The Role and Function of a Hierarchical Classification of Quadrilaterals [1]

MICHAEL DE VILLIERS

Introduction
A well-known and useful distinction between different types of understanding in mathematics is that of distinguishing between instrumental, relational, and logical understanding [e.g. Skemp, 1976; 1977, Byers & Herscovics, 1977]. Where instrumental “understanding” (the author actually prefers the term “proficiency”) refers to the ability of an individual to correctly and efficiently manipulate mathematical content (e.g. by using algorithms, rules and definitions), relational and logical understanding respectively refer to understanding the conceptual relationships between content and the underlying logic upon which these relationships are based.

A serious deficiency in this model is that no provision is really made for functional understanding, in other words, understanding the role, function, or value of specific mathematical content or of a particular process [compare Human, 1989]. Extensive experience with children in interview and classroom contexts seems to indicate that many of their problems with mathematical content and processes often do not lie so much with poor instrumental proficiency nor inadequate relational or logical understanding as in a poor understanding of the usefulness or function thereof. It should be noted that this functionality is not confined to applications to the real world outside of mathematics but includes the relative values or functions of content and processes within mathematics.

To a very large extent, it seems that the absence, presence, or level of an individual’s functional understanding determines that individual’s motivation to study and learn mathematics. Without functional understanding, mathematics simply degenerates into a useless, meaningless and arbitrary subject, demotivating the learner from attempting to learn and explore it. The adequate development of functional understanding is therefore an important criterion for evaluating any teaching approach.

In this article different types of classification will be distinguished, as well as a theoretical analysis done of the role and function of hierarchical classification in mathematics. Lastly, some brief comments regarding the teaching of a hierarchical classification of quadrilaterals will be made.

Interlude
The following extract is a fairly typical example of several interviews and experiences with children from Standard 7 to Standard 10 (Grades 9 to 12) over the past number of years [see De Villiers, 1987, 1990]:

I: If we define a parallelogram as any quadrilateral with opposite sides parallel, can we then say that a rectangle is a parallelogram?
S: Yes, ... because a rectangle also has opposite sides parallel ... But I don't like this definition of parallelograms... I know we are taught this definition at school and that squares and rectangles are parallelograms (pulls face), but I don't like it.
I: How would you define parallelograms instead?
S: As any quadrilateral with opposite sides parallel, but not all angles equal.
I: What about rhombi then? Would you say a rhombus is a parallelogram?
S: Hmm ... according to my definition, yes ... but I don't like that either ... I would therefore rather say a parallelogram is a quadrilateral with opposite sides parallel, but not all angles or sides equal.

Clearly this student has no problem with drawing correct conclusions from definitions and making hierarchical class inclusions but prefers not to do so. Furthermore, this student clearly exhibits the ability to formulate a definition. Clements & Battista [1992 63] have similarly reported two cases of students who were able to follow the logic of a hierarchical classification of squares and rectangles but had difficulty accepting it. The problem therefore seems to be not so much a lack of relational or logical understanding, or even of proficiency in defining, but one of a lack of functional understanding (i.e. what is the function or value of hierarchical classification of quadrilaterals).

Partition and hierarchical classification
By the term hierarchical classification is meant here the classification of a set of concepts in such a manner that the more particular concepts form subsets of the more general concepts. Several examples like the classification of the real numbers or the classification of various geometries from a transformation perspective (Erlangen program) can be provided, but for the purpose of this article we shall mainly focus on the classification of quadrilaterals.

In contrast to a hierarchical classification there also exists the possibility of a partition classification of concepts. In such a classification however the various subsets of concepts are considered to be disjoint from one another. For example, in Figure 1 a hierarchical classification of parallelograms, rectangles, rhombi and squares is contrasted with a partition classification (Two different types of representation for each classification are illustrated.) In the hierarchical classification we can clearly see that the rect-
angles and rhombi are *subsets* of the parallelograms, with the squares as the intersection between rectangles and rhombi. In contrast, in the partition classification squares are not rectangles or rhombi, nor the rectangles and rhombi parallelograms.

**The relationship between classifying and defining**
The classification of any set of concepts does not take place independently of the process of defining. For example, to hierarchically classify a parallelogram as a trapezium requires defining a trapezium as "a quadrilateral with at least one pair of opposite sides parallel." If on the other hand we want to exclude the parallelograms from the trapeziums we need to define a trapezium as a "quadrilateral with only one pair of opposite sides parallel."

Furthermore, it should be unequivocally stressed that a partition definition (and classification) is not mathematically "wrong" simply because it is partial (provided of course it contains sufficient information to ensure that all non-examples are excluded). For example, the partition definition for parallelograms given earlier by the student (i.e. a quadrilateral with opposite sides parallel, but not all angles or sides equal may be unconventional, but it is definitely not wrong. In fact it is a *correct economical* definition as it contains only necessary and sufficient properties. Of course, just as students often provide hierarchical definitions which are correct but uneconomical (i.e. containing superfluous information) the author has frequently also found them giving *correct, uneconomical* partition definitions like the following:

A parallelogram is a quadrilateral with opposite sides equal and parallel, opposite angles equal, diagonals of different length halving each other, but not perpendicularly.

(It is perhaps necessary to point out that even mathematicians do not always strictly adhere to economy of definitions and axioms. For example, in the definition of a group only a left inverse is really required as the right inverse is implied by it, but normally we simply state that there should be inverses for all elements. The reason for this is simply one of *convenience*, i.e. to avoid an extra, somewhat complicated proof. Similarly, it is customary to use five axioms for Boolean Algebra, although in fact only three axioms are necessary.)

Sometimes a partition classification and its corresponding definitions are useful and necessary to clearly distinguish between concepts. For example, consider the partition classification of convex, concave, and crossed quadrilaterals shown in Figure 2 with the following possible definitions:

![Hierarchical and Partitional Classification Diagrams](image-url)
A quadrilateral is any closed four-sided figure in the plane with four vertices.
A simple closed quadrilateral is a quadrilateral with sides only meeting at the vertices.
A crossed quadrilateral is a quadrilateral with two of the sides also crossing each other at a point other than the vertices.
A convex quadrilateral is a simple closed quadrilateral with none of its angles reflexive.
A concave quadrilateral is a simple closed quadrilateral with one of its angles reflexive.

Similarly, it is useful and necessary to partition the kites into convex and concave ones. Furthermore, when classifying and specific given quadrilateral, partitioning is a spontaneous and natural strategy. For example, we would not normally say when we have a square in front of us: aha! here we have a rectangle. Instead we would normally call a square a "square" and reserve the term "rectangle" only for a non-square (or general) rectangle. Similarly, we would normally employ the term "rhombus" only when facing a non-square (or general) rhombus. In precisely the same fashion, Dennis [1978] made use of partitioning to specify a computer programme for the classification of quadrilaterals (according to given coordinates).

In fact, partitioning is a generally accepted mathematical method in many areas of mathematics, but particularly in the study of topological surfaces and spaces where the fundamental problem is the subdivision of these surfaces and spaces in different disjoint types [e.g. see Patterson, 1956]. Furthermore, since a classification and its corresponding definitions are arbitrary and not absolute, we should acknowledge that the choice between a hierarchical and a partition classification is often a matter of personal choice and convenience. The author for instance recently came across the following partition definition in an old geometry textbook by Wentworth [1881:58] which was widely used in American colleges and universities during the previous century: A rhombus is a parallelogram which has its sides equal, but its angles oblique angles.

The fundamental question addressed later on in this paper is, therefore: why do we (conventionally) prefer a hierarchical classification of the various convex quadrilaterals rather than a partition classification? Or phrased differently, what advantages does a hierarchical classification in this instance have over a partition classification?

**Descriptive and constructive classification**

Analogous to similar distinctions for the process of axiomatization and defining [e.g. see Krygowska, 1971; Human, 1978; De Villiers, 1986], it is also possible to distinguish between two essentially different types of classification, namely descriptive (a posteriori) or constructive (a priori) classification, each of which can be either hierarchical or partitional.

In contrast, by a priori classification is meant here that the mathematical processes of generalization and special-
Generalization or specialization of course does not necessarily take place from only one concept, but can involve two or more concepts. For example, a new concept C may be generalized from two or more concepts by selecting one or more appropriate common properties (constraints) from these concepts. Similarly, a new concept C may also be specialized from two or more concepts by demanding that it combines all the properties (constraints) of these concepts. In general the most important function of an a priori classification is therefore clearly the discovery/creation of new concepts.

Let us now briefly look at some examples of a posteriori and a priori classification with regard to the quadrilaterals. An a posteriori classification would occur, for example, if the classification of the squares and rectangles was to be considered after they have already been known for some time, and their properties have been thoroughly examined.

On the other hand, with a priori classification we could start with the most special concept, a square, and generalize the rectangle and parallelogram consecutively as new concepts, as shown in Figure 3. For example, the rectangle can be generalized from the square by relaxing the requirement that all sides must be equal, but still retaining the property of equal angles. Similarly, the parallelogram can be generalized from the rectangle by relaxing the requirement that all angles must be equal, but still retaining the property of opposite sides parallel. In the same manner we can generalize via a rhombus to a parallelogram.

Or vice-versa, by starting from the more general concept, a parallelogram, we can specialize by imposing more and more properties to eventually produce a square. For example, the rhombus can be specialized from the parallelogram by requiring the additional property of equal sides. Similarly, the square can be specialized from the rhombus by requiring the additional property of equal angles (in other words, combining all the properties of rectangles and rhombi). It is however important to again emphasize that the generalization or specialization need not be hierarchical but could theoretically be partitional (although in actual practice this may be the exception rather than the rule).

Similarly we can generalize the concept kite to a new concept, say for example a perpendicular quadrilateral, by relaxing the conditions that two pairs of adjacent sides have to be equal but retaining the perpendicularity of the diagonals (see Figure 4). (Note that we can also get concave and crossed perpendicular quadrilaterals. An interesting property of perpendicular quadrilaterals is that if we connect the midpoints of adjacent sides we obtain a rectangle.) A hierarchical definition of a perpendicular quadrilateral would now simply be that it is a quadrilateral with perpendicular diagonals. In contrast, a partition definition of it would have to exclude the kites by adding that two pairs of adjacent sides may not be equal.

We can also specialize the concepts cyclic quadrilateral and (convex) kite to produce a new concept, say a right kite, by demanding that it is their intersection (i.e. has the properties of both) (see Figure 5). As before, one would now have to add further conditions to the cyclic quadrilaterals (i.e. two pairs of adjacent sides may not be equal) and the kites (i.e. may not be cyclic) if one wanted to exclude (partition) the right kites from the cyclic quadrilaterals and kites.

![Figure 4](image1)

![Figure 5](image2)
Some important functions of hierarchical classification
This brings us finally to the major focus of this paper, namely what is the value or function of hierarchical classification? Some of the most important functions are:

- it leads to more economical definitions of concepts and formulation of theorems
- it simplifies the deductive systematization and derivation of the properties of more special concepts
- it often provides a useful conceptual schema during problem solving
- it sometimes suggests alternative definitions and new propositions
- it provides a useful global perspective

Each of these will now be discussed in more detail

Economical definitions and formulations of theorems
Economy of definition and of formulation of theorems is probably one of the most important advantages of a hierarchical classification. As we have already seen earlier with the parallelograms, a hierarchical definition is shorter than a partitional one which has to include additional properties to exclude the rhombi, squares and rectangles. For another example consider a partition definition for an isosceles trapezium (see Figure 6).

A hierarchical definition which includes rectangles (and squares) as special cases would, for example, be to say that it is any quadrilateral with an axis of symmetry through (at least) one pair of opposite sides. (Note that it is then necessary to partition the isosceles trapezia into convex and crossed ones.) A partition definition on the other hand, which excludes the rectangles and squares, would have to include the additional condition that it may not have a right angle.

A partition classification also often makes the formulation of certain theorems clumsy and cumbersome. Consider, for example, the following two formulations of well-known results from a partition perspective:

If the midpoints E, F, G and H of the sides of any quadrilateral ABCD are consecutively connected, then EFGH is a parallelogram, rectangle, rhombus or square.

The exterior angle of a cyclic quadrilateral, isosceles trapezium, right kite, rectangle or square is equal to the opposite interior angle.

Simplification of deductive systematization
By classifying (defining) a concept A as a subset (special case) of a concept B, it becomes unnecessary to repeat any of the proofs of the properties of concept B for concept A as they are automatically implied for A by the hierarchical inclusion. For example, by classifying a rhombus as a kite, all the theorems which have already been proved for kites are immediately made applicable to the rhombi (and squares). In other words, it is unnecessary to prove, for instance, that the diagonals of a rhombus (and square) are perpendicular, since this is an easily proved property of the kites.

In contrast, if the rhombi (and squares) were to be excluded from the kites, one would strictly speaking have to again prove that the above property is also true from the chosen definition for rhombi (and squares), whatever it may be. Apart from an economy of definition or formulation, a hierarchical classification therefore also results in an economical deductive system.

A useful conceptual schema during problem solving
A hierarchical class inclusion is also often useful during problem solving; in particular, for proving riders. For example, suppose one wants to prove that a kite with one pair of opposite sides parallel is a rhombus. Using the hierarchical perspective that the rhombi are the intersection of the kites and parallelograms, it is sufficient therefore to prove that the figure is a parallelogram, since any kite with both pairs of opposite sides parallel must be a rhombus.

Another particularly illustrative example involves Von Aubel's theorem and a special case of it. Von Aubel's theorem states that if squares are constructed on the sides of any quadrilateral, then the line segments connecting the centers of opposite squares are equal and perpendicular. [A proof is given in Yaglom, 1962: 95-96] An interesting special case is that if squares are constructed on the sides
of a parallelogram, then the centers of these squares also form a square (see Figure 7). Although there are many different ways of proving this special case, an elegant way which utilizes hierarchical classification is simply to show that quadrilateral EFGH is a parallelogram, since a square is the only parallelogram with equal and perpendicular diagonals (the latter follows directly from Von Aubel).

**Figure 8**

**Alternative definitions and new propositions**

Consideration of a hierarchical relationship between concepts may sometimes suggest alternative definitions and new propositions. If, for instance, concept A is the intersection of two other concepts B and C, then it must obviously possess all the properties of both concepts B and C. By now considering various subsets of the total set of properties of concept A, alternative definitions for it, or new propositions, may be suggested.

For example, since any isosceles trapezium is cyclic, the isosceles trapezia may be considered as the intersection between the trapezia and the cyclic quadrilaterals. It therefore suggests immediately that a cyclic quadrilateral with at least one pair of opposite sides parallel would be an isosceles trapezium (see Figure 8a).

Similarly, since the diagonals of an isosceles trapezium are equal, the following alternative definition (or proposition) for isosceles trapezia is suggested:

An isosceles trapezium is a cyclic quadrilateral with equal diagonals (see Figure 8b).

The keeping in mind of a hierarchical classification can also sometimes enable the generalization of certain results. Suppose, for example, we accidentally discovered by experimentation that if we connected the centers of squares on the sides of any triangle to the opposite vertices of the triangle, then these three line segments are concurrent (Figure 9a). Since all squares are similar, and special rectangles, one might now conjecture that the same result should hold for similar rectangles, as shown in Figure 9b. (A proof of the result on which this is based, and a further generalization is given in De Villiers [1989].) Similarly, in Figure 7, connecting the centers (or other corresponding points) of (any) similar figures on the sides of the base parallelogram, would produce a parallelogram.

**A useful global perspective**

A hierarchical classification provides a useful global perspective which may lead to a more cohesive perspective on the underlying relationships between concepts, and therefore also to better retention. In addition, it is aesthetically pleasing and insightful to see how the various intersections between more general concepts produce the properties of more special concepts.

For example, since the rhombi are the intersection between the kites and parallelograms, it immediately follows from the diagonal properties of kites and parallelograms that the diagonals of a rhombus would perpendicularly bisect each other.

Similarly, since the rectangles are the intersection between the parallelograms and isosceles trapezia, it immediately follows that a rectangle must have opposite
angles equal (parallelogram property) as well as adjacent angles equal (isosceles trapezium property), from which we get the familiar property that all its angles are equal. In the same way, the rectangles inherit equal diagonals from the isosceles trapeziums, as well as bisecting ones from the parallelograms.

**Some brief comments regarding the teaching of a hierarchical classification of the quadrilaterals**

Unfortunately many teachers and textbook authors still seem to hold a perspective that only the conventional hierarchical classification is mathematically acceptable, whereas a partition classification is mathematically illogical and therefore unacceptable. However, as pointed out in this paper, a partition classification is equally acceptable and a frequently employed method in mathematics. The only reason for the conventional preference for a hierarchical classification of quadrilaterals lies in its greater functionality, as outlined earlier. Most textbooks and teachers however completely ignore discussing this fundamental aspect, simply imposing a hierarchical classification and definitions on students for which they have little or no functional understanding.

Many studies on the Van Hiele theory over the past number of years have clearly shown that many students have problems with the hierarchical classification of quadrilaterals [e.g. Mayberry, 1981; Usiskin 1982; Burger & Shaughnessy, 1986; Fuys, Geddes & Tischler, 1988] Research by the author and several of his students [e.g. Malan, 1986; De Villiers & Njisane, 1987; Smith, 1989; De Villiers, 1987, 1990] has further indicated that several children's difficulty with hierarchical class inclusion (especially older children) does not necessarily lie with the logic of the inclusion as such, but often with the meaning of the activity, both linguistic and functional: linguistic in the sense of correctly interpreting the language used for class inclusions, and functional in the sense of understanding why it is more useful than a partition classification.

At Van Hiele Level 1 (Visualization) and Level 2 (Exploration) the use of computer micro-worlds such as Logo Geometry [e.g. see Battista & Clements, 1992], or dynamic software such as *Cabri* or *Sketchpad*, offer great potential for conceptually enabling many children to see and accept the possibility of hierarchical inclusions (for example, letting children construct a square with a rectangle procedure in Logo, or letting them drag the vertices of a dynamic parallelogram in *Cabri* or *Sketchpad* to transform it into a rectangle, rhombus or square).

For a hierarchical classification of quadrilaterals to be meaningful to students at Van Hiele Level 3 (ordering), it is however also essential that an appropriate negotiation of linguistic meaning should already have taken place. From interviews with children individually and in classroom contexts, the author has for example found that many had difficulty with the meaning of the word "is" in a statement like "a square is a rectangle". They seemed to interpret it as meaning a square "is equivalent to" or "the same as" a rectangle, and therefore (quite correctly) rejecting the statement as ridiculous or false. Using the adjective "special", for example: "a square is a special rectangle", helped many students realize that what is actually meant is that the one is a subset of the other. Reference to analogous everyday or other mathematical situations where an object may be viewed as a special set of a larger set and therefore having two different "names" (e.g. "a mammal is a vertebrate" and "a horse is a mammal and a vertebrate"), was also useful.

Secondly, it is absolutely vital at Van Hiele Level 3 that a negotiation of functional meaning also take place; that is, sufficient opportunity and appropriate activities should be given for discussing the value or function of a hierarchical classification. The author has, for example, found it very useful to allow student firstly to formulate, compare, and choose their own definitions and classification of the squares, rectangles, and rhombi; many of them spontaneously prefer partitioning. By consistently now challenging these students to continue formulating partition definitions for more and more general quadrilaterals, and comparing them with the hierarchical alternatives, they soon start realizing and appreciating the economy of the latter. Simultaneously insisting that they prove all the properties of the partitioned quadrilaterals, and asking them to critically compare their system of definitions with a deductive system based on a hierarchical classification, most students gradually see the convenience of a hierarchical inclusion and make a transition towards it.

The idea that students should not be given ready-made definitions and classifications, but that they should actively participate in the process of defining and classifying, and critically comparing the alternatives, is strongly supported by constructivist epistemology and learning theory. Instead of simply ignoring or dismissing children's partitioning of quadrilaterals, we should address it with much greater empathy, and acknowledge that their approach is a rational and meaningful attempt at sense-making. It is rather alarming to see so many teachers and even researchers simply paying lip service to constructivism (i.e. professing to acknowledge children's autonomy in learning and constructing mathematics, but when it comes to the classification of quadrilaterals this is not applied at all).

**Note**

[1] This paper was presented at PME 17, University of Tsukuba, Japan, 18-23 July 1993. Attendance at this conference was made possible by a grant from the Fund for Research Development (FRD) of the Human Sciences Research Council (HSRC), Pretoria, South Africa. The views expressed in this paper are those of the author and not necessarily those of the HSC.

**References**


The 20th century has produced an assessment practice in education that is dominated the world over by psychometrics, the measurement of psyche. The challenge for the 21st century, as far as mathematics education is concerned, is to produce an assessment practice that does more than measure a person's mind and then assign that mind a treatment. We need to understand how people, not apart from but embedded in their cultures, come to use mathematics in different social settings and how we can create a mathematics instruction that helps them use it better, more rewardingly, and more responsibly. To do that will require us to transcend the crippling visions of mind as hierarchy, school as machine, and assessment as engineering.

Jeremy Kilpatrick