knowledge. In each case, there is a complex context involving the individual's reflection on her or his actions and perceptions For Dubinsky and colleagues, the theory grows from a focus on actions, which become interiorised as processes, then encapsulated as mental objects For Skemp, the theory is developed from a focus on the concepts, their properties and relationships.

Sfard (1991) provided a theory that resonates with both of these forms of mental growth. She postulated an *operational approach* focusing on mathematical processes (which later are 'reified' as mathematical objects) and then a structural approach focusing on the mathematical objects and their properties and relationships. She also proposed that operational activities invariably precede the development of structural theories (Sfard and Linchevski, 1994).

These two strands appear widely in the literature. In later developments of APOS theory (e.g. Czarnocha *et al.*, 1999), the methods of object construction are seen to occur not just by process-object construction (through the A-P-O sequence), but also by encapsulating *schemas* as objects (which we term S-O). The strict sequence of construction was formulated in a more flexible manner:

> although something like a procession can be discerned, it often appears more like a dialectic in which not only is there a partial development at one level, passage to the next level, returning to the previous and going back and forth, but also the development of each level influences both developments at higher and lower levels. (Czarnocha *et al.*, 1999, p. 98)

Tall (1999) suggested that this APOS scheme could be seen as having two parts: A-P-O and O-S. The second part (O-S) places an object within a broader schema, giving it greater meaning and applicability. For Dubinsky and colleagues, the object O in the O-S sequence must be built either through A-P-O or O-S encapsulation following an A-P-O encapsulation. Tall proposed that this view contradicted the manifest evidence that the brain is highly devoted to perceiving objects directly, tracing their edges, their movement and considering them as distinct entities. Therefore, he claimed, the brain was naturally constructed for direct perception of *objects*. In this light, the O-S construction could occur in the form of thought experiments, in which the individual imagines objects and investigates their properties and relationships to form a wider schema.

A more detailed analysis of these ideas is given in Tall *et al.* (1999). This takes a broader view, in which the theoretical ideas build from the fundamental abilities of *homo sapiens* to *perceive*, *act*, and *reflect*. The mental constructs based on *perception* correspond to empirical abstraction. Those based on *action* correspond to pseudo-empirical abstraction. Those based on *reflection* correspond to reflective abstraction. However, in the latter case, we do not believe that reflection only consists of acting on objects in order to routinise actions as processes and to encapsulate them as objects. The mental activity of reflection involves a variety of mental objects, actions, properties, relationships, theorems, and so on. In mathematics, it leads not only to the encapsulation of processes, such as counting, into mental objects, such as number; it also focuses on mental objects and their properties, using thought experiments, as in Euclidean geometry.

Tall (1995) observed that elementary mathematics involves both empirical abstraction in geometry and pseudo-empirical abstraction in counting and arithmetic, using symbols to represent the embodiment of actions interiorised as mental processes. The move from elementary to advanced (or formal) mathematics was characterised there in terms of shifting attention from the objects and actions/processes towards useful *properties* that are formulated as axioms in a propositional theory, one that deduces new properties by deduction from axioms and definitions.

However, the manner in which this shift is accomplished was not specified. Pinto (1998) provided empirical evidence that the transition to formal proof can successfully occur in different ways, one broadly following the encapsulation theories of Dubinsky and Sfard, another building and refining mental objects as in this case study involving Chris.

Refining properties of objects, rather than encapsulating processes is a central tenet of other theories. Wilensky (1991) writes:

> The more connections we make between an object and other objects, the more concrete it becomes for us. The richer the set of representations of the object, the more ways we have of interacting with it, the more concrete it is for us. Concreteness, then, is that property which measures the degree of our relatedness to the object (the richness of our representations, interactions, connections with the object), how close we are to it, or, if you will, the *quality of our relationship* with the object (p. 198)

Our case study highlights an individual who refines his own quality of relationship with a mental object to represent and translate his images and actions into the notion of convergence of a sequence. Such a process builds ideas that for him are, in this sense of Wilensky, concrete.

Conclusion

This article unfolds qualitative aspects and strategies of an individual's construction of knowledge. It presents data showing that the symbolic-propositional model generally adopted to understand learners' constructions of quantified statements (see, for example, Dubinsky *et al.*, 1988) does not capture the development of a successful 'natural' route of learning, such as the one pursued by the student Chris. Rather than constructing a concept image from a defined object, through abstracting from 'actions on objects', this successful natural learner exhibits how a defined object may be consistently understood by just reconstructing it from the concept image. For such a student, there seems to be no harm if (s)he still understands the concept definition as a description, provided that the property, or properties, which are 'describing' the object, must also *characterise* it completely. In particular, this study expands the theory of 'natural' learning presented in Duffin and Simpson (1993, 1994, 1995, 2000).

An alternative explanation that fits the 'process-object' sequence of development could be proposed, as follows. The student has an image of the process of convergence. He carries out *cognitive* processes on this image to refine and encapsulate it into a formal symbolic definition. This is a way of describing the development that seems to be consonant with the A-P-O part of APOS theory. The lecturer imposes a construct on him (the notion of convergence of a null-sequence) which he sees as being *external*. This is an *action* conception of convergence. He carries out his own cognitive processes (looking up alternative definitions, carrying out thought experiments) to make his own internal construct. This *interiorises* the action as a *process*. He then *encapsulates* his experiences into the mental *object* called 'limit'.

We have no quarrel with such a description. Indeed, Pinto (1998) reports students attempting to follow an APOS sequence of *extracting meaning* from the formal definition by formal deduction. What we find problematic is the beginning of a learning sequence on proof which begins only with externally-imposed actions, without attempting to get some kind of global grasp of what is going on at the outset.

There is an alternative route, one used by Chris in this case study It begins with an embodied representation that already includes the process of convergence and the concept of limit approached dynamically. Chris was able to take this concept image and use thought experiments to transform it into a formal approach.

Curriculum frameworks which are mainly concerned with 'encapsulation of processes into objects' as the basis for extracting meaning from a definition do not capture the whole complexity of the cognitive demands of a learner who builds his or her internal structures through *giving meaning* to definitions using thought experiments This global alternative, which sees the larger informal picture and reconstructs it as a symbolic sequential theory, is not only viable, it is also a 'natural' way for some students to learn.

Acknowledgement

The authors wish to thank CNPq (Brazil) for generous partial support in the preparation of this article.

References

Asiala, M., Brown, A., DeVries, D., Dubinsky, E., Mathews D and Thomas, K (1996) 'A framework for research and curriculum development in undergraduate mathematics education', *Research in Collegiate Mathematics Education II*, CBMS 6, pp 1-32

Barnard, A. and Tall, D (1997) 'Cognitive units, connections and mathematical proof', in Pekhonen, E (ed), *Proceedings of the Twenty-First International Conference for the Psychology of Mathematics Education*, Lahti, Finland, **2**, pp 41-48.

Bills, L. and Tall, D. (1998) 'Operable definitions in advanced mathematics: the case of the least upper bound', *Proceedings of the Twenty-Second International Conference for the Psychology of Mathematics Education*, Stellenbosch, South Africa, **2**, 104-111

Confrey, J and Costa, S (1996) 'A critique of the selection of "mathematical objects" as a central metaphor for advanced mathematical thinking', *International Journal of Computers for Mathematical Learning* **1**(2), 139-168.

Cottrill, J, Dubinsky, E, Nichols, D., Schwingendorf, K, Thomas, K and Vidakovic, D (1996) 'Understanding the limit concept: beginning with a coordinated process schema', *Journal of Mathematical Behavior* **15**(2), 167-192.

Czarnocha, B, Dubinsky, E, Prabhu, V. and Vidakovic, D (1999) 'One theoretical perspective in undergraduate mathematics education research', in Zaslavsky, O. (ed), *Proceedings of the Twenty-Third International Conference for the Psychology of Mathematics Education*, Haifa, Israel, **1**, pp 95-110.

Dubinsky, E (1986) 'A theoretical perspective for research in learning mathematics concepts: genetic decomposition and groups' (Informal summary of presentation)

Dubinsky, E (1991) 'Reflective abstraction in advanced mathematical thinking', in Tall, D. (ed) *Advanced Mathematical Thinking*, Dordrecht, Kluwer, pp. 95-123

Dubinsky, E., Elterman, F and Gong, C (1988) 'The student's construction of quantification', *For the Learning of Mathematics* **8**(2), 44-51

Duffin, J and Simpson, A (1993) 'Natural, conflicting and alien', *Journal of Mathematical Behaviour* **12**(4), 313-328

Duffin, J and Simpson, A (1994) 'Towards a theory of learning', in Da Ponte, J (ed), *Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education*, Lisbon, Portugal, **2**, pp. 288-295

Duffin, J and Simpson, A (1995) 'Characteristics of understanding', *Proceedings of BCME-3*, Loughborough, Leicestershire, pp. 7-12.

Duffin, J and Simpson, A (2000) 'A Search for understanding', *Journal of Mathematical Behaviour* **18**(4), 1-13

Eysenck, M. and Keane, M (2000) *Cognitive Psychology: a Students Handbook* (fourth edition), Hove, Sussex, Psychology Press

Gray, E. and Tall, D (1994) 'Duality, ambiguity and flexibility: a "proceptual" view of simple arithmetic', *Journal for Research in Mathematics Education* **25**(2), 116-140

Hadamard, J (1945) *The Psychology of Invention in the Mathematical Field*, Princeton, NJ, Princeton University Press (reissued as Dover edition, New York, NY, 1954)

Kuyk, W. (1982) 'A neuropsychodynamical theory of mathematics learning', *For the Learning of Mathematics* **3**(1), 16-23.

Lakoff, G. and Johnson, M (1999) *Philosophy in the Flesh*, New York, NY, Basic Books.

Mac Lane, S (1994) 'Responses to theoretical mathematics', *Bulletin (new series) of the American Mathematical Society* **30**(2), 190-193.

Mason, J (1989) 'Mathematical abstraction as the result of a delicate shift of attention', *For the Learning of Mathematics* **9**(2), 2-8

Mason, J and Pimm, D (1984) 'Generic examples: seeing the general in the particular', *Educational Studies in Mathematics* **15**(3), 277-289.

Paivio, A (1971) *Imagery and Verbal Processing*, New York, NY, Holt, Rinehart and Winston

Paivio, A (1986) *Mental Representations: a Dual Coding Approach*, Oxford, Oxford University Press.

Piaget, J (1972) *The Principles of Genetic Epistemology* (trans Mays, W), London, Routledge and Kegan Paul

Pinto, M. (1998) *Students' Understanding of Real Analysis*, Unpublished Ph D Thesis, Coventry, Warwickshire, Warwick University

Pinto, M and Tall, D (1999) 'Student constructions of formal theory: giving and extracting meaning', in Zazlavsky, O (ed), *Proceedings of the Twenty-Third International Conference for the Psychology of Mathematics* Education, Haifa, Israel, **2**, pp 41-48

Poincaré, H (1913) *The Foundations of Science* (trans. Halstead, G), New York, NY, The Science Press (reissued in a University Press of America edition, Washington, DC, 1982)

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

Sfard, A (1994) 'Reification as the birth of metaphor', *For the Learning of Mathematics* **14**(1), 44-55.

Sfard, A and Linchevski, L. (1994) 'The gains and pitfalls of reification – the case of algebra', *Educational Studies in Mathematics* **26**(2/3), 191-228.

Simpson, A (1995) 'Developing a proving attitude', *Proceedings of the 'Justifying and Proving in School Mathematics' Conference*, London, Institute of Education, University of London, London, pp 39-46

Skemp, R (1971) *The Psychology of Learning Mathematics*, London, Penguin Books Ltd

Strauss, A (1987) *Qualitative Analysis for Social Scientists*, Cambridge, Cambridge University Press.

Strauss, A and Corbin, J. (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, London, Sage Publications

Tall, D (ed.) (1991) *Advanced Mathematical Thinking*, Dordrecht, Kluwer

Tall, D (1995) 'Mathematical growth in elementary and advanced mathematical thinking', in Carraher, D and Meira, L (eds), *Proceedings of the Nineteenth International Conference for the Psychology of Mathematics Education*, Recife, Brazil, **1**, pp. 61-76

Tall, D. (1999) 'Reflections on APOS theory in elementary and advanced mathematical thinking', in Zaslavsky, O (ed), *Proceedings of the Twenty-Third International Conference for the Psychology of Mathematics Education*, Haifa, Israel, **1**, pp 111-118

Tall, D, Thomas, M, Davis, G., Gray, E and Simpson, A (1999) 'What is the object of the encapsulation process?', *Journal of Mathematical Behavior* **18**(2), 223-241.

Wilensky, U (1991) 'Abstract meditations on the concrete and concrete implications for mathematics education', in Harel, I. and Papert, S. (eds), *Constructionism*, Norwood, NJ, Ablex Publishing Corporation, pp 192-203