A Characterization of Van Hiele's Level 5 of Geometric Reasoning Using the Delphi Methodology

Alberto Arnal-Bailera[1] · Víctor Manero[2]

Abstract

The Van Hiele model of geometric reasoning establishes five levels of development, from level 1 (visual) to level 5 (rigor). Despite the fact that this model has been deeply studied, there are few research works concerning the fifth level. However, there are some works that point out the interest of working with activities at this level to promote the acquisition of previous levels. Our goal is to describe this level through the construction and validation of a list of indicators for each of the processes involved in geometrical reasoning (definition, proof, classification, and identification). Due to the lack of previous research, we have decided to use the Delphi methodology. This approach allowed us to collect information from a panel of experts to reach a consensus through a series of phases about the indicators that describe each of the processes. The final product of the iterative application of this method is a list of validated indicators of the fifth Van Hiele level of reasoning. In particular, proof and definition processes have turned out to be the most relevant processes at this level.

Keywords Definition · Delphi methodology · Proof · Van Hiele

Introduction

In recent years, research in mathematics education at university level has received an increasing interest shown by the work of specific groups in the most relevant international conferences in the area such as CERME, ICME, INDRUM, or RUME.[1] However, the number of contributions specifically focused on geometry at university

---

[1] CERME: Congress of the European Society for Research in Mathematics Education; ICME: International Congress on Mathematical Education; INDRUM: International Network for Didactic Research in University Mathematics; RUME: Research in Undergraduate Mathematics Education.

✉ Víctor Manero
vmanero@unizar.es

1 Mathematics Department, University of Zaragoza, Zaragoza, Spain

2 Mathematics Department, University of Zaragoza, Huesca, Spain

level is really scarce, apart from those related to prospective secondary school teachers. This work has made it possible to increase the knowledge of the mathematical learning processes at tertiary level.

Among these processes, we are interested in the teaching and learning of geometry. The Van Hiele model has been widely used both to study the level of geometric thought of different groups of students or teachers (Arnal-Bailera & Manero, 2021; González et al., 2022; Jaime & Gutiérrez, 1994; Manero & Arnal-Bailera, 2021; Pandiscio & Knight, 2010; Wang & Kinzel, 2014) and to design learning activities (Abdullah & Zakaria, 2013; Corberán et al., 1994; Guillén, 2004; Howse & Howse, 2014; Wahab et al., 2017). However, there are few works concerning the fifth level (Blair, 2004; Mayberry, 1983; Usiskin, 1982). In order to deepen the knowledge of this level of thought, we decided to give credit to the subjective collective judgment of experts, in this case geometry experts, via the Delphi methodology.

The aim of this study is to gain understanding on the fifth level as a base to start a line of research on the possible effectiveness of an instructional approach to strengthen the acquisition of the previous levels (Blair, 2004; Martin, 2008). In particular, we are thinking about ways of increasing the degree of acquisition of the fourth level of mathematics majors and pre-service mathematics teachers (Manero & Arnal-Bailera, 2021; Demiray & Işiksal, 2017; Pandiscio & Knight, 2010; Sears, 2019). These ideas show the relevance of describing level 5 in terms of a set of indicators (sentences expressing the different abilities of a person at this level). These indicators could serve as a starting point to design level 5-activities that could increase the degree of acquisition of level 4 among mathematics majors and pre-service mathematics teachers. In terms of instructional design, these indicators could help to design the goal of a high-level geometry course. Our research question is: Which characteristics describe better the highest level of reasoning according to the Van Hiele theory?

In order to answer this question, we establish the following specific objectives:

(1) Design a list of indicators for every key process present in the fifth Van Hiele level
(2) Validate the relevance of the indicators obtained to describe the fifth Van Hiele level

Theoretical Framework

Van Hiele Model

The Van Hiele model (Van Hiele, 1957) has been one of the most relevant theoretical frameworks concerning teaching and learning geometry at all educational levels. This model states the existence of five different levels of geometric reasoning (Burger & Shaughnessy, 1986; Hoffer, 1983; Jaime & Gutiérrez, 1990; Van Hiele, 1986) that can be summarized as follows:

- Level 1 (visualization). Students at this level recognize geometric figures by their appearance and as a whole. Also, they describe figures using their physical characteristics or comparisons with everyday objects by means of a nonmathematical language
- Level 2 (analysis). This level is characterized by the students' ability to handle the parts and properties of figures, which allows them to deal with mathematical descriptions of geometric concepts
- Level 3 (informal deduction). The reasoning of this level uses logical deductions in the first place, which enable students to interrelate properties of geometric figures. Thus, these students can understand logical classifications of families of figures, construct definitions as sets of necessary and sufficient conditions, and provide some general arguments to justify the validity of a mathematical statement
- Level 4 (formal deduction). Students at this level can produce formal proofs and deal with equivalent definitions of a concept
- Level 5 (rigor). Students at this level can compare systems based on different axioms and can study several geometries in absence of concrete models

The main characteristics of these levels consist of its sequential and hierarchical nature, meaning that they are acquired in a specific order throughout the learning process.

Despite the fact that the Van Hiele model contains five levels, it must be emphasized that almost all the related literature has been focused on the four first ones. This lack of work on the fifth Van Hiele level can be explained by the fact that during the scholar years very few students start the acquisition of level 4. In particular, according to Gutiérrez and Jaime (1995), only 22.6% of the last secondary year students show a certain degree of acquisition of level 4.

Among the very few studies in which level 5 has been considered, the works of Usiskin (1982), Mayberry (1983), and Blair (2004) stand out. Usiskin (1982) designed a test whose objective was to determine the Van Hiele level of the students. That test contains 25 questions in which level 5 was also considered. Also, Mayberry (1983) gave some ideas of the properties that questions of a Van Hiele test must have for every level. In particular, the author pointed out that questions in level 5 must be related with propositions considering finite geometries; notice that this idea is coherent with the questions proposed by Usiskin. The level 5 is also considered in the work of Blair (2004). Concretely, the author describes tasks involving classic geometry but considering non-conventional metrics (like Taxicab geometry) as a possible way to develop level 5.

Key Processes

Some authors, such as De Villiers (1987), describe the Van Hiele levels attending to the different processes that are involved in them. Jaime and Gutiérrez (1994) present a description of the Van Hiele levels (from 1 to 4) that is organized in terms of the different key processes: definition (use and formulation of definitions of geometrical objects); proof (the way of convincing ourselves or someone else of the

truth of a statement); classification (sort geometrical objects into different families or create a new group of families to sort the objects); and identification (establishing the family to which a particular geometrical object belongs to). Historically these are the process that have been considered in order to describe the Van Hiele levels (Burger & Shaughnessy, 1986; De Villiers, 1987; Gutierrez & Jaime, 1998). Currently, there exists a consensus among the Van Hiele research community that these four processes are the crucial ones to consider as indicators for the Van Hiele levels; for example, recent studies that consider the potential of the Van Hiele framework for Graph Theory have considered these four levels (Manero & Arnal-Bailera, 2021; González et al., 2021).

Concerning the definition process, level 1 definitions consider only physical global properties or consist in descriptions based on diary objects like "a rectangle is something like a window". Students at level 2 can state definitions in terms of a list of mathematical properties but they do not care if the list of properties contains all the necessary and sufficient ones. Level 3 definitions are stated as a set of necessary and sufficient mathematical properties. Students at level 4 accept equivalent definitions for the same mathematical object and are able to prove the equivalence of both definitions by double implication. Previous research works have provided some key ideas into high-level reasoning: Martín-Molina et al. (2018), in a study about the way professional mathematicians define, present the definition inscribed in a more general process of generalizing while Larsen and Zandieh (2008) point out to proof as a motivation to define when working across different geometries.

Proofs at level 2 consist of verifying the concerning property or statement in one or several concrete examples. Students at level 3 are able to make proofs that only require a few logical steps. The arguments given in order to explain the truth of a statement are based on mathematical properties but usually are informal or lack rigor. In contrast, level 4 students understand proofs and are able to produce formal proofs. Regarding the proof of a proposition, people on this level are conscious that it requires a sequence of implications based on already established properties. Previous research works have provided some key ideas into high-level reasoning: Fernández-León et al. (2020) state the ability of a researcher of selecting and applying different proving methods; Weber and Mejia-Ramos (2011) showed that mathematicians will eventually construct sub-proofs (e.g. lemmas) to help the understanding of the proof; finally, Boero (1999) distinguishes between mathematical proofs as a product, and proving as a (cyclic) process. The work across geometries is a situation that could promote comparing several geometries. This comparison could be based on the study of the transferability of a demonstration (Blair, 2004).

Level 1 classifications are produced attending to global physical properties. Students at level 2 are able to classify geometrical objects into disjoint families attending to their mathematical properties. Inclusive classifications (classifications into families with relations between them) may be done at level 3. Regarding the identification process, it is characterized at level 1 by the recognition of geometrical objects based on their global appearance and physical properties while at levels 2, 3, and 4, this process is based on mathematical properties. The classic descriptions of the Van Hiele levels do not include specifications about what is or what is not classification and recognition at level 5. Moreover, these processes are supposed to be fully

developed at lower levels (Jaime & Gutiérrez, 1994). However, it is reasonable to think that working in different geometries or with different metrics could develop different skills to identify or classify objects or environments.

Delphi Method

The Delphi technique (Linstone & Turoff, 1975) consists of a methodology whose objective is to discuss and obtain consensus about a certain question through the iterative administration of a series of questionnaires to a panel of experts. The use of the Delphi method for this research is justified by the facts that there is little literature concerning the related topic and there are many experts in the field of geometry at a tertiary level. Thus, a subjective collective judgment of experts can be an appropriate way to obtain reliable information.

Since the Delphi method allows us to obtain a reliable consensus of opinion from a group of experts, it has been used in many different areas of social sciences like economy, policy making, or social services. It has also been useful in educational settings due to the fact that it eases to form guidelines, standards, or predict trends (Green, 2014). For example, in the last decade, it has been a usual tool in order to develop and validate competencies that teachers should have acquired; for primary school science teachers (Alake-Tuenter et al., 2013) and secondary school mathematics teachers (Muñiz-Rodríguez et al., 2017).

This method can be described as follows: A questionnaire is designed by a small monitor team (usually called research group) who sends it to a larger respondent group (also called panel of experts). Once the completed questionnaires are received, the research group summarizes the obtained answers and, based upon them, develops a new questionnaire for the panel of experts who should evaluate the ideas presented on it. In successive phases, the panel reevaluates the same or a modified questionnaire in which is also given a controlled feedback of the group responses to the previous assessments. This iterative process finishes when a termination criterion is achieved (Linstone & Turoff, 1975).

The implementation of this method can be summarized explaining the initial preparation, the administration of questionnaires, and the termination criterion.

The initial preparation requires to determine the expert panel and the elaboration of the first questionnaire. On one hand, Hsu and Sandford (2007) describe that people with experience and related background on that issue can be considered eligible. Also, heterogeneity on the members of the panel is desirable. On the other hand, the number of panelists is in general lower than 50 (Witkin & Altschuld, 1995), and most of Delphi studies contain 15–20 experts (Ludwig, 1997). The first questionnaire is traditionally formed by open-ended questions; however, structured questionnaires on the first phase can be used if they are obtained after an exhaustive review of the related literature (Hsu & Sandford, 2007).

After the first questionnaire is fulfilled and sent back, the research group analyzes the experts' productions. This analysis consists of a revision of their answers in order to clarify, avoid duplications, and classify them into different categories. With the results of the previous analysis, a structured questionnaire is designed.

Afterwards, the expert panel is required to evaluate the relevance of the different items presented there, generally using a Likert scale (Cox, 1996). In the following phases, the panel of experts receive some controlled feedback possibly including all the panelists' answers to the previous phase (Andranovich, 1995; Scheibe et al., 1975) or just some measures of centrality; in this latter case, the use of the median is preferred in the literature (Eckman, 1983; Hill & Fowles, 1975; Jacobs, 1996). Along with this feedback, the panelists are asked to reevaluate the items.

Regarding the termination criterion, Linstone and Turoff (1975) established that three phases could be sufficient to complete the study. Another termination criterion consists in finishing the study when consensus is achieved. Usually, consensus is said to have been obtained when a high percentage of ratings concentrates in a concrete range (Miller, 2006). Another finishing criterion can be validity: A certain item or idea is considered to be validated when there exists consensus about its relevance or convenience. Other authors (Scheibe et al., 1975) consider stability (no significant differences in the evaluations of consecutive phases) in order to terminate the Delphi study.

Methodology

Application of the Delphi Method

The iterative administration of the questionnaires has been done via online, concretely using Google questionnaires. The number of phases was previously fixed in three as stated by Linstone and Turoff (1975).

The first questionnaire was validated by two professors of mathematics education who are experts in the Van Hiele model, none of them being involved in this work.

In the first phase, a personalized email was sent to the participants including a cover letter and a link to the questionnaire. In the letter, the study and their objectives were presented as well as a brief description of the Delphi method and a schedule of the phases.

At the beginning of the questionnaire, we included an introduction to the Van Hiele model (see Fig. 1). The questionnaire consisted of five questions. The first four questions were related to the four key processes (Burger & Shaughnessy, 1986) and were introduced by a short description of the corresponding process, for each level, according to the one given in Gutiérrez and Jaime (1995). After this introduction, we posed the following open question: *Regarding this process, what elements, capabilities … Describe someone who is at level 5 of geometric reasoning?* Thus, the panel of experts was required to elaborate comments that, in their opinion, described the different processes in level 5. The last question was: *If you consider that Van Hiele level 5 involves processes other than those described above, you can indicate it here.* This question intended to explore possible different processes involved in the highest Van Hiele level.

Finally, the research group analyzed all these comments following a thematic analysis (Clarke & Braun, 2016): We clustered them into similar categories, clarified ambiguous wordings, discarded some due to their inappropriateness, and deleted duplicities. As a result, we obtained the first list of indicators.

Van Hiele's model (Van Hiele, 1957) has been a theoretical reference in the field of teaching and learning geometry at all educational levels. However, the more advanced levels have been very little studied, being very few works focused on this topic.

The general purpose of our research project is to develop scientific knowledge of the characteristics that distinguish the most advanced level of geometric reasoning (level 5 - rigor). In the fifth level, as described by Burger and Shaughnessy (1986), a person is able to compare systems based on different axioms and can study various geometries in the absence of concrete models. Aspects dealt with by people at this level would be non-Euclidean geometries, non-Euclidean metrics or finite geometries, among others.

A good way to understand Van Hiele's levels is through how the different key processes that appear in geometry are understood in each of them: definition (use and establishment), proof, classification and identification

Fig. 1 Introduction to the Van Hiele model presented in the questionnaire

In the second phase of the Delphi method, the experts were asked to evaluate the relevance of the indicators obtained after phase 1 using a four-point Likert scale ((1) no relevance; (2) insignificantly relevant; (3) relevant; and (4) highly relevant). Experts were allowed to make comments in order to justify their ratings.

In order to validate the different indicators, we followed a similar approach to the one described in Muñiz-Rodríguez et al. (2017). We classified the indicators into three different groups according to its validity, i.e. the consensus about its high relevance:

- Validated. When at least 70% of the experts' ratings were 4
- Reevaluate. Between 40 and 70% of the experts' ratings were 4
- Rejected. When less than 40% of the experts' ratings were 4

The consensus percentage has been established at 70% since according to Hsu and Sandford (2007) and Green (2014), this is the percentage used to determine consensus for a four-point Likert scale. Then, if there exists consensus concerning the high relevance of a certain indicator, we consider it as validated by the panel of experts.

In addition to validation, it would be interesting to follow a subsequent process that would detect gaps or repetitions among the indicators or allow them to be written in a more functional way for their use in teaching. We refer to a process that would include the design of tasks for university mathematics students in which it would be possible to detect the presence of such indicators.

The questionnaire of the third, and final, phase asked for the rating of indicators that have not been validated nor rejected in the previous phase. As Andranovich (1995) and Scheibe et al. (1975) recommend, we included all the previous ratings of these indicators in the questionnaire. For the indicators that have received relevant comments in the second phase, we provided an alternative wording to be rated.

Sample

According to Hsu and Sandford ([2007](#)), the panel of experts' must be a heterogenic sample of people with expertise or experience in the concerning issue. In this regard, the sample was designed to be representative of the geometry experts' population. Specifically, we have considered only mathematicians with expertise in geometry and topology. Concretely, the whole sample of our study consisted of 25 university teachers, all of them holding a PhD in geometry and topology. A small group of the respondents (5 out of the 25 teachers) had prior experience on mathematics education having held positions at the mathematics education university department or in secondary education. Thus, our sample combines the required expertise in mathematics and teaching experience at different educational levels.

The respondents to our questionnaires belong to different European universities, mostly in Spain. The reasons to choose these respondents were related to the following: accessibility, they were easy to contact with through their university websites; relevance of their research, the respondents belong to research groups with an important presence in conferences such as ICM[2] and a large number of indexed publications; heterogeneity of their research, the respondents belong to research groups focused on different sub-fields of geometry: algebraic geometry, differential geometry …; and commitment, they agreed to answer the different questionnaires and to participate in the study until its completion; in some of the cases, this commitment had to do with the professional relation with the authors, and note that the Delphi literature stablishes the importance of a motivated panel in order to obtain high-quality responses (Andranovich, [1995](#)).

We have not found worldwide or European statistics on the number of geometry and topology teachers at tertiary level. However, to gain some understanding of the size of the intended population, we considered that in the last 30 years, only 453 thesis were defended in Spain related to the area of geometry and topology.[3] Moreover, in this country, there are only around 200 university teachers of geometry and topology.[4]

The sample is balanced in terms of gender and university position (see Table [1](#)). Considering the teaching area of the experts, there are 15 teachers of geometry and topology, 7 of other areas of mathematics (mainly applied mathematics), and 3 part-time university/secondary school teachers. Out of the 25 teachers, 19 are currently enrolled in research projects (with research papers published in the last 5 years in the area of geometry and topology), while the other 6 are mainly dedicated to teaching or management tasks.

---

[2] ICM: International Congress on Mathematics. Spanish participants in the last three conferences were 2.37% of participants, whereas the Spanish population is 0.61% of the world population.

[3] The Spanish Department of Universities supports the Teseo database which contains the data of any doctoral thesis read in Spain since 1977. In the last 30 years, 453 thesis were read in Spain containing the UNESCO codes 120400 (geometry) or 121000 (topology) respectively. https://www.educacion.gob.es/teseo/irGestionarConsulta.do

[4] The Spanish Department of Universities supports the University Information System (SIIU). In the 2020/2021 academic year, there were 218 teachers working in the disciplines of geometry and topology in Spanish universities. https://n9.cl/jyz1k

Table 1 Distribution of the sample by gender and position

	Men	Women	Total
Assistant professors or equivalent position	6	4	10
Associate professors or equivalent position	3	5	8
Professors or equivalent position	4	3	7
Total	13	12	25

The panel of experts involved two different groups, namely A and B, who were formed by 14 and 11 experts respectively. Group A participated along the three phases of the study while group's B participation started at the second phase. This incorporation of members to the panel of experts during the different phases of the study helps to increase the validity and reliability of the results (Linstone & Turoff, 1975; Muñiz-Rodríguez et al., 2017). Two of the participants did not answer at the third phase of the study and thus, the number of respondents has been different at every phase: 14, 25 and 23 respectively.

Results

First Phase

By condensing all the respondents' answers to the first phase questionnaire, we got to formulate a list of indicators (see Table 2). These indicators express potential competences of a person reasoning at level 5.

After receiving the answers, the researchers performed a process of classifying the 81 sentences according to their content on each of the four processes: 27 for definition, 25 for proof, 15 for classification, and 14 concerning identification.

In the case of the definition process, the comments were classified forming six indicators (see Table 2). We describe how the definition indicator Def2[5] was built according to the three comments in Fig. 2. Respondent 9 (R9) explains that, in different environments the defined object can have different properties. Something similar is pointed out by R4 but using the term *mathematical structures-theories* instead of *environments*. Finally, R1 focuses on the similarities and differences of the same definition when considered in different axiomatic systems. Thus, indicator Def2 tries to summarize that people at level 5 are able to understand that the properties of a defined object depend not only on its definition but also on the geometric context[66] By geometric context, we mean the different fields in which we can work in geometry. For

---

[5] Indicators are labeled as Defx, Prox, Clax, and Idex for definition, proof, classification, and identification processes, respectively.

example are different metrics, axiomatic systems, spaces, or geometries. We have chosen this expression to encompass others that have appeared in the answers such as "domains," "structures," "categories," or "theories". in which it is defined.

This indicator can be exemplified as follows: If we consider the definition of the sphere as the set of points of R^3 which are at the same distance of a certain point, indicator Def2 shows that someone at level 5 is able to understand that the properties of such object would depend, not only on its definition, but also on the metric (Euclidean, Taxicab …).

Concerning the proof process, we obtained 25 comments in phase 1 that were summarized into five indicators (see Table 2). In order to illustrate the indicator construction process, we describe how the indicator Pro5 was built (see Fig. 3). Respondent R6 remarked that proving at level 5 can be done even if the definitions of new concepts are required. Another respondent, R9, described that a person at this level is able to structure a demonstration into partial interesting results as lemmas. Finally, respondent R11 highlighted that a relevant characteristic of level 5 proofs was to decompose long proofs into definitions, lemmas, and other proofs. From all these comments, we designed indicator Pro5 based on the idea that proofs at this level can be structured introducing different lemmas or definitions.

The idea of structure a proof into different interesting results such as lemmas is very common among professional mathematicians. There also exist many examples of mathematical objects which have been defined with the purpose to be used or structure properly a theorem or its proof. A classic example of this idea in differential geometry is Ado's Theorem which states that every Lie algebra admits a finite representation. This result can be found in classical works of geometry such as Varadarajan (1984) or Lee (2013, p. 199). In both cases, it can be noticed that the definition of finite representation appears only a few lines before the theorem, showing that it has been introduced only because it is required to state and prove the theorem.

In the case of the classification process, we obtained 15 comments that were classified forming four indicators (see Table 2). Respondent 1 highlighted the competence of comparing classifications in different geometries while respondent 13 stated that the careful specification of the equivalence relation is a key element in this process at fifth level. From these comments, we designed indicator Cla2 pointing to the comprehension of the relativity of the classification process with respect to the geometric context (see Fig. 4).

This indicator tries to underline the importance of the geometric context when classifying objects. For example, as R13 points out, from a differential point of view, an ellipsoid and a sphere are the same object while this is no longer true if we consider a metric point of view since the curvature of both objects is different.

In the case of the identification process, we obtained 14 comments that were classified forming four indicators (see Table 2). As a matter of example, R3 commented on the relevance of the Reidemeister's movements to identify knots while R8 considered that some objects can be identified by studying the possibility of constructing new objects from the given ones via, for example, fiber bundles. From these comments, we designed indicator Ide1 expressing the ability of recognizing objects through transforming processes (Fig. 5).

Table 2 Phase 1 results. Indicator's list

Definition

Def1. Constructs and uses definitions in different axiomatic systems

Def2. Understands that defining a given mathematical object is not absolute, but is an action relative to the geometric context in which one works, implying for example that the defined object may have different properties in each context

Def3. Defines new objects, for example, because it may be necessary to generalize existing ones or to prove a statement

Def4. Understands that a definition arises out of the necessity to introduce a new mathematical object or to emphasize a property

Def5. Compares equivalent definitions to choose the most *interesting* one,[a] depending on the work to be done

Def6. Constructs new definitions of known objects

Proof

Pro1. Constructs proofs in different geometric contexts

Pro2. Understands that proving a given mathematical result is not absolute, but is an action relative to the geometric context in which one is working, implying for example that a proof may, or may not, be transferable in whole or in part to another geometric context

Pro3. Compares proofs on the basis of different criteria, e.g. for the possibility of using them to prove more general results

Pro4. Performs hybrid formulation-proof processes by means of which one not only proves results, but also shapes them[b]

Pro5. Structures a proof by recognizing partial results that can be of use in another proof and to which he gives his own entity in the form of lemmas, definitions …

Classification

Cla1. Classifies mathematical theories, e.g. to see Euclidean geometry as a particular case of a family of geometries

Cla2. Understands that classifying a given mathematical object is not something absolute, but an action relative to the geometric context in which one works, implying for example that the equivalence relationship between geometric objects varies from one context to another

Cla3. Establishes new classifications of mathematical objects, in addition to using already elaborated classifications

Cla4. Compares different classifications to choose the most interesting one, depending on the work to be done

Identification

Ide1. Identifies geometric objects through processes that transform the given object into an equivalent object that is directly recognizable

Ide2. Understands that recognizing a given mathematical object is not absolute, but is an action relative to the geometric context in which one works, implying for example that the properties that characterize the object vary from one context to another

Ide3. Identifies objects by analyzing invariants associated with them

Ide4. Intuits what aspects of an object will help the most in its recognition

[a]By *interesting definition* or *classification*, we mean, according to the panelist's comments, the one that best suits his/her interests, e.g. the one that allows subsequent reasoning to be simpler. For instance, in some situations, it is *more interesting* to use the definition of continuity in terms of the pre-image of an open set than use the delta-epsilon definition. Regarding classification, in some situations, it is interesting to classify surfaces attending to geometrical criteria, like curvature, while in other situations, it can be more useful to classify surfaces attending to topological criteria, like fundamental group, betti numbers, etc. …

[b]By shaping a result, we mean change the initial statement, for instance, to limit its validity. It is usual to include in a theorem statement some initial conditions because of their necessity in the proof process. For example, the Banach fixed-point theorem includes a condition on the completeness of the space which is a technical requirement to complete the proof.
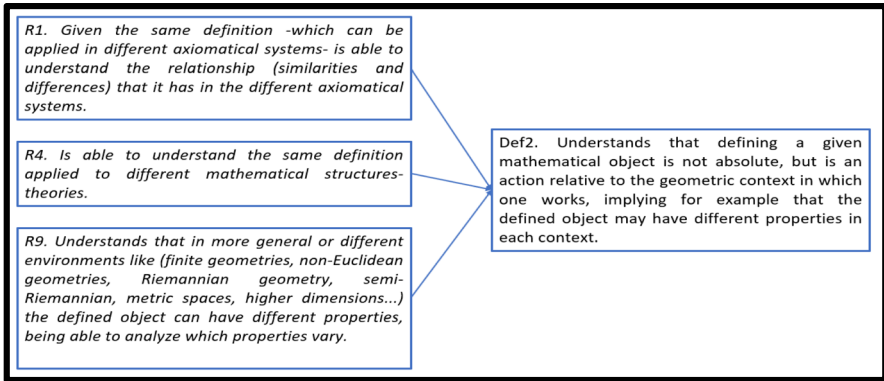
Fig. 2 Design of a definition indicator

There are several examples of these transformation processes that can be identified in advanced geometric work. For example, in the mathematical area of knot theory, a knot can be represented in very complicated manners; however, the Reidemeister's moves consist of transformations that preserve the knots, and maybe simplify its presentation which eventually allows us to identify them more easily. Another example, in the field of Riemannian geometry, consists on identifying a six-dimensional manifold (M,g) as Calabi-Yau if the product manifold $N = M \times R$ endowed with the cone metric $h = g + dt^2$ (with t the coordinate in R) is a parallel G_2 manifold (Boyer & Galicki, 2007, p. 544), which can be recognizable in terms, for example, of the behavior of its fundamental 3-form.

We can see in Table 2 the initial list with the 19 indicators obtained after the first phase: The processes of definition and proof have been described through six and five indicators respectively while the processes of classification and identification have been described only with four indicators each.

Notice that, from the revision of the 19 indicators, we can observe that related competences or topics appear in different processes. There are three indicators related to the comparison of definitions, proofs, or classifications. There are indicators highlighting the relevance of different contexts in all the processes. Finally, we have designed several indicators that point to the conceptual differences of the
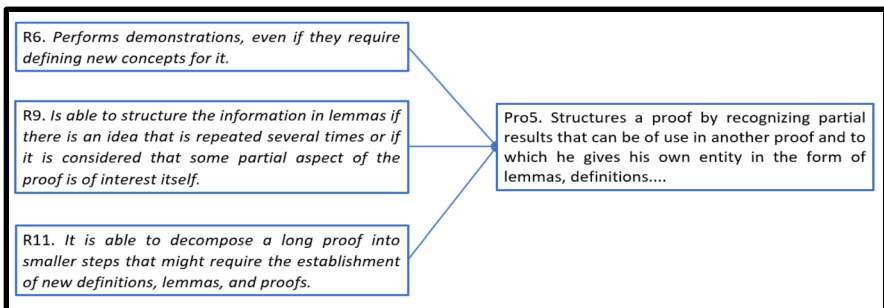

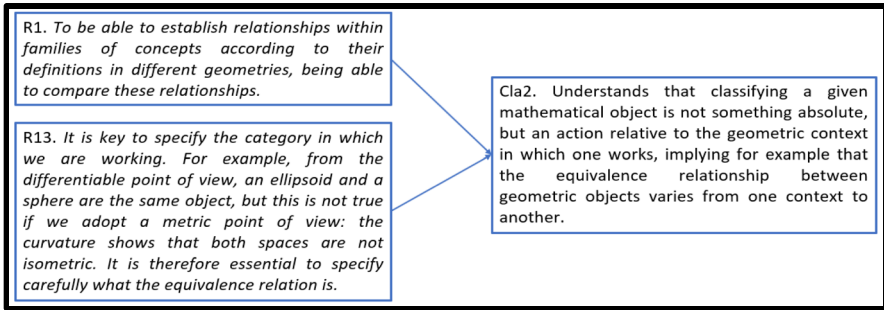**Fig. 3** Design of a proof indicator

Fig. 4 Design of a classification indicator

processes with the previous levels. Among others, be aware of the necessity to define in order to introduce a new mathematical object (Def4) or the ability to perform mixed formulation-proof processes (Pro4).

Second Phase

After designing the list of indicators, we administered another questionnaire to assess their relevance to describe the fifth level. We present now the different indicators with the percentage of persons awarding the maximum degree of relevance to each of them (Table 3).

Some of the 13 indicators to be reevaluated received comments suggesting changes, which led to their rewriting (see Table 4). For percentages between 40 and 70%, there were two different cases: If the commentaries of the respondents gave enough information, we designed a new indicator that was included in the third phase questionnaire along with the original one; otherwise, we included only the original indicator. As a matter of example, the indicator Pro1 for the proof process was rewritten in a similar form but including some modifications due to the respondents' comments: Since the comments expressed doubts about the meaning of *geometric context*, we included some examples in the new indicator Pro1.1 (see Table 4).

During this phase, the panelists provided information about their concerns; thus, in order to illustrate the Delphi process, we describe some of the information of the panel's reasoning in this phase. About Pro1 and Pro2, some of the respondents stated that, on the one hand, these items would correspond to Van Hiele level 4. Regarding Pro4, the whole item was not clear to some respondents; in particular, the expression "…but also shapes them" was not well understood. With respect to the idea of the relevance of different geometric contexts, respondents agreed with Cla2 but not with Def2: Some respondents noted that the question is not about an object having different properties in different contexts but about how different contexts focus their attention in different types of properties.

For a better understanding of the modified indicators, we provide now some examples. A way to illustrate Ide2.1. consists of thinking in the identification of a circle using Taxicab or Euclidean metric. If we establish the definition of circle as the set of points which are at the same distance from a given point, the shape of this
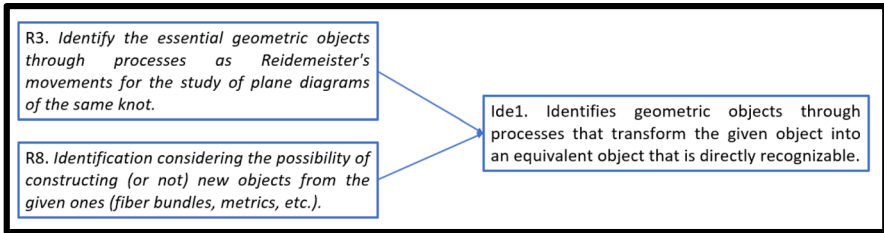
Fig. 5 Design of an identification indicator

object will depend on the metric of the ambient space. Concretely, if we consider the Taxicab metric, this object looks like a square while if we consider the Euclidean metric, it looks like a "classical" circle.

For the case of Pro2.1. it can also be exemplified using the comparison between Taxicab and Euclidean metrics. For example, the classical proof of the result that the sum of the three angles of a triangle is 180° can also be done using Taxicab metric due to the fact that the only requirement of that proof is that 180° spins are isometries, which is also satisfied with the Taxicab metric despite the fact that isometries in this later metric are not the same as the Euclidean ones.

Third Phase

Out of the original 19 indicators, 13 needed to be reevaluated in the third phase (see Table 3). Apart from these, 5 indicators were re-written (see Table 4). Thus, the third phase questionnaire consisted of 18 indicators. We present in Table 5 the results obtained in this final phase. Note that all the percentages increased in the third phase. However, since some of them were still lower than 70%, this was not enough to validate all the proposed indicators.

For the definition process, there were 6 indicators at the questionnaire proposed in the second phase. With the results of the second phase, one of them was rejected and two were validated. The other three indicators were validated after the third phase questionnaire.

Considering the proof process, there were 5 indicators at the questionnaire proposed in the second phase. With the ratings obtained in that phase, one was validated and the other four were proposed for a reevaluation. Once we analyzed the comments, indicators Pro1, Pro2, and Pro4 were proposed with an alternative writing (Pro1.1, Pro2.1, and Pro4.1) along with indicator Pro3. According to the third phase results, indicators Pro2.1, Pro3, and Pro4.1 were validated. Indicators Pro1 and Pro1.1 were classified as non-validated since they did not get more than 70% of high-relevance rating. Note that indicators Pro2 and Pro2.1 had the same percentage, so we decided to choose Pro2.1 due to the presence of comments of respondents expressing their preference for it.

For the classification process, there were 4 indicators at the questionnaire proposed in the second phase (see Table 2). With the results of the second phase, one of them was validated and three were proposed for a reevaluation. Since there were

Table 3 Percentage of persons awarding the maximum degree of relevance—phase 2

Indicator	Percentage of high-relevance ratings	Decision
Def1	72.0%	Validated by the panel
Def2	68.0%	Reevaluate
Def3	80.0%	Validated by the panel
Def4	60.0%	Reevaluate
Def5	64.0%	Reevaluate
Def6	28.0%	Reject
Pro1	64.0%	Reevaluate and propose Pro1.1
Pro2	64.0%	Reevaluate and propose Pro2.1
Pro3	68.0%	Reevaluate
Pro4	56.0%	Reevaluate and propose Pro4.1
Pro5	76.0%	Validated by the panel
Cla1	64.0%	Reevaluate
Cla2	72.0%	Validated by the panel
Cla3	52.0%	Reevaluate
Cla4	48.0%	Reevaluate
Ide1	56.0%	Reevaluate and propose Ide1.1
Ide2	48.0%	Reevaluate and propose Ide2.1
Ide3	36.0%	Reject
Ide4	52.0%	Reevaluate

Table 4 Modified indicators according to respondents' comments

Proof
Pro1.1. Constructs proofs in different geometric contexts (projective geometry, algebraic geometry, differential geometry …)
Pro2.1. Is able to consider whether or not a proof can be totally or partially transferable to another geometric context, understanding that proving a certain mathematical result is an action linked to the geometric context in which one works
Pro4.1. Performs mixed formulation-proof processes, eventually being able to modify the statement according to the development of the proof
Identification
Ide1.1. Identifies geometric objects through processes that transform the given object into an equivalent object that is recognizable
Ide2.1. Understands that the identification of a given mathematical object is not absolute, but is an action relative to the geometrical context, since if an object is changed in context, it will probably become a different object

no comments on this process, we could not propose a rewriting of these three indicators. According to the third phase results, indicator Cla1 was validated, while indicators Cla3 and Cla4 were classified as non-validated.

Considering the identification process, there were 4 indicators at the questionnaire proposed in the second phase. With the results of the second phase, one was

rejected and the other three were proposed for a reevaluation, two of them along with an alternative rewriting. According to the third phase results, indicator Ide1.1 was validated, while the rest of the indicators were classified as non-validated.

We present in Table 6 the list of validated indicators. We can observe the relevance of definition and proof processes in comparison with the rest of them.

Discussion

Our results highlight the processes of definition and proof as the most relevant in Van Hiele's level 5. However, classification and identification processes seem to play a certain role at this level. In quantitative terms, the processes of definition and proof have been described through five and four indicators respectively while the processes of classification and identification have been described by two and one indicators respectively.

In the definition process, the use and formulation of definitions are usually studied separately. At this level, the indicators that have appeared are mainly related to the formulation of definitions. These five indicators can be clustered into two different groups: on one hand indicators Def1 and Def2, and on the other hand indicators Def3, Def4, and Def5.

Indicators Def1 and Def2 form the first set of indicators which is related to the influence of working in diverse geometrical contexts. Indicator Def1 (*constructs and uses definitions in different axiomatic systems*) is a core part of the Van Hiele level 5 (Burger & Shaughnessy, 1986) and expresses the ability of working at different axiomatic systems, whereas indicator Def2 (*understands that defining a given mathematical object is not something absolute, but is an* action *relative to the geometric context in which one works, implying for example that the defined object may have different properties in each context*) points out the relevance of reasoning across several geometries as opposed to within a particular geometry (Blair, 2004). In this regard, Silfverberg (2019) states that, a person reasoning at level 5 can investigate Euclidean-equivalent concepts in geometries such as spherical or Taxicab which means that this person can move between different axiomatic systems.

Indicators Def3, Def4, and Def5 are related to the reasons that lead to formulating new definitions or choosing between definitions. Concerning the formulation of new definitions, Martín-Molina et al. (2018) point out: "a popular activity in the field of differential geometry consists of defining new spaces by taking a well-known definition and generalizing it" (p. 1076). Several of our respondents suggested similar ideas leading to the construction of indicator Def3 (*defines new objects, for example, because it may be necessary to generalize existing ones or to prove a statement*). In their study, Martín-Molina and colleagues identified four phases in the defining process (finding an opportunity to generalize an existing concept, proposing a new definition, justifying that the new definition is valid, and continuing the chain of definitions). Notice that while Def3 can be associated with the generalization phase, indicator Def4 (*understands that a definition arises from the need to introduce a new*

Table 5 Percentage of persons awarding the maximum degree of relevance—phase 3

Indicator	Percentage of high-relevance ratings	Decision
Def2	73.9%	Validated by the panel
Def4	73.9%	Validated by the panel
Def5	82.6%	Validated by the panel
Pro1	69.6%	Non-validated
Pro1.1	47.8%	Non-validated
Pro2	73.9%	Validated by the panel
Pro2.1	73.9%	Validated by the panel
Pro3	82.6%	Validated by the panel
Pro4	69.6%	Non-validated
Pro4.1	73.9%	Validated by the panel
Cla1	78.3%	Validated by the panel
Cla3	56.6%	Non-validated
Cla4	56.6%	Non-validated
Ide1	65.2%	Non-validated
Ide1.1	78.3%	Validated by the panel
Ide2	65.2%	Non-validated
Ide2.1	60.9%	Non-validated
Ide4	56.5%	Non-validated

mathematical object or to emphasize a property) highlights the reasons that lead to proposing a new definition.

Once the need for a definition of a concept has been established, it may be necessary to choose among several definitions of the same concept depending on the objectives; this need is reflected in the indicator Def5 (*compares equivalent definitions to choose the one that interests him/her most, depending on the work to be done*). Kemp and Vidakovic (2021) report on the problems faced by students enrolled in a college geometry course when they had to extend the concept of midset from Euclidean to Taxicab geometry. To solve this problem, they needed to compare and analyze the differences between two equivalent definitions of midset in order to determine which one could be transferred to Taxicab geometry.

Regarding the proof process, we have found four different indicators (five if we consider Pro1[7]) which can be grouped as Pro1 and Pro2.1 on one hand and Pro3, Pro4.1, and Pro5 on the other hand.

Indicators Pro1 and Pro2.1 form a first set of indicators which is related to the construction of proofs in different geometric contexts. Indicator Pro1 (*performs proofs in different geometric contexts*) is a fundamental part of Van Hiele's level 5 (Burger & Shaughnessy, 1986) and expresses the ability of working at different axiomatic systems. Indicator Pro2.1 (*is able to consider whether or not a proof can be totally or partially transferable to another* geometric *context, understanding that*

---

[7] Pro1 has a percentage slightly lower than 70% (69.6%) but we considered it interesting to mention it in this section due to its presence in the description of the level in classical works.

Table 6 Indicators validated by the panel describing the fifth Van Hiele level

Definition

Def1. Constructs and uses definitions in different axiomatic systems

Def2. Understands that defining a given mathematical object is not absolute, but is an action relative to the geometric context in which one works, implying for example that the defined object may have different properties in each context

Def3. Defines new objects, for example, because it may be necessary to generalize existing ones or to prove a statement

Def4. Understands that a definition arises out of the necessity to introduce a new mathematical object or to emphasize a property

Def5. Compares equivalent definitions to choose the most interesting one, depending on the work to be done

Proof

Pro2.1. Is able to consider whether or not a proof can be totally or partially transferable to another geometric context, understanding that proving a certain mathematical result is an action linked to the geometric context in which one works

Pro3. Compares proofs on the basis of different criteria, e.g. for the possibility of using them to prove more general results

Pro4.1. Performs mixed formulation-proof processes, eventually being able to modify the statement according to the development of the proof

Pro5. Structures a proof by recognizing partial results that can be of use in another proof and to which he gives his own entity in the form of lemmas, definitions …

Classification

Cla1. Classifies mathematical theories, e.g. to see Euclidean geometry as a particular case of a family of geometries

Cla2. Understands that classifying a given mathematical object is not something absolute, but an action relative to the geometric context in which one works, implying for example that the equivalence relationship between geometric objects varies from one context to another

Identification

Ide1.1. Identifies geometric objects through processes that transform the given object into an equivalent object that is recognizable

proving a given mathematical result is an action linked to the geometric context in which one works) is related to the importance of reasoning across several geometries as opposed to within a particular geometry. In relation to this indicator, Blair (2004) documented instances when mathematics students compared geometries by attending to whether the proof of a theorem in Euclidean geometry could be transferred or adapted to Taxicab or spherical geometry.

Indicators Pro3, Pro4.1, and Pro5 form a second set of indicators related to the knowledge of the internal structure of the proof: indicator Pro3 (*compares proofs on the basis of criteria that are of interest to them, e.g. the possibility of using them to prove more general results*). In this regard, Fernandez-León et al. (2020), in a study with professional mathematicians, identified the ability of selecting between different proving methods to prove the same result. This idea can be connected with Pro3 since this selection requires comparison according to a criterion.

During the proof process, the hypotheses or the conclusions of a proposition can be found not adequate and it can be decided to generalize or to limit its validity; this has been expressed by our respondents leading to the construction of indicator Pro4.1 (*performs mixed formulation-proof processes being able to eventually modify the statement he/she is trying to prove depending on the development of the proof*). This fact has been expressed by Boero (1999) by distinguishing between mathematical proofs as a product, and proving as a process. He stated that the proving process consists of six phases which are usually interconnected in non-linear ways in mathematicians' normal work. Based on these ideas, Fernandez-León et al. (2020) found two activities that connect to Boero's ideas: (i) modifying statements and (ii) formalizing modifications of statements related to modifying the hypotheses or conclusions of a theorem.

Indicator Pro5 (*structures the proof recognizing partial results that can be useful in another and to which it gives its own entity in the form of lemmas, definitions*) is of great interest at this level and some of the ideas described on it have already been pointed out by several authors. Concerning the identification of partial results, Weber and Mejia-Ramos (2011) showed that mathematicians will eventually construct sub-proofs to verify that each line in a proof follows from previous assertions. Also, Fernández-León et al. (2020) stated that it is common to search for proof techniques or tools used in other proofs that may fit in well with a new proof. Attending the construction of definitions during the proving process, Larsen and Zandieh (2005) showed an example of a proof analysis that supported the development of a new concept: the idea of a *small triangle.*

The classification process has only two validated indicators: Cla1 (*classifies mathematical theories, for example, to see Euclidean geometry as a particular case of a family of geometries*) and Cla2 (*understands that classifying a given mathematical object is not something absolute, but an action relative to the geometric context in which one works, implying for example that the equivalence relation between geometric objects varies from one context to another*). The former indicator is related to the faculty of comparing axiomatic systems; in this sense, Silfverberg (2019) states that, at the highest Van Hiele level, different geometries can be compared by observing their differences and similarities as axiomatic systems.

The identification process has only one validated indicator: Ide1.1 (*identifies geometric objects through processes that transform the given object into an equivalent one that is directly recognizable*). For instance, in differential geometry, it is common to identify manifolds via processes like warped products (different metrics) of the original manifold (Boyer & Galicki, 2007) or considering the latter as the fiber of a certain fiber bundle like for example in the mapping torus constructions.

Based in our results, we consider that the Van Hiele level 5 indicators are consistent with the previous descriptions found in classical works such as Jaime and Gutiérrez (1994) where they give the greatest relevance to the proof and definition processes at the highest levels. Moreover, we have stated in our discussion the close relation between both processes at this level. This relationship can be seen in indicators Pro5 and Def3 which explain how a proving process can lead to a definition of a new object while a definition can help to prove a result.

On the other hand, working with a combination of different geometries has resulted in the re-appearance of indicators related to the processes of identification and classification; note that level 4 is presented only in terms of definition and proof processes (Gutiérrez & Jaime, 1995). The Van Hiele model proposes that mathematical objects become explicit (Van Hiele, 1986) as higher levels are reached (e. g., when a person achieves level 2, the elements of a geometrical object become explicit in contrast with the global appreciation of these objects at level 1). In the fifth level, the possibility of working across geometries becomes explicit (in contrast with a somehow internal work in previous levels) giving a new meaning to the classification process. This meaning includes two different aspects: On the one hand, we can classify different geometric contexts (e.g. Euclidean, hyperbolic, and spherical geometries), and on the other hand, we can understand that different classifications of geometric objects are provoked by different equivalence relationships (the classification of surfaces in geometry depends on the diffeomorphism equivalence while in topology the equivalence relation is given by the homeomorphism).

Concerning the identification process, the mastery of the relationships between different geometries, metrics … may allow that, in certain cases, the identification at this level is carried out by working with a different metric from that of the initial problem and then somehow transferring the results to the original metric. In both processes, it is fundamental to work between geometries, which is a distinctive feature of level 5.

Another classical work (Burger & Shaugnessy, 1986) established that, at level 5, systems based on different axioms can be studied and compared. The obtained results in this study agree with this idea; in particular, indicators Def1, Def2, Pro1, Pro2.1, Cla1, and Cla2 focus on the relationship between geometric contexts more than on the geometric contexts considered individually. This idea is related with Blair's (2004) considerations: "…reasoning within Taxicab or spherical (or any other geometry for that matter) characterized according to the complexity of the students' object of thought is not inherently more or less 'advanced' as that within Euclidean geometry" (p. 337). In this regard, Guven and Baki (2010) highlighted that working with spherical geometry has similar difficulties to working with Euclidean geometry. In particular, they were able to show a certain parallelism between Van Hiele's levels of reasoning with spherical geometry and Euclidean geometry.

Some limitations have to be discussed. Firstly, the answers of the respondents could depend on their particular field of expertise. It seems reasonable that experts in differential geometry and algebraic topology would give different ratings to the same indicator. Moreover, if the panelists in future studies were researchers in different areas of geometry, the examples they would propose would be more varied. Secondly, in order to obtain more information and to go deeper into the particular reasons that lead to value each indicator in one way or another, it could be of interest to include interviews with the panel members in such studies. A third limitation can be found in the application of the Delphi method. Instead of using the number of phases as a termination criterion, we could have decided to study the stability of the ratings. Probably, this would have produced more reliable results about the validity of indicators such as Pro1 with percentages really close to 70%. However, this could have led to a higher number of phases with the risk of losing respondents in the final phases.

Our future research interests include the design of a questionnaire including mathematical tasks in which the obtained descriptors of the fifth level could be identified among the responses. The implementation of such a questionnaire with college/university mathematics majors and the subsequent analysis of their productions could give empirical support to the indicators obtained in this study. This future work could provoke a refinement of the indicators as well as a deeper validation of them.

On the other hand, our future research also includes the study of how level 5 is related to previous levels and how including activities at that level can improve the teaching of geometry. In particular, Pirie and Kieren (1991) established that a person can recur to inner (lower) level of understanding when challenged by outer (higher) level activities. This process, called folding back, "allows for the reconstruction and elaboration of inner level understanding to support and lead to new outer level understanding" (p. 172). This leads to the interest in determining the characteristics of fifth level reasoning which would allow to create activities adapted to such level that can promote the fully acquisition of the fourth level. This research interest is related to previous works of the authors (Arnal-Bailera & Manero, 2021; Manero & Arnal-Bailera, 2021) where we found that graduated students enrolled in teacher training programs had not completely acquired the fourth level. Our work can also facilitate different current research lines as comparative studies between different groups (in-service teachers at different levels, mathematics majors …) with respect to aspects such as demonstration following those initiated by Buchbinder et al. (2022) or university student's interactions when working with non-Euclidean geometries (Kaisari & Patronis, 2010).

Declarations

Competing Interests The authors declare no competing interests.

References

Abdullah, A. H., & Zakaria, E. (2013). The effects of Van Hiele's phases of learning geometry on students' degree of acquisition of Van Hiele levels. *Procedia-Social and Behavioral Sciences, 102*, 251–266. https://doi.org/10.1016/j.sbspro.2013.10.740

Alake-Tuenter, E., Biemans, H., Tobi, H., & Mulder, M. (2013). Inquiry-based science teaching competence of primary school teachers: A Delphi study. *Teaching and Teacher Education, 35*, 13–24. https://doi.org/10.1016/j.tate.2013.04.013

Andranovich, G. (1995). *Developing community participation and consensus: The Delphi technique*. California State University.

Arnal-Bailera, A. & Manero, V. (2021). La demostración en matemáticas. Perfiles de profesores en formación según sus niveles de Van Hiele [Proof in mathematics. Profiles of preservice teachers according to their Van Hiele levels]. In P. D. Diago, D. F. Yáñez, M. T. González-Astudillo, & D. Carrillo (Eds.), *Investigación en Educación Matemática XXIV* (pp. 149–156). SEIEM.

Blair, S. (2004). *Describing undergraduates' reasoning within and across Euclidean, taxicab and spherical geometries* [Ph.D. thesis]. Portland State University.

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof*. https://n9.cl/jyz1k

Boyer, C., & Galicki, K. (2007). *Sasakian geometry*. Oxford University Press.

Buchbinder, O., McCrone, S., Capozzoli, M., & Butler, R. (2022). Mathematical knowledge for teaching proof: Comparing secondary teachers, pre-service secondary teachers, and undergraduate majors. *International Journal of Research in Undergraduate Mathematics Education*. https://doi.org/10.1007/s40753-022-00187-8

Burger, W. F., & Shaughnessy, J. M. (1986). Characterizing the Van Hiele levels of development on geometry. *Journal for Research in Mathematics Education, 17*(1), 31–48. https://doi.org/10.2307/749317

Clarke, V., & Braun, V. (2016). Thematic analysis. *The Journal of Positive Psychology, 17*(3), 297–298.

Corberán, R., Gutiérrez, A., Huerta, M., Jaime, A., Margarit, J., Peñas, A., & 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 curricular proposal for learning Geometry in Secondary Education based on the Reasoning Model of van Hiele]. M.E.C.

Cox, J. (1996). *Your opinion please: How to build the best questionnaires in the field of education*. Corwin Press.

De Villiers, M. D. (1987). *Research evidence on hierarchical thinking, teaching strategies and the Van Hiele theory: Some critical comments*. University of Stellenbosch.

Demiray, E., & Işiksal, I. (2017). An investigation of pre-service middle school mathematics teachers' ability to conduct valid proofs, methods used, and reasons for invalid arguments. *International Journal of Science and Mathematics Education, 15*(1), 109–130. https://doi.org/10.1007/s10763-015-9664-z

Eckman, C. A. (1983). *Development of an instrument to evaluate intercollegiate athletic coaches: A modified Delphi study* [Ph.D. thesis]. West Virginia University.

Fernández-León, A., Gavilán-Izquierdo, J. M., & Toscano, R. (2020). A case study of the practices of conjecturing and proving of research mathematicians. *International Journal of Mathematical Education in Science and Technology, 52*(5), 767–781. https://doi.org/10.1080/0020739X.2020.1717658

González, A., Gallego Sánchez, I. M., Gavilán Izquierdo, J. M., & Puertas, M. L. (2021). Characterizing levels of reasoning in graph theory. *EURASIA Journal of Mathematics, Science and Technology Education, 17*(8), 1–16. https://doi.org/10.29333/ejmste/11020

González, A., Manero, V., Arnal-Bailera, A., & Puertas, M. (2022). Proof levels of graph theory students under the lens of the Van Hiele Model. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/10.1080/0020739X.2022.2113467

Green, R. A. (2014). The Delphi Technique in Educational Research. *SAGE Open, 4*(2), 1–8. https://doi.org/10.1177/2158244014529773

Guillén, G. (2004). El modelo de van Hiele aplicado a la geometría de los sólidos: Describir, clasificar, definir y demostrar como componentes de la actividad matemática [The van Hiele model applied to the geometry of solids: Describing, classifying, defining and proving as components of mathematical activity]. *Educación Matemática, 16*(3), 103–125.

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, 22*(3), 237–251.

Gutiérrez, A., & Jaime, A. (1995). Towards the design of a standard test for the assessment of the students' reasoning in geometry. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19th PME conference* (vol. 3, pp. 11–18). PME.

Guven, B., & Baki, A. (2010). Characterizing student mathematics teachers' levels of understanding in spherical geometry. *International Journal of Mathematical Education in Science and Technology, 41*(8), 991–1013. https://doi.org/10.1080/0020739X.2010.500692

Hill, K. Q., & Fowles, J. (1975). The methodological worth of the Delphi forecasting technique. *Technological Forecasting and Social Change, 7*, 179–192. https://doi.org/10.1016/0040-1625(75)90057-8

Hoffer, A. (1983). Van Hiele based research. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematical concepts and processes* (pp. 205–227). Academic Press.

Howse, T. D., & Howse, M. E. (2014). Linking the Van Hiele theory to instruction. *Teaching Children Mathematics, 21*(5), 304–313. https://doi.org/10.5951/teacchilmath.21.5.0304

Hsu, C. C., & Sandford, B. A. (2007). The Delphi Technique: Making Sense of Consensus. *Practical Assessment, Research & Evaluation, 12*(10), 1–8.

Jacobs, J. M. (1996). *Essential assessment criteria for physical education teacher education programs: A Delphi study* [Ph.D. thesis]. West Virginia University.

Jaime, A. & Gutiérrez, A. (1994). A model of test design to assess the Van Hiele levels. In J. P. da Ponte & J. F. Matos (Eds.), *Proceedings of the 18th PME conference* (vol. 3, pp. 41–48). PME.

Jaime, A., & Gutiérrez, A. (1990). Una propuesta de fundamentación para la enseñanza de la geometría: El modelo de van Hiele [A proposal for the foundation for the teaching of geometry: The van Hiele model]. In S. Llinares & V. M. Sánchez (Eds.), *Teoría y práctica en educación matemática* (pp. 295–384). Alfar.

Kaisari, M., & Patronis, T. (2010). So we decided to call "straight line"(…): Mathematics students' interaction and negotiation of meaning in constructing a model of elliptic geometry. *Educational Studies in Mathematics, 75*, 253–269.

Kemp, A., & Vidakovic, D. (2021). Ways secondary mathematics teachers apply definitions in Taxicab geometry for a real-life situation: Midset. *The Journal of Mathematical Behavior, 62*, 100848.

Larsen, S., & Zandieh, M. (2005). Conjecturing and proving as part of the process of defining. In G. M. Lloyd, M. Wilson, J. L. M. Wilkins, & S. L. Behm (E.ds.), *Proceedings of the 27th PME-NA* (pp. 797–804). PME-NA.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67*(3), 205–216. https://doi.org/10.1007/s10649-007-9106-0

Lee, J. M. (2013). *Introduction to smooth manifolds*. Springer.

Linstone, H. A. & Turoff, M. (1975). *The Delphi method: Techniques and applications.* Addison-Wesley.

Ludwig, B. (1997). Predicting the future: Have you considered using the Delphi methodology? *Journal of Extension, 35*(5), 1–4.

Manero, V., & Arnal-Bailera, A. (2021). Understanding proof practices of pre-service mathematics teachers in geometry. *Mathematics Teaching Research Journal 13*(3), 99–130.

Martin, L. C. (2008). Folding back and the dynamical growth of mathematical understanding: Elaborating the Pirie-Kieren theory. *The Journal of Mathematical Behavior, 27*(1), 64–85.

Martín-Molina, V., González-Regaña, A. J., & Gavilán-Izquierdo, J. M. (2018). Researching how professional mathematicians construct new mathematical definitions: A case study. *International Journal of Mathematical Education in Science and Technology, 49*(7), 1069–1082. https://doi.org/10.1080/0020739X.2018.1426795

Mayberry, J. (1983). The Van Hiele levels of geometric thought in undergraduate preservice teachers. *Journal for Research in Mathematics Education, 14*(1), 58–69. https://doi.org/10.2307/748797

Miller, L. E. (2006). *Determining what could/should be: The Delphi technique and its application.* Paper presented at the 2006 annual meeting of the Mid-Western Educational Research Association, AERA.

Muñiz-Rodríguez, L., Alonso, P., Rodríguez-Muñiz, L. J., & Valcke, M. (2017). Developing and validating a competence framework for secondary mathematics student teachers through a Delphi method. *Journal of Education for Teaching, 43*(4), 383–399. https://doi.org/10.1080/02607476.2017.1296539

Pandiscio, E. A., & Knight, K. C. (2010). An investigation into the Van Hiele levels of understanding geometry of preservice mathematics teachers. *Journal of Research in Education, 20*(1), 45–53.

Pirie, S. E. B., & Kieren, T. E. (1991). Folding back: Dynamics in the growth of mathematical understanding. In F. Furinghetti (Ed.), *Proceedings of the fifteenth annual meeting of the International Group for the Psychology of Mathematics Education,* (vol. 3, pp. 169–176).

Scheibe, M., Skutsch, M., & Schofer, J. (1975). Experiments in Delphi methodology. In H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications* (pp. 262–287). Addison-Wesley Publishing Company.

Sears, R. (2019). Proof schemes of pre-service middle and secondary mathematics teachers. *Investigations in Mathematics Learning, 11*(4), 258–274. https://doi.org/10.1080/19477503.2018.1467106

Silfverberg, H. (2019). Geometrical conceptualization. In A. Fritz, V. Haase, & P. Räsänen (Eds.), *International handbook of mathematical learning difficulties: From the laboratory to the classroom* (pp. 611–630). Springer. https://doi.org/10.1007/978-3-319-97148-3_36

Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry* (ERIC Document Reproduction Service Nº ED 220 288). ERIC.

Van Hiele, P. M. (1957). De problematiek van het inzicht [The problem of insight] [Ph.D. thesis]. University of Utrecht.

Van Hiele, P. M. (1986). *Structure and insight*. Academic Press.

Varadarajan, V. S. (1984). *Lie Groups, Lie Algebras, and Their Representations*. Springer.

Wahab, R. A., Abdullah, A. H., Mokhtar, M., Atan, N. A., & Abu, M. S. (2017). Evaluation by experts and designated users on the learning strategy using sketchup make for elevating visual spatial skills and geometry thinking. *Bolema: Boletim de Educação Matemática, 31*(58), 819–840.

Wang, S., & Kinzel, M. (2014). How do they know it is a parallelogram? Analysing geometric discourse at Van Hiele Level 3. *Research in Mathematics Education, 16*(3), 288–305.

Weber, K., & Mejia-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329–344. https://doi.org/10.1007/s10649-010-9292-z

Witkin, B. R. & Altschuld, J. W. (1995). *Planning and conducting needs assessment: A practical guide*. Sage Publications, Inc.