Cite as: Gutiérrez, A.; Jaime, A. (1995): Towards the design of a standard test for the assessment of the student's reasoning in geometry, Proceedings of the 19th PME Conference (vol. 3, pp. 11-18) Recife, Brasil: PME.

# TOWARDS THE DESIGN OF A STANDARD TEST FOR THE ASSESSMENT OF THE STUDENTS' REASONING IN GEOMETRY

Angel Gutiérrez and Adela Jaime
Dpto. de Didáctica de la Matemática. Universidad de Valencia. Valencia (Spain)

**Abstract.** *In previous publications, we outlined a theoretical framework for the design of tests to evaluate students' Van Hiele level of thinking and for the assignment of Van Hiele levels to the students. Based on this framework, we present here a test aimed to assess students in Primary and Secondary Schools. The subject area of the test are polygons and other related concepts. The test is integrated by open-ended super-items, each one of them having several related questions dealing with the same problem. In this paper we describe the items, analyze the structure of the test, and present the results of the administration to a sample of 309 primary and secondary students.*

## Introduction.

A constant in the research on the Van Hiele Model of Geometric Reasoning over the years is the expressed need of an assessment instrument fitting the usual requirements of reliability and validity, and also fitting the requirement of easy administration in a short time to big samples. Unfortunately, the third requirement seems to be incompatible with the previous ones. Clinical interview is considered the most valid and reliable technique, but it can hardly be used with medium sized samples. Written tests do not have this inconvenience, but it is usually harder to verify their reliability and validity, since written answers are poorer than oral answers. Some attempts have been done in the direction of building written tests to assess the Van Hiele levels but, unfortunately, they have not been successful enough (Mayberry, 1981; Usiskin, 1982; Senk, 1983; and Crowley, 1989).

We are working in a research program, a part of which is presented in this paper, whose main objective is to build a written test with a structure as close as possible to semi-structured clinical interviews. The research is divided into three related parts:

- Definition of a model for the evaluation of tests and assignment of Van Hiele levels (Gutiérrez, Jaime, Fortuny, 1991 and Gutiérrez et al., 1991): The continuity of the Van Hiele levels has been showed by lots of students' answers. Then we proposed a method to assign to students a degree of acquisition of each level, mirroring the reality that most students use a level of thinking or another depending on the task they are solving.

- Identification of a framework for the design of tests to evaluate students' levels of thinking (Jaime, Gutiérrez 1994): The reasoning of each level is characterised by several key processes or abilities, so a balanced and valid test should assess everyone of them. On the other side, most tasks can be answered from several levels of thinking, so students' level of thinking is determined by their answers, not by the statements of the tasks. Then, researchers should consider the range of possible levels of answer to each item of a test.

Furthermore, this helps to obtain a reasonable amount of items assessing each Van Hiele level without the inconvenience of a long test taking too much time to be answered.

- Design of a test to assess the Van Hiele level of reasoning of students in Primary, Secondary, and University. In this paper we present results of this part of the research. Namely, we describe and analyze a test made of open-ended items based on polygons and other related concepts, and present the results of the administration of this test to students from 6th grade of Primary to the last grade of Secondary. Previous versions of this test were also administered to university students (preservice teachers).

To conclude this introduction, just to mention that we consider Van Hiele levels 1 (recognition), 2 (analysis), 3 (classification), and 4 (formal deduction).

## Description of the Test.

Each Van Hiele level is integrated by several key processes of thinking so, for a student to attain a certain level, the student should show mastery in all the processes characteristic of this level. These processes are (Jaime, Gutiérrez, 1994):

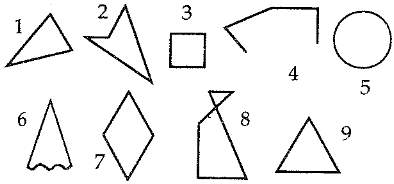
- **Identification** of the family a geometric object belongs to.
- Work with the **Definition** of a concept, in two ways: To *use* a known definition, and to *state* a definition for a class of geometric objects.
- **Classification** of geometrical objects into different families.
- **Proof** in some way of a property or statement.

The table below summarises the key processes characteristic of each Van Hiele level 1 to 4. The "---" mark means that this process is not a part of the reasoning of the level. Then, any test designed to assess the Van Hiele levels of thinking should have items evaluating every process of each level.

| | Identification | Definition | Classification | Proof |
|---|---|---|---|---|
| **Level 1** | √ | √ (State) | √ | --- |
| **Level 2** | √ | √ (Use & State) | √ | √ |
| **Level 3** | --- | √ (Use & State) | √ | √ |
| **Level 4** | --- | √ (Use & State) | --- | √ |

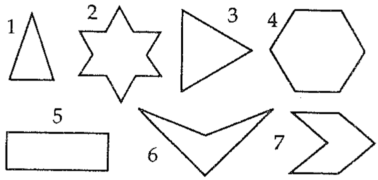
For several years, we have piloted and improved a set of paper and pencil items. Usually, pilot trials of the items were followed by clinical interviews of some students, to check the reliability of their written answers. The final result in the subject area of polygons is the test we describe below. It is integrated by 8 open-ended items, each one of them having several questions. When evaluating a student's answers, we do not consider each answer independently, but the answers to all the questions in the same item or section. In this way, we can have a clearer picture of the student's way of reasoning, and we can judge inconsistencies, contradictions, etc. among different answers. These are the items in the test (due to the limited space, the text of some items is shortened):

**Item 1.** - Write a **P** on the polygons, write an **N** on the non-polygons, write a **T** on the triangles, and write a **Q** on the quadrilaterals. If necessary, you may write several letters on each figure.




- Write the numbers of the figures which <u>are not polygons</u> and explain, for each of them, why it is not a polygon.
- The same questions for figures which <u>are triangles</u>, and figures which <u>are quadrilaterals</u>.
- Is figure 8 a polygon? Why? - Is figure 2 a triangle? Why?

**Item 2.** - Write an **R** on the regular polygons, an **I** on those that are irregular, a **V** on those that are concave, and an **X** on those that are convex. If necessary, you may write several letters on each figure.



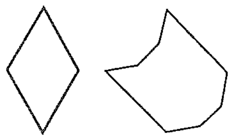

- For polygons 2, 4, 5, and 7, explain your choice of letters or why you did not write any letter.

**Item 3.** A) - Write all the important properties which are <u>shared</u> by *squares* and *rhombi*.
- Write all the important properties which are <u>true</u> for *squares* but <u>not</u> for *rhombi*.
- Write all the important properties which are <u>true</u> for *rhombi* but <u>not</u> for *squares*.

B) - The same questions as in A) for *equilateral triangles* and *acute triangles*.

**Item 4.** A) - You can see a shape in figure -a- (a rhombus). Make a list of all the properties that you find for this shape (you can draw to explain the properties).

B) - The same question for shape in figure -b-.



- a - - b -

**Item 5.1.** - Recall that a *diagonal* of a polygon is a segment that joins two non adjacent vertices of the polygon. How many diagonals does an n-sided polygon have? Give a proof for your answer.

**Item 5.2.** - Complete the three following statements (you can draw if you want):

In a 5-sided polygon, the number of diagonals which can be drawn from each vertex is ..... and the total number of diagonals is .....

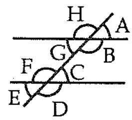
In a 6-sided polygon, the number of diagonals which can be drawn from each vertex is ..... and the total number of diagonals is .....

In an n-sided polygon, the number of diagonals which can be drawn from each vertex is ..... Justify your answer.

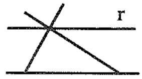
- Using your answers above, tell how many diagonals an n-sided polygon has. Prove your answer.

**Item 6.1.** - Prove that the sum of the angles of <u>any</u> acute triangle is 180°.

**Item 6.2.** - Recall that, if you have two parallel straight lines cut by another straight line: All the acute angles marked in the figure (A, G, C, E) are equal. All the obtuse angles marked in the figure (B, H, D, F) are equal.




Taking into account the figure on the right (line r is parallel to the base of the triangle) and the properties mentioned above, prove that the sum of the angles of <u>any</u> acute triangle is 180°.




**Item 6.3.** - Here is a complete proof that the sum of the angles of <u>any</u> *acute* triangle is 180°. Read it and try to understand it.

- The sum that we are supposed to calculate is $M + R + T$ (figure 1).
- Construct a parallel to the base of the triangle through the opposite vertex R (figure 1). Extending a side, we have two parallel lines cut by a transverse, so $T = t$ (figure 2).
- Extending the other side we have two parallel lines cut by a transverse, so $M = m$ (figure 3).
- Therefore, $M + R + T = m + R + t = 180°$, as the latter three angles form a straight angle (figure 4).

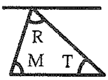


Figure 1.

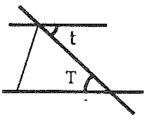


Figure 2.

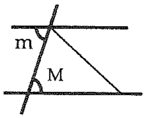


Figure 3.

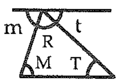
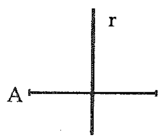


Figure 4.

- You have seen above a proof that the sum of the angles of an acute triangle is 180°. Is it true that the sum of the angles of a *right* triangle is 180°? Prove your answer.
- Tell how much is the sum of the angles of an *obtuse* triangle: <u>Exactly</u> 180°, <u>more than</u> 180°, or <u>less than</u> 180°. Prove your answer.

**Item 7.** A) - Prove that the two diagonals of <u>any</u> rectangle have the same length.

B) - Recall that the *perpendicular bisector* of a segment is the line perpendicular to that segment that cuts it through its midpoint (line *r* is the perpendicular bisector of segment AB). Prove that any point of the perpendicular bisector of a segment is equidistant from the endpoints of the segment.




**Item 8.** - Usually a *parallelogram* is defined as a quadrilateral having two pairs of parallel sides.

Could a *parallelogram* also be defined as a quadrilateral in which the sum of any two consecutive angles is 180°? Justify your answer: If your answer is affirmative, prove that both definitions are equivalent. If your answer is negative, draw some example.

Usually all the questions of an item are presented in the same sheet. However, each section in items 5 and 6 is presented in a different sheet, and students are not allowed to "go back" to correct or complete answers in previous sheets after they have moved to the second or third section. In both items, the first section just states a property to be proved; then, the second and third sections provide some hints to help students to understand and complete the proof. In this manner, we allow more able students to complete the proofs on their own, and less able students to work on the problems with some help and to produce some kind of answer, in a way similar to the procedure used in clinical interviews, were the researcher, when necessary, guides the student with a hint, comment, question, etc.

The table below summarizes the key processes evaluated in each item and the possible Van Hiele levels of students' answers. It supports the validity and reliability of the test, since i) every process is considered at least by an item, and ii) for each level of thinking, there are several items that can be answered by students in that level. In Spain, students in upper Primary and Secondary are likely acquiring thinking level 2 or 3; for this reason we have included in the test a high number of items assessing these levels.

| | **Van Hiele levels** | | | | | **Definition** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Item** | 1 | 2 | 3 | 4 | **Identif.** | Use | State | **Classif.** | **Proof** |
| **1** | • | • | | | • | | | | |
| **2** | • | • | | | • | | | | |
| **3** | • | • | • | | | • | | • | |
| **4** | • | • | | | • | • | | | |
| **5.1, 5.2** | | • | • | • | | • | | | • |
| **6.1** | | • | • | • | | | | | • |
| **6.2, 6.3** | | • | • | | | | | | • |
| **7** | | • | • | • | | • | | | • |
| **8** | | | • | • | | • | • | | • |

This test was administered to students in upper Primary (grades 6 to 8) and Secondary (grades 1 to 4) (ages from 11 to 18). To optimize the administration, we did not present the eight items to all the students, but we took into account the particular characteristics of students in different grades: Since primary school students were most likely reasoning in levels 1 or 2, we reduced the number of items evaluating levels 3 and 4 in their test. In the same way, we reduced the number of items evaluating levels 1 or 2 in the test for upper secondary school students. On the other side, the mathematical content of an item may increase its difficulty in certain grades, so we avoided items whose contents likely still had not been studied in some of the grades 6 to 8. Then, we administered three different sub-tests:

A) The test for students in grades 6, 7, and 8 had items: 1, 3, 4, 6, 7.

B) The test for students in grades 9 and 10 has items: 1, 2, 3, 5, 6.

C) The test for students in grades 11 and 12 has items: 1, 3, 5, 6, 8.

All the test have five item, three of them being the same items, to guarantee the possibility of comparison of results, and the other two items selected depending on the expected students' level of thinking and their knowledge of geometric facts. Tests A and B do not assess the process of statement of definitions. Although it may be considered as a weakness of these tests, in pilot trials, we noticed that questions asking to compare definitions or to build a definition (i.e. a list of necessary and sufficient properties) from a list of given properties were meaningless for most students in those grades. Then, we decided to exclude this kind of questions, to have a shorter and more efficient test.

## Results of the Administration of the Tests.

The test was administered to 309 students. The table on the right shows the number of students in each grade. The answers were codified according to the method of levels and types of answers defined in Gutiérrez, Jaime, Fortuny (1991), and a vector with 4 percentages was assigned to each student, showing the student's acquisition of Van Hiele levels 1 to 4. Both researchers made independent assignations of level and type to each answer, then both assignations were compared, and the disagreements were analyzed. Some times this analysis resulted in an improvement of the marking criteria, and a new marking of some answers if necessary.

| Grade | Students |
|---|---|
| 6th Primary | 34 |
| 7th Primary | 62 |
| 8th Primary | 83 |
| 1st Secondary | 35 |
| 2nd Secondary | 36 |
| 3th Secondary | 28 |
| 4th Secondary | 31 |

In order to make more meaningful the results of the evaluation of the tests, the numeric scale of percentages of acquisition of a Van Hiele level can be translated into a qualitative scale of degrees of acquisition of the level, as follows:

[ No acquisit. ] Low acquisit. [ Intermed. acq. ] High acquisit. [ Complete acq. ]

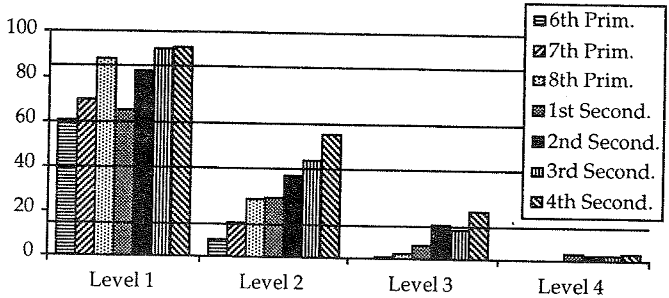
0% 15% 40% 60% 85% 100%

The vectors of the degrees of acquisition of the four levels provide information about the behaviour of every individual student. Analysing those data, we have identified several profiles, that correspond to different styles and qualities of reasoning. The table in the next page presents the most significant of such profiles and the percentages of students having each profile (for instance, in CHI≤L, C stands for complete acquisition of level 1, H for high acquisition of level 2, I for intermediate acquisition of level 3, and ≤L for low or null acquisition of level 4).

| Profiles | 6th | 7th | 8th | 1st | 2nd | 3rd | 4th |
|---|---|---|---|---|---|---|---|
| C C I,H ≤I | 0 | 0 | 1,2 | 0 | 5,6 | 3,6 | 12,9 |
| C H I ≤L | 0 | 0 | 0 | 0 | 5,6 | 3,6 | 9,7 |
| C H ≤L N | 0 | 4,8 | 4,8 | 0 | 0 | 25,0 | 12,9 |
| C I ≤L N | 0 | 3,2 | 10,8 | 11,4 | 5,6 | 7,1 | 25,8 |
| C ≤L ≤L N | 26,5 | 43,6 | 59,0 | 17,1 | 44,4 | 42,9 | 22,6 |
| H ≤I ≤L N | 20,6 | 9,7 | 12,1 | 28,6 | 26,8 | 0 | 0 |
| I N N N | 29,4 | 8,1 | 7,2 | 2,9 | 0 | 14,3 | 0 |
| L ≤L N N | 20,6 | 25,8 | 2,4 | 34,3 | 2,8 | 0 | 0 |
| N N N N | 2,9 | 4,9 | 1,2 | 0 | 0 | 0 | 0 |

An evolution in the students' kind of reasoning along the grades can be observed: The higher the course, the bigger the number of students showing better profiles of reasoning, with the exception of 1st and 2nd grades of Secondary.

It is also interesting to analyze the relationship between the results of students in different grades. Next chart shows the means of the acquisition of each Van Hiele level by the students in each grade, so it provides with a global picture of the differences from students in different grades. Some conclusions can be drawn from both table and chart:




- Most students had a high or complete acquisition of level 1. There is a progress in the acquisition of this level along the Primary grades, and also along the Secondary grades. However, it is noticeable the reduction of the acquisition of this level in the first grade of Secondary. A reason for such reduction may be that some students in this grade had not enough time to complete the test, since 8 students (23% of the group) did not answer the last item in the test (item #3), and half of them neither answered the previous item (#2). Both items assess levels 1 and 2. The influence of this problem in the results of level 1 is rather important since only three items evaluate this level, but its influence in the results of level 2 is smaller because it is assessed by all the items in the test.

- Students in 7th grade of Primary and 1st, and 2nd grades of Secondary had not completely acquired the first level, although they showed at the same time a low acquisition of level 2. In the same way, students in the upper grades of Secondary showed an intermediate acquisition of level 2 and also a certain acquisition of level 3. One of the main characteristics of the Van Hiele Model is the hierarchy of the levels (a student is supposed to begin the acquisition of a level only after s/he has completed the acquisition of the previous level), but the reality of the teaching of mathematics is that often students are being taught in the higher level of reasoning, and teachers force them to answer according to that level. The result is that students are not allowed to complete the acquisition of the lower level but, sometimes, they acquire practice, although only a few, in the higher level. However, the phenomenon of "reduction of level" has to be taken into consideration (Van Hiele, 1986).

- Only 17 students in the sample had an intermediate or better acquisition of level 3, and only 7 students showed a low or intermediate acquisition of level 4. So, most Spanish students leave the Secondary School having a low or null acquisition or level 3, i.e. almost completely unable to make any kind of mathematical deductive reasoning (neither formal nor informal). These poor results may be a consequence of the usual way of teaching Mathematics in Secondary School, where many teachers emphasizess formal proofs (level 4) when, as we see in the previous graph, students are reasoning only in level 1 or 2.

## References.


Crowley, M.L. (1989): *The design and evaluation of an instrument for assessing mastery Van Hiele levels of thinking about quadrilaterals.* (U.M.I.: Ann Arbor, USA).

Gutiérrez, A.; Jaime, A.; Fortuny, J.M. (1991): An alternative paradigm to evaluate the acquisition of the van Hiele levels, *Journal for Research in Mathematics Education* vol. 21.3, pp. 237-251.

Gutiérrez, A.; Jaime, A.; Shaughnessy, J.M.; Burger, W.F. (1991): A comparative analysis of two ways of assessing the van Hiele levels of thinking, in *Proceedings of the 15th P.M.E. Conference* vol. 2, pp. 109-116.

Jaime, A.; Gutiérrez, A. (1994): A model of test design to assess the Van Hiele levels, in *Proceedings of the 18th P.M.E. Conference* vol. 3, pp. 41-48.

Mayberry, J. (1981): *An investigation of the van Hiele levels of geometric thought in undergraduate preservice teachers.* (U.M.I.: Ann Arbor, USA).

Senk, S.L. (1983): *Proof-writing achievement and van Hiele levels among secondary school geometry students* (Doct. Diss.). (University of Chicago: Chicago, USA).

Usiskin, Z. (1982): *Van Hiele levels and achievement in secondary school geometry.* (ERIC: Columbus, USA).

Van Hiele, P.M. (1986): *Structure and insight. A theory of mathematics education.* (Academic Press: London, G.B.).