

## A COMPARATIVE ANALYSIS OF TWO WAYS OF ASSESSING THE VAN HIELE LEVELS OF THINKING

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

J. Mike Shaughnessy and William F. Burger. Dept. of Mathematics; Oregon State University. Corvallis, OR (USA).

### Abstract

*This study provides the results of a comparison between two different ways of determining the Van Hiele Levels of reasoning. On one hand, we have compared a clinical interview versus a paper and pencil test. On the other hand, we have compared the assignation of students to a single level of reasoning versus their assignation to a degree of acquisition of each Van Hiele level. Subjects from both test environments were assigned to the Van Hiele levels by using the two assessment procedures. Comparisons of the two assessment procedures yield close agreement on Van Hiele levels acquisition of many of the subjects, particularly the American students.*

### Introduction and Rationale

In the late Seventies, a growing interest for the Van Hiele Model of Reasoning started in the Western Countries. Since then, there has been a continuous research activity analyzing and applying the Van Hiele Levels; the assessment of the students' level of thinking has played a relevant role in this research. Several types of tests have been used for this assessment:

- a) Paper and pencil tests with multi-choice questions (Usiskin, 1982).
- b) Paper and pencil tests with open-ended questions (De Villiers, 1987; Jaime, Gutiérrez, 1990).
- c) Clinical interviews with open-ended questions (Burger, Shaughnessy, 1986; Mayberry, 1981).
- d) Learning sequences (Fuys, Geddes, Tischler, 1988).

Each one of these kinds of tests has been used in several works, and has proved to have both advantages and inconveniences, although the usefulness of the first one has been questioned because it offers serious doubts about its reliability to reflect the student's thinking (Crowley, 1989; Wilson, 1990).

On the other hand, independent of the kind of test used, several ways of assignment of the students to the Van Hiele Levels have been used by researchers:

a) Each question in the test is assigned to a specific Van Hiele level, and the answers are marked as good or bad. The student is assigned to an overall level on the basis to the number of good/bad answers for each level (Mayberry, 1981; Usiskin, 1982).

b) Each student's answer is assigned to the specific Van Hiele level it reflects. The student is assigned to an overall level on the basis of the number of answers in each level (Burger, Shaughnessy, 1986; De Villiers, 1987; Fuys, Geddes, Tischler, 1988).

c) The Van Hiele levels are considered in a continuous way, and each student's answer is assigned to a point in that continuum. The degrees of the student's acquisition of each level are determined on the basis of that assignments (Gutiérrez, Fortuny, Jaime, (w.d.); Jaime, Gutiérrez, 1990).

A current problem in the research related to the Van Hiele Levels is the absence of standardized tests and criteria for determining the students' level of thinking, valid and reliable enough as to be used as reference when designing a new test or way of assessment. The aim of the research we are reporting is to compare and analyze two kinds of tests and two ways of determining the Van Hiele levels of thinking, by means of a crossed application: We have selected subsamples of the students who participated in the researches reported in Burger, Shaughnessy (1986) and Jaime, Gutiérrez (1990), and we have made a twin assignment of each student in those subsamples to the Van Hiele levels by using our two assessment criteria independently. In this report we present the results of the study.

There are several complete descriptions of the Van Hiele Levels in the literature; in particular, in Burger, Shaughnessy (1986) and Jaime, Gutiérrez (1990) appear the descriptors we have used in this research. Here we just make short statements of the characteristics of the levels to help the reader:

Level 1 (Recognition): Students judge a geometrical object by its appearance and consider it as a whole.

Level 2 (Analysis): Students identify the components of geometrical figures, and they describe them by means of their properties. Deductions are based on observation and measurement.

Level 3 (Informal deduction): Students are able to logically classify families of

figures. Definitions are meaningful for students, and they can give informal arguments for their deductions.

Level 4 (Formal deduction): Students understand the role of the elements of an axiomatic system, and they can perform formal proofs.

### Method

A) The Tests. The basis for this work were two tests built and administered as part of previous research projects carried out by the members of the team.

- One was the test used in the Project "Assessing Children's Intellectual Growth in Geometry" (Burger, Shaughnessy, 1986 and 1990). It had 8 open-ended tasks dealing with geometric shapes, and it was administered by means of clinical interviews to a range of North-American students from grades K to 12 and university mathematics majors. We will refer to them as the American test and American students. In short, the items were: 1) Drawing of several different triangles. 2) Identifying and defining several sorts of triangles. 3) Sorting triangles according to attributes chosen by the student and by the interviewer. 4, 5, and 6) Similar to the previous ones, but for quadrilaterals. 7) "What's my shape?" that is, identifying a shape from a list of clues. 8) Equivalence of definitions of parallelogram, and knowledge about axioms, theorems and postulates.

- The other test was the one used in the Project "Design of a Curricular Proposal for the Learning of Secondary Geometry According to the Van Hiele Levels" (Jaime, Gutiérrez, 1990). This test had 6 open-ended items (some of them divided into two parts), and it was administered in a written form to several classes of first grade students in Spanish Secondary Schools (students aged from 14 to 16). We will refer to them as the Spanish test and Spanish students. In short, the items were: 1) Identifying regular, irregular, concave, and convex polygons. 2) Identifying several sorts of quadrilaterals, according with the student's definitions. 3) Identifying squares and rectangles according with given definitions. 4) Identifying appearances of "anla" (a non-standard polygon) in a set of polygons by using its definition. 5) Classifying "anlas" and regular polygons or quadrilaterals. 6) Building a definition for obtuse triangle from a given relationship of properties.

The main difference among the two tests was the way of administration. There is no doubt that a clinical interview is the best way for assessing the

student's thinking level, since the interaction between student and interviewer may give an in depth knowledge and the interviewer may ask for more explanations when necessary. But, on the other hand, clinical interviews are very time-consuming and are not appropriate for big samples; in this case, researchers are forced to use a written test. The weakness of a written test is that usually students tend to write short explanations, and then it is difficult to know their real level of thinking; the result may be the assignment to students of a Van Hiele level lower than the real one. When designing the Spanish test, we were aware of this problem, and we stated the items asking the students to give complete explanations and stating concrete questions about the reasons for their choices.

B) The Marking Schemes. The two ways of assigning students to the Van Hiele levels used in our researches represented two different interpretations of the process of acquisition of a level: The American Project assumed the hypothesis of the discreteness of the Van Hiele levels, while the Spanish Project assumed the hypothesis of the continuity of the levels.

- The American researchers made assignments (each researcher independently) of each student's answer to the predominant level of thinking exhibited saying that "it was in level n", although in some cases they realized that the answer had clear indications of two consecutive levels, and then they assigned it to a rating like 1-2, indicating that the answer showed a transition from level 1 to level 2. Finally, each researcher made an overall assignment of students to the Van Hiele levels, based on the ratio between the levels of the different student's answers, which was either a level or a transition between two consecutive levels (Burger, Shaugnessy, 1990).

- The Spanish researchers tried to reflect in their assignments how strongly each student's answer was rooted in a given level (they made independent assignments and afterward they put them together, looking for a consensus in the discordant assignments). Every answer was assigned both to the Van Hiele level it better reflected, and to a certain "type of answer", depending on how clearly the level was reflected and on its mathematical accuracy (from the point of view of the reflected Van Hiele level). The types of answers were:

Type 0. No reply or answers which cannot be codified.

Type 1. Answers indicating that the learner has not attained a given Van Hiele level but which give no information about any lower level.

Types 2/3. Wrong/correct answers which contain very few explanations but giving some indication of a given level of thinking.

Type 4. Answers with clear and sufficient justifications which clearly reflect characteristic features of two consecutive levels of thinking.

Type 5. Wrong answers which clearly reflect a given level of thinking.

Types 6/7. Incomplete/complete, correct, and sufficiently justified answers which clearly reflect a given level of thinking.

After marking the answers to a test, the types of answer were quantified and the student's "degree of acquisition" of a given Van Hiele level was determined by calculating the arithmetic average of the values of the student's answers to those items that could have been answered at that level. The overall assignation for a student was a vector with four values, from 0 to 100, reflecting the student's degree of acquisition of the Van Hiele levels 1 to 4. For a detailed description, see Gutiérrez, Fortuny, Jaime (w.d.).

C) The Sample. The American researchers selected 6 students from the sample in the American project, aiming to represent the different ages and thinking abilities present in the whole American sample. In the same way, the Spanish researchers selected 6 students from the sample in the Spanish project, aiming to represent the different thinking abilities present in the whole Spanish sample.

D) The Process. Our purpose was to obtain a twin assignation of both American and Spanish students to the Van Hiele levels, in order to compare the results and to have conclusions referring to the kind of test (oral/written) and the marking criteria (discrete/continuous Van Hiele levels).

Then, the Spanish researchers marked, according to their own marking schemes explained above, the students in the American subsample. The working material were the audiotapes and students' drawings from the clinical interviews carried out by the American researchers. Similarly, the American researchers marked, according to their own marking schemes explained above, the students in the Spanish subsample. The working material were the texts from the written tests administered by the Spanish researchers.

## Results

Tables 1 and 2 present the twin assignation of students to the Van Hiele levels of thinking. The first rows contains vectors with the student's numeric

degrees of acquisition of the Van Hiele levels 1, 2, 3, and 4. The second and third rows contain the independent assignments of students to the levels made by the American member of the team.

Table 1. The results for the Spanish students (written tests).

	Susana	Mayte	Juan
Spanish	(100, 39, 13, 0)	(50, 25, 3, 0)	(90, 27, 3, 0)
American (S)	1	1	1-2
American (B)	1	1-2	1*

  

	Maria	Salud	Yolanda
Spanish	(38, 2, 0, 0)	(100, 58, 32, 0)	(100, 17, 0, 0)
American (S)	1-2	2	1
American (B)	1-2	2*	1

\* With indication of the upper level.

Table 2. The results for the American students (oral tests).

	Amy	Tyrone	Don
Spanish	(100, 81, 0, 0)	(95, 27, 0, 0)	(100, 98, 64, 33)
American (S)	2	1	3
American (B)	2	1	3

  

	Karen	David	Tom
Spanish	(100, 84, 26, 0)	(98, 55, 0, 0)	(100, 100, 74, 52)
American (S)	2	1	4
American (B)	2	1	4

The first thing that catches the attention is the difference between the two kinds of results. While the assignment to a single level (or a range of two levels) just tells us which was the dominant level (or the two dominant levels), the degrees of acquisition tell us how confidently and satisfactorily the student used each of the four levels in the tasks.

It is very difficult a perfect agreement between such different ways of assignment, but an analysis of Tables 1 and 2 looking for similarities and discrepancies shows that there is a rather good agreement between the three assessments. The numeric degrees of acquisition of the levels may be transformed in qualitative ones (as seen in Table 3): From 0 to 15, there is no

acquisition of the level; from 15 to 40, there is a low acquisition; from 40 to 60, an intermediate acquisition; from 60 to 85, a high acquisition; and from 85 to 100, there is a complete acquisition of the level. For example, the vector (100, 58, 32, 0) means complete acquisition of level 1, intermediate acquisition of level 2, low acquisition of level 3, and no acquisition of level 4. In this way, we can easily compare the results from the different assignments.

Table 3. Qualitative results for the Spanish assignments.

Susana	Mayte	Juan	María	Salud	Yolanda
(C, L, N, N)	(I, L, N, N)	(C, L, N, N)	(L, N, N, N)	(C, I, L, N)	(C, L, N, N)
Amy	Tyrone	Don	Karen	David	Tom
(C, H, N, N)	(C, L, N, N)	(C, C, H, L)	(C, H, L, N)	(C, I, N, N)	(C, C, H, I)

C = Complete; H = High; I = Intermediate; L = Low; N = No acquisition

The better agreement between the three assignments was in the American Students; only the cases of David and, mainly, Tom presented a slight disagreement between the American and the Spanish assignments. Tom was one of the university majors, and may be the American researchers were so expectant of his high Van Hiele level that his answers were unconsciously over-valued. This may be a danger of any method of assessment based on subjective estimations without clear objective directions. Similarly, in a group of good students, we may have an unconscious tendency to under-value a poor student.

The bigger disagreements appeared in the Spanish subsample (Mayte and María); it has relation with the fact that these were the poorest students, since a poor student usually gives few answers and short explanations, making more difficult the task of analyzing the answers and giving an accurate assessment; in fact, María answered only 4 out of the 9 questions.

There is another reason for a better agreement in the American Students: In an interview there is the possibility of asking the student for an explanation of a confusing answer, so the interviewer may modify the script if necessary. When marking the written tests, there was a general claim between the researchers that we would like to have the opportunity of interviewing them for asking for some explanations.

### Conclusions

In this paper we have stated some point that should receive further

attention. There is still a lot of work to do in relation to the Van Hiele Levels, mainly in curricular development, and, previously, it is necessary a detailed work for improving the current ways of evaluating the levels of thinking. With respect to the tests, there is not a clear direction; we think that the most promising possibility is a kind of test based on paper and pencil open-ended questions followed by short interviews devoted to ask students for explanations on the dark parts of their written answers. On the other hand, it is quite clear that the traditional assignation of students to a single level is a simplistic view which lost part of the richness of the student's answers, so research should be done aiming to develop new methods of evaluation based on the observation of the ability of students in using the four Van Hiele levels, as a way for obtaining a more complete picture of the student's thinking.

### References

- Burger, W.F.; Shaughnessy, J. (1986): Characterizing the van Hiele levels of development in geometry, *Journal for Research in Mathematics Education* 17.1, pp. 31-48.
- Burger, W.F.; Shaughnessy, J.M. (1990): *Assessing children's intellectual growth in geometry* (final report). (Oregon State University: Corvallis, USA).
- Crowley, M.L. (1989): *The design and evaluation of an instrument for assessing mastery Van Hiele levels of thinking about quadrilaterals* (Ph.D. dissertation). (University of Maryland: Maryland).
- De Villiers, M.D. (1987): *Research evidence on hierarchical thinking, teaching strategies and the van Hiele theory: Some critical comments* (Research Unit for Math. Educ.; Univ. of Stellenbosch: Stellenbosch, R. South Africa).
- Fuys, D.; Geddes, D.; Tischler, R. (1988): *The van Hiele model of thinking in geometry among adolescents*, "Journal for Research in Mathematics Education Monograph" n° 3. (N.C.T.M.: Reston, USA).
- Gutiérrez, A.; Fortuny, J.M.; Jaime, A. (w.d.): An alternative paradigm to evaluate the acquisition of the van Hiele levels, *Journal for Research in Mathematics Education*, to appear.
- Jaime, A.; Gutiérrez, A. (1990): Study of the degree of acquisition of the van Hiele levels by secondary school students, in *Booker; Cobb; Mendicuti (1990): Proceedings of the 14th International Conference of the P.M.E.* (Authors: México) vol. 2, pp. 251-258.
- Mayberry, J. (1981): *An investigation of the van Hiele levels of geometric thought in undergraduate preservice teachers*. (University Microfilms Internat.: USA).
- Usiskin, Z. (1982): *Van Hiele levels and achievement in secondary school geometry*. (ERIC: Columbus, USA).
- Wilson, M. (1990): Measuring a van Hiele geometry sequence: A reanalysis, *Journal for Research in Mathematics Education* 21.3, pp. 230-237.