

Full Length Research Paper

An instrument for measuring performance in geometry based on the van Hiele model

Ana B. Sánchez-García^{1*} and Ana Belén Cabello²

¹Department of Education, Paseo de Canalejas 169, Faculty of Education, Salamanca. INICO-University of Salamanca-37008. Spain

²Department of Didactic of Mathematics. Facultad de Educación. Rector Royo Villanova s/n Ciudad Universitaria. Complutense University of Madrid 28040 – Madrid, Spain.

Received 02 April, 2016; Accepted 15 June, 2016

In this paper we present the process of constructing a test for assessing student performance in geometry corresponding to the first year of Secondary Education. The main goal was to detect student errors in the understanding of geometry in order to develop a proposal according to the Van Hiele teaching model, explained in this paper. Our research methodology took into account reliability using Cronbach's alpha coefficient, as well as the construct validity, with the extraction of 13 factors that accounted for a high percentage of variance. This result leads us to conclude that the instrument constructed has the appropriate technical and pedagogical features to be considered an original and significant contribution to the field of geometry teaching. The final version of the test constructed after the extraction of factors is shown in the Appendix.

Key words: Assessment, geometric concepts, mathematics curriculum, research, Van Hiele Model.

INTRODUCTION

Although there have been continuous changes in Mathematics teaching in Spain since 1970, perhaps the greatest transformation lies in our way of understanding how to construct geometric thinking in the classroom. Currently, learning is conceived as an active and constructive process. However, both the international PISA 2012 (OECD, 2014) findings and the Spanish Educational System Diagnosis (INEE, 2010, 2011), as well as popular opinion, all reflect certain deficiencies in academic performance in the area of mathematics. Some of the current problems in the teaching-learning of mathematics in general, and of geometry in particular,

may be the result of a lack of mathematical knowledge on the part of individuals studying to become teachers (Sánchez and López, 2011), as well as a result of their training in mathematics education (Rico, 2012), an aspect that should be taken into account in the teaching of geometry.

Nonetheless, solving these professional challenges is something that can be expected of the field of research in mathematics teaching (Rico, 2012; Sierra, 2011). For all of these reasons and as a result of other problems typical of geometry teaching, students have difficulty in understanding and learning the subject, and for the same

*Corresponding author. E-mail: asg@usal.es.

reasons, teaching has focused on developing students' deductive abilities, omitting interaction (conjecturing, exploring).

Being aware of the problems involved in this sphere of teaching-learning, we believe that more educational research is needed to broach and solve the dysfunctions detected in the different reports assessing our educational system. We therefore posited investigating the most representative theories in this area of study, which then led to the need to design an instrument to measure performance in geometry of students in the first year of Compulsory Secondary Education in Spain (acronym in Spanish: ESO), which corresponds to 12 to 13 years old students. This is the year and stage of education that our research into geometry education focuses on. The findings of this work are important for their repercussions on student learning and improved teaching. Our concern as mathematics teachers is to incorporate to our teaching practice the contributions that have come from the field of educational research, specifically regarding geometry teaching, and to investigate and propose efficient teaching models.

As regards the specific research objectives, an educational research approach was taken, based on the need to solve problems of comprehension in geometry learning that come up daily in the classroom, and in this study the results of the design and validation of an instrument that allows us to probe into the mistakes that students make in the geometry block of contents in the curriculum were presented. To this effect, a test (Appendix) to measure the students' knowledge before and after completing the geometry subject matter in the first year of ESO was devised. With this instrument and in successive stages of research, our objective was to detect student errors in understanding at both moments of learning.

When devising the measuring instrument in question, we took as a reference the curricular framework for the first year of compulsory secondary education in Spain. We likewise took into account the related bibliography considered relevant by the scientific community, but above all, the students' errors in comprehension.

The curricular proposals for learning geometry that was analyzed are based mainly on the Van Hiele Model (1957, 1986), analyzed by Corberán et al. (1994), Jaime (1993), Gutiérrez et al. (1991, 1994), and Peeg et al. (1997), in addition to assessment tests designed to evaluate the contents in this area, such as those by Gutiérrez et al. (1991), Mayberry (1981, 1983), Usiskin (1982) and other contributions to geometry teaching such as those by Alsina et al. (1997), and Rico et al. (2002). Furthermore, other studies, such as those by Guillén (2000) and Huerta (1999), suggest focusing the study on the detection of errors as part of the research stage in the Van Hiele Model and adapting the theoretical frame to the reality of the classrooms in which the research is to

be carried out. Thus, in our case the instrument had 13 items with three responses to choose from and was applied to a sample of 177 students in two schools of secondary education (Table 1).

After a detailed analysis of the theories related to geometry learning in order to make a theoretical selection of the items for the final test, it can be said that these theories are linked to two focal points: the first focuses on the cognitive processes that deal with the comprehension of geometric concepts (Hoffer, 1983; Jaime, 1993; Van Hiele, 1957, 1986) and the second draws together all the theories emphasizing the formation of geometric concepts, such as the Theory of Figural Concepts (Fischbein, 1993), the Concept Image - Concept Definition model (Vinner and Hershkowitz, 1980) and the Visualization-based Interpretation of Geometric Concepts (Duval, 1995, 1998). These approaches are not independent, but rather complement each other. Although in no way underestimating the influence of the theories underlying any of these theoretical models, in this study our central focus revolves around the basic level defined by the Van Hiele, which has to do with the aspect of visualization (the image of the concept) in the forming of geometric concepts, and which we describe briefly in the next section. Other relevant studies about the forming of geometric concepts which will also serve as a reference point in the shaping of the thematic nucleus underlying our research were likewise approached.

THEORETICAL BACKGROUND

The theory known as the Van Hiele model emerged in Holland in 1957. Pierre and Dina Van Hiele defended their doctoral dissertations and carried out successive research studies until they developed a model which explains how students gradually construct geometric thinking in five consecutive levels, guided by the teacher through five stages of learning. Their contribution to the field of geometry learning has been discussed by different authors, the most outstanding being those by Fuys et al. (1988) and Pegg (1992).

The theoretical construction of the model has a dual nature: a descriptive aspect, in which the levels of geometric thinking that the student goes through are specified, and a prescriptive aspect, which establishes the learning stages the teacher must guide the students through so that they can acquire a determined level of knowledge.

The descriptive aspect is hierarchically ordered into five levels of thinking (Patkin, 2014), although according to Patkin and Barkai (2014), owing to doubts that have emerged from mathematics educators (including van Hiele himself), today the usual practice is to speak of the following four levels: (i) Recognition or visualization, (ii) analysis or description, (iii) informal deduction, and (iv)

Table 1. Percentages of distribution by sex in the sample (aged: 12 to 13).

School	No. of students	Females	%	Males	%	NA	%
IES Joaquin Araujo	81	41	50.61	40	49.39	0	0
IES Dolores Ibarruri	96	49	51.04	45	46.88	2	2.08
Total	177	90	50.85	85	48.02	2	1.13

formal deduction

In addition to the specific thinking that takes place at each level, Van Hiele identified five general properties that characterize the model: sequential, advancement, intrinsic and extrinsic, linguistics, and mismatch. These properties are significant to educators because they provide guidance for making instructional decisions (Crowley, 1987).

On the other hand, the prescriptive aspect consists of five learning stages: (i) Inquiry; (ii) Directed orientation; (iii) Explanation, (iv) Free orientation, and (v) Integration.

According to Clements (2003) progression from one level to the next can be facilitated through planned instruction. Therefore geometry instruction should be designed with these stages in mind (Choi-koh 1999).

Authors such as Koester (2003) and Patkin and Sarfaty (2012) have focused on investigating the different levels of knowledge acquisition in geometry and the difficulties involved. The existence of an important line of research in this field have also been noted, related to the work done by Vinner and Hershkowitz (1980) and Hershkowitz (1990, 1999), which is focused on knowledge of the cognitive processes that take place in students' minds when they are learning a new geometric concept. In this sense, the authors differentiate the image of the concept from its definition. The results of their studies led them to formulate a theory focused on the concept, which is called mental drawing, defined as the set of all the drawings that have been associated with the concept in the learner's mind. Thus, the image of the concept is composed of its mental drawing together with the properties that have been associated with that drawing in the mind of the learner. The definition of the concept is a verbal definition that accurately explains its comprehension.

These authors affirm that knowing the conceptual images that the students have in their minds is very important for teaching, since it gives teachers a better understanding of students' learning processes and helps them to suggest improvements for teaching to help avoid the formation of images of erroneous concepts.

In parallel, Fischbein (1993) developed the Theory of Figural Concepts to refer to mental entities that have both conceptual and figural characteristics. From the point of view of this model, images and concepts interact constantly during the reasoning process and sometimes produce conflictive situations until the geometric concept

is formed.

Duval (1993, 1995, 1998) developed his theory by focusing on the framework of acquisition of geometric concepts based on different systems of representation located in the field of semiotics, a field subject to important cognitive activity. For this author, the construction of a geometric concept is directly related to the construction of its graphic representation.

From our point of view, at the basic level of recognition or visualization (the stage at which the students in our sample find themselves) students have not yet formed the figural concept; they know only images without knowledge of the concepts and think that the figures possess irrelevant attributes such as their orientation on the plane; that is why they find themselves in a conflictive situation in which the concept loses strength in relation to the image.

For this reason, and as indicated earlier, in order to detect the level of formation of geometric concepts and errors in comprehension as part of the inquiry stage of the Van Hiele model and to adapt the theoretical framework to the reality of the classroom, our study focuses on the basic level of recognition or visualization in the Van Hiele model, analyzing the conceptual images that students show at this level. The objective is highly relevant, since we are not aware of any previous research that has broached the relation between the recognition and the visualization of the concept (Van Hiele level 0) on one hand, and the curricular design proposal for geometry in Spain, on the other. Thus, this study revolves around the drawing up and validating of a test for measuring the performance of students in their first year of compulsory secondary education in geometry. In the next section, we describe the process of validating the instrument.

MATERIALS AND METHODS

This research was carried out using a quantitative analytical empirical approach with a non-experimental design in order to validate the instrument in question.

Technical characteristics: Sample, process, reliability and validity

As shown in Table 1, the target population was comprised of male and female students in their first year of compulsory secondary

education (ESO) aged between 12 and 13. The schools were selected for their convenience. Following the current terminology, it can be said that the sample was incidental, because it was the one available for the study at the time it was being carried out (Pereda, 1987).

The first question posited was the number of sample units necessary and, as pointed out by Morales (2012) the question to be answered is "What is the study for? The literature is abundant in this sense and different authors maintain different criteria. The issue was related mainly to educational research. In our case, the fitting question is: How many students are needed to construct and analyze a measurement instrument and to subject the instrument to item and factor analysis, as well as interpret its reliability and validity?"

Nunnally (1978) and Afifi and Clark (1990) proposed that there should be at least between five and ten subjects per item. This may be the most widespread criterion for carrying out factor analysis when trying to establish construct validity (Argibay, 2006). Other authors, such as Guilford (1954) and Kline (1986, 1994), consider a small sample adequate as long as the number is not much below 200.

To determine the number of sample units we adopted the criterion established by the above-mentioned authors. Since the recommendation is to use between 150 and 200 subjects, we assume that our sample, comprised of 177 students, is sufficient and adequate, and thus acceptable. The sample characteristics are shown in Table 1.

The process was began with an updated review of the literature concerning the contents to be assessed. This analysis allowed us to identify a series of variables that appeared as constants in the assessment of the geometrical concepts to be measured. Taking as a reference the theoretical foundation described in the previous section, and in awareness of the complexity involved in the construction of an instrument to assess the acquisition of geometric concepts, we decided to use a multiple choice test as the measurement instrument because it would allow us to collect and quantify the information in a simple way. In the wording of the questions for the pilot test, we took into account the specifications of Blaxter et al. (2000) and followed the indications of Spector (1992) for the stages used to devise it.

Once the items were selected (32 questions with three responses to choose from [96 items]) a group of experts evaluated the suitability of each item for the attribute to be measured. This is the most common way to detect the quality of the contents, especially in the educational context (Prieto and Muñiz, 2000; Prieto and Delgado, 2010). This is why we turned to a group of experts in the contents of the instrument in order to evaluate its items, instructions, and design (Millman and Greene, 1989).

Once the contents of the items had been analyzed, some of the items (19 questions with three responses to choose from [57 items]) were eliminated and changed the content of others in line with the indications from the group of experts. A cognitive analysis of the items was also carried out in order to learn the strategies that the students would use to answer them.

Since these procedures alone are not sufficient (Visser et al., 2000), an attempt was made to strengthen the process using the analyses presented in the following sections.

Subsequently, with the sample selected, and before imparting the teaching programmed for the contents in geometry, a pilot test was carried out and the test was given to 177 students in order to analyze its internal consistency. This processes yielded a Cronbach's alpha coefficient = 0.938, which according to authors such as Web (1983) or Kline (2000) more than satisfies the reliability requirements for the instrument.

It can thus be concluded that the results obtained show very high reliability, which according to the Cronbach interpretation would be

within the coefficient value between 0.85 and 0.90, whose meaning is catalogued as almost perfect. Nonetheless, in order to be as rigorous as possible with the process we should also take into account the reflections made by Morales (1995) to the effect that when reliability is calculated in the human sciences, one should specify what kind of reliability it is. A low coefficient does not necessarily mean that the instrument is bad and should not be used. Reliability is not a characteristic of the instrument, but rather a characteristic of the results, of the scores obtained by a specific sample. And it is very important to underscore this fact even though it is commonly referred to as if it were a characteristic of the instrument; this frequent way of referring to it must be understood in order to understand what it means. One same instrument can measure and classify one sample very well but do so very badly with a high margin of error when applied to another sample. One same instrument can measure well if the subjects differ widely among themselves but have a low reliability if the sample is homogeneous. Reliability must thus be calculated for each sample, regardless of the reliability of the instrument. The data obtained from applying our test were analyzed using the statistical program SPSS 20.0

Subsequently, items were empirically selected based on having administered each one to the sample of subjects selected to this effect. Statistical analysis of the test was begun by analyzing the reliability of each item (Morales, 2003; Lukas, 1998; Muñiz and Yela, 2003). The term "difficulty index" is usually used to indicate the ratio of correct to incorrect answers to an item in the student sample in question, but to be consistent with the formula it was agreed to call it the "facility index." This analysis, together with the evaluation by experts, allowed us to shorten the test with a wide range of facility indices that gave the test a rating of average difficulty, desirable for measuring student learning. Subsequently, the psychometric characteristics of the test were measured in order to verify its reliability and validity to then draw up the definitive version, even though most of the items discriminated satisfactorily.

Before presenting the reliability study, it is worthwhile to recall that reliability expresses the accuracy of the measuring instrument and validity means that the instrument actually measures what we want it to measure.

The reliability of the results of the test, that is, its internal consistency as measured by Cronbach's alpha coefficient, was 0.938. Cronbach (1960) affirmed that only those tests with a reliability coefficient of at least 0.90 should be used for educational purposes (which is our case), whereas Nunnally (1978) proposed a minimum value of 0.70. After reviewing the works of several authors, Webb (1983) proposed an interpretation of the reliability coefficient that would fall within a coefficient value between 0.85 and 0.90. If we take that as a reference, it can be concluded that our results show very high reliability. Factor analysis was carried out, identifying the fit of the data, using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's sphericity test.

Some authors, such as Lukas (1998) and Muñiz (2003), report that attempts are being made to extract certain inferences from the test data. Validity refers to the fits, significance and usefulness of these inferences. It is traditionally defined as the degree to which a test measures what it is supposed to measure. This definition, however, needs to be qualified and requires specification with the validity of the contents, of the criterion and of the construct, as stated previously.

The content validity of the test seems to be guaranteed by the study carried out by the group of experts, the result of which was the elimination of certain items.

Finally, construct validity, the most important level of validation (Cronbach and Meehl, 1955), aimed at determining the degree to which the test measures the theoretical trait in question, is a complex process that usually includes several procedures. As

Table 2. Explanatory percentage of variance in the factor analysis.

Factor	Eigen values	%Variance	%Accumulative
Factor 1	12.846	26.762	26.762
Factor 2	3.201	6.668	33.430
Factor 3	2.679	5.582	39.012
Factor 4	2.473	5.151	44.163
Factor 5	2.390	4.978	49.141
Factor 6	1.916	3.992	53.133
Factor 7	1.810	3.772	56.905
Factor 8	1.649	3.435	60.340
Factor 9	1.462	3.047	63.387
Factor 10	1.307	2.723	66.110
Factor 11	1.289	2.685	68.794
Factor 12	1.166	2.429	71.223
Factor 13	1.101	2.293	73.516

Messick (1980) states: "...construct validity is indeed the unifying concept of validity that integrates criterion and content considerations into a common framework for testing rational hypotheses" (p.1015).

In regard to the technical demands of a study of these characteristics, we can affirm that all of them have been fulfilled; there is usually consensus about the importance of construct validity (Cronbach, 1988; Chacón and Moreno, 2000; Loewinger, 1957; Messick, 1995; Moss, 1992; Pérez-Gil; Tenopyr, 1977); in addition, factor analysis is usually run to analyze the structure being measured (Cronbach, 1980, 1982, 1988; Guion, 1977; Tenopyr, 1977; Thurstone, 1931, 1947). Besides observing the high correlation between the items, the two tests mentioned earlier were carried out to analyze the appropriateness of applying factor analysis (García et al., 2000; Pérez, 2009). In line with technical data obtained, Table 2 and 3 show the descriptive statistics.

This structure, which maintains an unequivocal relation to the curriculum of the first year of upper secondary education in Spain, is what allows us to confirm the validity construct that can be seen in the Table 3.

The analyses performed yielded very satisfactory results. The definitive version of the test (Appendix) consists of 13 items with three responses to choose from.

RESULTS

This paper describes the process of devising an instrument (test) for measuring performance in geometry of students in their first year of compulsory secondary education (ESO) in Spain. After analyzing the theoretical framework and other preliminary assessment tests in this field, the test was given to a sample of 177 students in order to validate it. The empirical treatment and validation process consisted of two stages: the first, aimed at obtaining the number of factors and percentage of explained variance, yielded a test comprised of 13 factors, and the second, confirmatory analysis, determined that the test was comprised of 13 factors and 13 items.

The items on the test are related to the theory of acquisition of geometrical concepts and to the contents included in the prescriptive curricular design in Spain. It can be affirmed that the test is related to the attribute that it is intended to measure and fits the target population. Likewise, the data obtained from the (exploratory and confirmatory) factor analyses showed adequate values, allowing us to establish a definitive version consisting of 13 items grouped into 13 factors with eigenvalues higher than one that explain 73.516% of the variance.

The instrument was carried out in a pilot study. A quasi-experimental study with one non-randomized study group (n=137) was also conducted using a pre-and post-test design with the purpose of searching for differences before and after the Van Hiele learning process and detecting persistent errors (Cabello et al., 2014). The instrument allowed us to observe the initial concepts that the students had in relation to the contents of geometry and their most common errors. There were significant differences between the average performance of each of the two groups of the study in favor of the experimental group taught with the Van Hiele Model ($F= 0.317$, $p =0.006$). Eight persistent errors were also detected after implementing the learning process, only two of which were detected in control group. Therefore, the results confirm that the instrument is valid to detect the initial errors of students in relation to the geometrical concepts that they have acquired. We think it is important to have real data on the initial situation of the students before implementing the Van Hiele Model for effective teaching. The relevance of persistent error detection is that it allows us to determine the efficacy of the teaching methodology and that it indicates the errors that students make. The positive effects of employing the Van Hiele model in geometric concepts acquisition are clear (Al-ebous, 2016), but a tool is needed to detect geometric

Table 3. Factor composition.

Factor	Items	Weighting Factors	M	DS
1 Classification of triangles according to their sides and angles	7b2	0.875	0.63	0.803
	7c2	0.859	0.71	0.868
	7a2	0.858	0.62	0.790
	7b1	0.722	0.82	0.897
	7a1	0.713	0.72	0.866
2 Recognition of polygons (difficult)	11h	0.669	0.43	0.672
	13e1	0.669	0.42	0.728
	11j	0.658	0.55	0.746
	11g	0.653	0.69	0.839
	11c	.538	0.63	0.844
3 Identification of the relative position of two angles	5b	0.854	1.40	0.733
	5a	0.835	1.41	0.734
	5c	0.809	1.38	0.737
4 Determining the angles of similar triangles and obtaining an angle from knowledge of the other two angles	6b	0.939	0.88	0.874
	6a	0.919	0.90	0.853
	6c	0.831	0.76	0.776
5 Identifying types of angles	4b	0.814	1.90	0.314
	4c	0.790	1.88	0.348
	4a	0.756	1.93	0.282
6 Recognition of the elements of a circle	16c	0.699	0.97	0.907
	16b	0.676	0.75	0.836
	16e	0.664	0.51	0.770
7 Identification of (difficult) planes related to the circle	17e	0.847	1.25	0.433
	17c	0.808	1.30	0.459
	17b	0.705	1.33	0.473
8 Determining the area of a polygon	13b3	0.851	0.79	0.927
	13c3	0.851	0.76	0.923
	13e3	0.444	0.22	0.556
	18c	0.445	0.67	0.822
9 Recognition of the elements of a regular polygon	12b	0.634	0.42	0.696
	12a	0.628	0.98	0.935
	12c	0.619	0.44	0.689
10 Recognition of (easy) polygons	c1	0.894	1.26	0.965
	b1	0.884	1.29	0.956
11 Identification of the relative position of two lines on a plane	2a	0.848	1.56	0.737
	2c	0.723	1.35	0.700
	2b	0.491	1.92	0.351
12 Identification of the perpendicular bisector and the bisector	8	0.802	1.20	0.936
	9	0.785	1.24	0.923
13 Identification of the (easy) planes relating to the circle	17a	0.763	1.77	0.423

mistakes before implementing the Van Hiele Model.

Conclusion

It can be concluded that the instrument presented here has construct validity and there is evidence of its predictive potential in relation to the grades obtained by the students in the subject.

This instrument therefore will allow us to explore students' errors in geometry comprehension and seek the best way to correct them through application of the Van Hiele Model. This test was considered to be a contribution to the field of teaching/learning geometric concepts, one that can be very useful in the detection of mistakes made by students in this area.

Conflict of interests

The authors have not declared any conflict of interests.

REFERENCES

- Al-ebous T (2016). Effect of the Van Hiele Model in Geometric Concepts Acquisition: The Attitudes towards Geometry and Learning Transfer Effect of the First Three Grades Students in Jordan. *Int. Educ. Stud.* 9(4):87-98. doi:10.5539/ies.v9n4p87
- Afifi AA, Clark V (1990). *Computer-aided multivariate analysis*. New York, NY: Chapman & Hall.
- Alsina C, Fortuny J, Pérez R (1997). ¿Por qué Geometría?. Propuesta didáctica para la ESO [Why Geometry? Didactic proposal for high school]. Madrid, Spain: Síntesis.
- Argibay JC (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad [Psychometric techniques. Issues of validity and reliability]. *Subjetividad y Procesos Cognitivos*, 8:15-33. Retrieved from <http://dspace.uces.edu.ar:8180/xmlui/handle/123456789/765>
- Baxter L, Hughes C, Tight M (2000). ¿Cómo se hace una investigación? [How is an investigation?]. Barcelona, Spain: Gedisa.
- Bollen KA, Long JS (1994). *Testing structural equation models*. Newbury Park, CA: SAGE Publications.
- Cabello AB, Sánchez-García AB, López R (2014) Persistent geometric errors in students of 7th grade: detection and correction methodology. *Epsilon. Revista de Educ. Matemática*, 31(1):21-36.
- Corberán R, Gutiérrez A, Huerta M, Jaime A, Margarit J, Peñas, Ruiz E (1994). Diseño y evaluación de una propuesta curricular de aprendizaje de la Geometría en Enseñanza Secundaria basada en el modelo de razonamiento de Van Hiele [Design and evaluation of a proposed curriculum of learning of geometry in high school based on reasoning model of Van Hiele]. Madrid, Spain: M.E.C.
- Choi-koh SS (1999). A Student's Learning of Geometry Using the Computer. *J. Educ. Res.* 92(5):301-11. doi:<http://dx.doi.org/10.1080/00220679909597611>
- Clements DH (2003). Teaching and Learning Geometry. In: J. Kilpatrick, W. Gary Martin & D. Schifter (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp. 15–78). Reston, VA.: National Council of Teachers of Mathematics
- Cronbach LJ (1980). Validity on parole: How can we go straight? In W. B Schrader (Ed.) *New directions for testing and measurement, planning achievement: Progress over a decade* (pp. 99-108). San Francisco, CA: Jossey-Bass.
- Cronbach, LJ (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.
- Cronbach LJ (1988). Internal consistency of tests: Analyses old and new. *Psychometrika*, 53(1):63-70. Retrieved from <http://hdl.handle.net/10.1007/BF02294194>.
- Cronbach LJ, Meehl PE (1955). Construct validity in psychological tests. *Psychol. Bull.* 52:281-302. Retrieved from <http://dx.doi.org/10.1037/h0040957>
- Duval R (1995). *Sémiosis et pensée humaine. Registres sémiotiques et apprentissages intellectuel* [Semiosis and human thought. Learning intellectual and semiotic registers]. Berne, Switzerland: Peter Lang.
- Duval R (1998). Registro de representación semiótica y funcionamiento cognitivo del pensamiento [Registration of representation of semiotic and cognitive functions in thought]. In F. Hitt (Ed.) *Investigaciones en Educación Matemática II* [Research in mathematics education II] (pp. 173-201). México, México: Grupo Editorial Iberoamérica.
- Fischbein E (1993). The theory of figural concepts. *Educ. Stud. Mathe.* 24(2):139-162. doi: 10.1007/BF01273689
- Fuys D, Geddes D, Tischer R (1988). *The Van Hiele model of thinking in Geometry among adolescents*. Reston, VA: National Council of Teachers of Mathematics.
- García E, Gil J, Gómez G (2000). *Análisis factorial* [Factorial design]. Madrid, Spain: Hespérides.
- Groves R, Fowler F, Couper M, Lepkoski J, Singer E, Tourangeau R (2004) *Survey methodology*. New Jersey, NJ: John Wiley & Sons.
- Crowley ML (1987) The van Hiele Model of the Development of Geometric Thought." In: M. M. Lindquist & A.P. Shulte (Eds.), *Learning and Teaching Geometry, K-12, 1987 Yearbook of the National Council of Teachers of Mathematics* (pp.1-16). Reston, Va.: National Council of Teachers of Mathematics.
- Guilford JP (1954). *Psychometric methods*. New York, NY: McGraw-Hill.
- Guillén G (2000). Sobre el aprendizaje de los conceptos geométricos relativos a los sólidos. Ideas erróneas [On the learning of geometrical concepts to solids. Misconceptions]. *Enseñanza de las Ciencias* 18(1):35-53.
- Guion RM (1977). Content validity: Three years of talk-what's the action? *Public Personnel Manage.* 6:407-414.
- Gutiérrez A, Jaime A, Fortuny JM (1991). An alternative paradigm to evaluate the acquisition of the Van Hiele levels. *J. Res. Math. Educ.* 22:237-251.
- Hershkowitz R (1990). Psychological Aspects of Learning Geometry. In P. Neshet & J. Kilpatrick (Eds.), *Mathematics and Cognition*. Cambridge, England: Cambridge University Press pp. 70-95.
- Hershkowitz R (1998). On reasoning in geometry. In: C. Mammana & V. Villani (Eds.) *Perspectives on the teaching of geometry for the 21st century*. (An ICMI study) Dordrecht, The Netherlands: Kluwer Academic Publishers. pp. 29-37.
- Hoffer A (1983). Van Hiele based research. In R. Lesh & M. Landau (Eds.) *Acquisition of Mathematical Concepts and Processes* (pp. 205-227). New York, NY: Academic Press.
- Huerta P (1999). Los niveles de Van Hiele y la taxonomía SOLO: un análisis comparado, una integración necesaria [Van Hiele levels and SOLO Taxonomy: a comparative analysis, a required integration]. *Revista Enseñanza de las Ciencias*, 17(2):291-309.
- Instituto de Evaluación Educativa (INEE) (2010). *PISA 2009. Programa para la Evaluación Internacional de los Alumnos. Informe Español* [PISA 2009. Programme for the international assessment of students. Madrid, Spain: Ministerio de Educación.
- Instituto Nacional de Evaluación Educativa (INEE) (2011). *Evaluación general de diagnóstico 2010. Educación Secundaria Obligatoria. Segundo curso. Informe de resultados* [General Diagnostic 2010 evaluation. Compulsory Secondary Education. Report of results]. Madrid, Spain: Ministerio de Educación.
- Jaime A (1993). Aportaciones a la interpretación y aplicación del modelo de Van Hiele: la enseñanza de las isometrías del plano. La evaluación del nivel de razonamiento [Contributions to the interpretation and application of Van Hiele model: teaching of the isometrics in the drawing. Evaluation of the level of reasoning]. (Unpublished Doctoral dissertation). Universidad de Valencia, Valencia, Spain.
- Jaime A, Gutiérrez A (1994). A Model of Test Design to assess the Van Hiele Levels. In: J. P. Daponte & J.F. Matos (Eds.) *Proceedings of*

- the 18th Conference of the International Group for the Psychology of Mathematics Education, Lisboa, Portugal: Universidad de Lisboa. 3:41-48.
- Kline P (1986). *A Handbook of Test Construction*. New York, NY: Methuen.
- Kline P (1994). *An Easy Guide to Factor Analysis*. Newbury Park, CA: SAGE Publications.
- Kline P (2000). *The handbook of psychological testing (2nded.)*. New York, NY: Routledge.
- Koester BA (2003). Prisms and pyramids: Constructing three-dimensional models to build understanding. *Teach. Children Math.* 9(8):436-442.
- Loevinger J (1957). Objective tests as instruments of psychological theory. *Psychol. Reports*, 3:635-694.
- Lukas JF (1998). Análisis de ítems y de test con ITEMAN [Analysis of items and test with ITEMAN]. Bilbao, Spain: Universidad País Vasco.
- Mayberry JW (1981). An investigation of the Van Hiele levels of geometric thought in undergraduate pre-service teachers. (Unpublished doctoral dissertation) University of Georgia, Athens, GA.
- Mayberry JW (1983). The Van Hiele levels of geometric thought in undergraduate preservice teachers. *J. Res. Math. Educ.* 14(1):58-69.
- Messick S (1980). Test validity and the ethics of assessment. *Am. Psychol.* 35(11):1012-1027.
- Messick S (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 15:5-12.
- Millman J, Greene J (1989). The specification and development of tests of achievement and ability. In R.L. Linn (Ed.) *Educational measurement*. London, England: Macmillan pp. 335-336.
- Moos PA (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Rev. Educ. Res.* 62:229-258.
- Morales P, Urosa Sanz B, Ángeles BB (2003). Construcción de escalas de actitudes tipo Likert: una guía práctica [Construction of attitudes Likert-type scales: a practical guide]. Madrid, Spain: La Muralla, S.A.
- Morales P (2012). Tamaño necesario de la muestra: ¿Cuántos sujetos necesitamos? [Required sample size: how many subjects do we need?] Retrieved from: <http://www.upcomillas.es/personal/peter/investigacion/Tama%F1oMuestra.pdf> [February 2nd, 2012]
- Muñoz J (2003). *Teoría clásica de los tests [Classical test theory]*. Madrid, Spain: Pirámide.
- Organization for Economic Co-operation and Development (OECD) (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014)*. PISA: OECD Publishing.
- Patkin D (2014). Global van Hiele (GVH) Questionnaire as a tool for mapping knowledge and understanding of plane and solid geometry. *Journal of the Korean Society of Mathematical Education. Series D: Res. Mathe. Educ.* 18(2):103-128.
- Patkin D, Barkai R (2014). Geometric thinking levels of pre- and in-service mathematics teachers at various stages of their education. *Educ. Res. J.* 29(1/2):1-26.
- Patkin D, Sarfaty Y (2012). The effect of solid geometry activities of pre-service elementary school mathematics teachers on concepts understanding and mastery of geometric thinking levels. *Journal of the Korean Society of Mathematical Education Series D: Res. Math. Educ.* 16(1):31-50.
- Pegg J (1992). Students' Understanding of Geometry: Theoretical Perspectives. In B. Southwell, B. Perry & Owens, K. (Eds.) *Space The First and Final Frontier* (pp. 18-36). Sydney, Australia: Mathematics Education Research Group of Australasia.
- Peeg J, Gutiérrez A, Huerta MP (1997). Assessing Reasoning Abilities in Geometry. In V. Villani & C. Mammana (Eds.) *Perspectives on the teaching of Geometry for the 21st Century*. Dordrecht, The Netherlands: Kluwer Academic Publishers pp. 275-295.
- Pereda S (1987). *Psicología Experimental. I. Metodología [Experimental Psychology. I. Methodology]*. Madrid, Spain: Pirámide.
- Pérez-Gil JA, Chacón S, Moreno R (2000). Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez [Construct validity: The use of factor analysis]. *Psicothema*, 12(2):442-446.
- Prieto G, Delgado A (2010). Fiabilidad y validez [Reliability and validity]. *Papeles del Psicólogo* 31(1):67-74.
- Prieto G, Muñoz J (2000). Un modelo para la evaluación de los tests utilizados en España [A model for the evaluation of the tests used in Spain]. *Papeles del Psicólogo* 77:65-72.
- Rico L (2012). Aproximación a la Investigación en Didáctica de la Matemática [Approach to the research on Didactics of Mathematics]. *Avances de Investigación en Educ. Matem.* 1:39-63.
- Rico L, Sierra M, Castro E (2000). Didáctica de la Matemática [Didactics of mathematics]. In L. Rico & D. Madrid (Eds.) *Fundamentos didácticos de las áreas curriculares*. Madrid, Spain: Síntesis pp. 351-406.
- Rico L, Sierra M, Castro E (2002). El área de conocimiento Didáctica de la Matemática [Didactics of the mathematical knowledge field]. *Revista de Educ.* 328:35-58.
- Sánchez AB, López R (2011). La transferencia del aprendizaje algorítmico en el origen de los errores [The transfer of algorithmic learning and the origin of errors in subtraction]. *Revista de Educ.* 354:429-445.
- Sierra M (2011). Investigación en educación matemática: objetivos, cambios, criterios, métodos y difusión [Research in mathematics education: objectives, changes, criteria, methods and dissemination]. *Educación Siglo XXI*, 29(2):173-198.
- Spector PE (1992). *Summated rating scale construction*. Newbury Park, CA: SAGE Publications, Inc.
- Tenopyr ML (1977). Content-construct confusion. *Personnel Psychol.* 30:47-54.
- Thurstone LL (1931). *Multiple Factor Analysis*. *Psychol. Rev.* 38:406-427.
- Thurstone LL (1947). *Multiple Factor Analysis*. Chicago, IL: University of Chicago Press.
- Usiskin Z (1982). *Van Hiele levels and achievement in secondary school geometry*. Chicago, IL: University of Chicago.
- Van Hiele PM (1957). El problema de la comprensión en conexión con la comprensión de los escolares en el aprendizaje de la Geometría. (Doctoral Dissertation). Utrecht, Holland: University of Utrecht.
- Van Hiele PM (1986). *Structure and insight. A theory of Mathematics Education*. New York, NY: Academic Press.
- Vinner S, Hershkowitz R (1980). Concept images and common cognitive paths in the development of some simple geometrical concepts. In: R. Karplus (Ed.) *Proceedings of the International Conference for the Psychology of Mathematics Education* (pp. 177-184). Berkeley, CA :California University.
- Visser P, Krosnick J, Lavrakas P (2000). Survey Research. In: H. Reis & C. Judd. (Eds.) *Handbook of research methods in social and personality psychology*. Cambridge, England: Cambridge University Press) pp. 223-252.