

Evaluation of the Degree of Coherence Found in Students' Conceptions Concerning the Particulate Nature of Matter

Enrique Jiménez Gómez,¹ Alicia Benarroch,² Nicolás Marín³

¹*Dpto. de Didáctica de las Ciencias Experimentales, Facultad de Educación, Universidad de Murcia, 30100 Murcia, Spain*

²*Dpto. de Didáctica de las Ciencias Experimentales, Facultad de Educación y Humanidades de Melilla, Universidad de Granada, Spain*

³*Dpto. de Didáctica de las Matemática y de las Ciencias Experimentales, Facultad de Ciencias de la Educación, Universidad de Almería, 04120 Almería, Spain*

Received 28 December 2004; Accepted 1 July 2005

Abstract: Students' conceptions are characterized by some authors as having a high degree of coherence while, in the minds of others, they show little coherence and great heterogeneity. The objective of this study was to throw light on this problem by reference to *the particulate nature of matter*, a topic where great discrepancies have been observed in the degree of coherence shown by students. We interviewed 43 students aged 9–22 years to evaluate the coherence of their replies, using a questionnaire in which different methodological strategies are incorporated, as will be explained below. A qualitative analysis of the students' replies to the interviewer's questions permitted a group of empirical categories to be established, which were then arranged hierarchically. The quantitative analysis of these categories gave rise to new categories, which we call structural categories. These permitted us to select the most significant variables, to identify the different types of conceptions held by the subjects and to assign to each a given level of conceptualization. They also revealed the high degree of coherence in the replies given to different tasks. The usefulness of the proposed method for studying students' conceptions and for evaluating their degree of coherence is confirmed. © 2006 Wiley Periodicals, Inc. *J Res Sci Teach* 43: 577–598, 2006

The great research effort put into the study of students' conceptions has yielded a vast body of knowledge concerning their conceptions on different scientific topics, which has undoubtedly contributed to improving the way in which science is taught (Carmichael et al., 1990; Driver, Guesne, & Tiberghien, 1985; Hierrezuelo & Montero, 1991; Pfundt & Duit, 1994, among others).

Contract grant sponsor: DGICYT; Contract grant number: BS02001-0496.

Correspondence to: E. Jiménez Gómez; E-mail: ejimenez@um.es

DOI 10.1002/tea.20130

Published online 22 June 2006 in Wiley InterScience (www.interscience.wiley.com).

In their examination of students' conceptions, several authors, including Driver (1986; 1988), Driver et al. (1985), Osborne and Freyberg (1985), Pozo, Gómez Crespo, Limón, and Serrano Sanz (1991), have observed a series of regularities concerning, for example, how they are shared by subjects of different cultures and ages, are dominated by perception, are marked by spatial and temporal causal lineal reasoning, etc. (For a more extensive revision, see Marín, Benarroch, & Jiménez Gómez, 2000). However, contradictions arise when authors begin to evaluate these characteristics and others: for some the conceptions are coherent (Brown, 1989; Gamble, 1989; Ioannides & Vosniadou, 2002; Samarapungavan & Wiers, 1977; Watts & Zylbersztajn, 1981), while for others they are diffuse, poorly differentiated, and fragmented (Pozo et al., 1991; diSessa, Gillespie, & Esterly, 2004; diSessa & Sherin, 1998; Kuiper & Mondlane, 1994; Stavy & Tirosh, 1993).

Such contradictions may be due to the fact that the data used are based on whatever reply is given by the subject, because there is a temptation to accept "any reply" as valid and not to differentiate between the replies (Driver, 1988; Vosniadou & Brewer, 1992). The way in which the replies are obtained, the theoretical context of the investigator, the subject under investigation, etc., may also play a part here (Marín, Jiménez Gómez, & Benarroch, 2004).

Before it can be affirmed that not all the replies given by students are valid to describe their conceptions and, especially, to establish the degree of coherence of their replies referring to a given scientific topic, it is necessary to use a series of methodological strategies and statistical techniques to help discriminate between them (Marín et al., 2004). This is because some of the students' replies will arise from well-structured information and be related with certain schemes of knowledge that will remain unchanged in the face of a series of questions concerning their experiences, while other replies, either because the subjects have no appropriate scheme or because they do not use them, will lack regularity. The regularity of the first type of reply regarding the way in which they are coherently related will permit the investigator to assign a level of reasoning, which will reflect the degree of coherence of the subject's thinking.

Indeed, the aim of this study is to illustrate that students' thoughts are coherent but do not always appear so, either because the questions asked are too academic, the context provided seems strange or because the questions asked are simply too difficult. However, it is possible to find questions for a context that does not appear strange and that has a level of difficulty that is not too high that will permit the researcher to identify subjects whose replies are coherent.

A Glossary of the Terms Associated with the Concept of Coherence

The starting point of this study is that the cognitive system of the subject is constructed coherently during a process of self-regulation similar to that belonging to any living thing (Piaget, 1978). For the subject, his/her knowledge is coherent, useful, and effective to respond to the usual demands of daily life (Pozo, 2003). However, an observer may well think differently.

To understand fully the meaning of the term *coherence*, it is necessary to distinguish between *scheme* and *conception*. The first, *scheme*, refers to a construct that forms part of the unobservable cognitive network of the subject, while *conception* is used to refer to students' replies that show a certain degree of regularity and that are observable manifestation of their cognitive structure (see, e.g., Brown, 1989; Brumby, 1979; Marín et al., 2000; Meheut, Larcher, & Chomat, 1988; Nussbaum & Sharoni-Dagan, 1983; Terry & Jones, 1986). Any reply may represent a scheme of knowledge, although many replies may be given without the intervention of such schemes and may be used by the investigator to identify and describe students' conceptions (Halldén, 1999; Marín et al., 2000; McClelland, 1984; Vosniadou & Brewen, 1992).

The information provided by students, then, will take on more significance if the researcher can ask questions that are suited to and that involve the use of some scheme of knowledge or group thereof. If this is achieved, and because schemes are stable entities, the replies, too, will assume a degree of soundness. We shall speak, then, of coherence when a certain degree of regularity is perceived in the students' replies and in the relations and connections that they establish in the face of a variety of experiences involving a similar academic content, about which a series of increasingly difficult questions is formulated. Looked at in this way, the problem of coherence is of a methodological (although also theoretical) type, and it is necessary to discern between replies in which schemes of knowledge are involved and those where this is not the case.

Given this, incoherence, fragmentation, and confusion are characteristics of conceptions that will arise when students' knowledge is confronted by excessively academic questions and will be, in most cases, a "collateral" effect of deficiencies in the methodology used, rather than cognitive attributes of the subject (Marín et al., 2000; Oliva, 1886).

It is accepted that not all information provided by the student is equally significant, and that the significance will depend on the degree of implication of the subject's cognitive system. In this sense, any perturbation and conflict strategies applied in a questionnaire should widely involve the subject's cognitive system (Hewson & Thorley, 1989; Piaget, 1978; Posner et al., 1982). The replies obtained after such conflict or perturbation will provide much richer and more meaningful information than that obtained otherwise, which means that a series of methodological strategies and statistical techniques must be used that will permit different types of replies to be distinguished.

Methodology

Content

To describe students' conceptions and to reveal the degree of coherence of their answers we have used as subject the particulate nature of matter. There were two main reasons for choosing this subject: first, because we wished to improve students' performance in this topic, which they frequently find difficult (see, e.g., Benson, Wittrock, & Baur, 1993; Brook, Briggs, & Driver, 1984; Griffiths & Preston, 1992; Haidar & Abraham, 1991; Lee, Eichinger, Anderson, Berkheimer, & Blakeslee, 1993; Novick & Nussbaum, 1978, 1981; Snir, Smith, & Raz, 2003; among others); and second, because many authors have described students' conceptions on this subject as being coherent and therefore presenting a certain degree of regularity (Eilam, 2004; Nakhleh & Samarapungavan, 1999; Yair & Yair, 2004), while others have found the conceptions to be heterogeneous and of little coherence (DiSessa & Sherin, 1998; Gómez, Pozo, & Sanz, 1995; Pozo et al., 1991; Stavy, 1988, 1990; Stavy & Tirosh, 1993).

Methodological Strategies

To obtain significant information, a questionnaire-based personal interview is a good starting point. A double strategy involving *confrontation* and *contextual variation* was the basis of the questionnaire used here, which attempted to involve students' knowledge schemes to the greatest extent possible. At the same time, the level of difficulty (as regards cognitive demand) of the questions proposed for each of the contextual variations was gradually increased.

The first of these strategies, *confrontation*, basically comprises two stages:

1. *Prediction stage*, where the student is asked to anticipate possible outcomes for a physical fact or event.

2. *Confrontation stage*, where the researcher or student carries out an experiment (empirical verification), the results of which are confronted with the answers given in the previous stage.

The second strategy, *contextual variation*, consists of presenting different physical situations with the same underlying research topic. For each context there are two new types of variation:

1. *Relevant variations*, which are modifications of some factors involved in the situation, and which cause some kind of relevant alteration. Usually, but not necessarily, a relevant variation involves some change of context.
2. *Nonrelevant variations*, which are modifications of certain factors involved in the situation, which do not involve any kind of relevant alteration.

Similar conflict-producing techniques were used by Hewson and Thorley (1989), Piaget (1946), Piaget (1975, 1978), and Posner et al. (1982). For more detail, see Marín et al. (2004). The combination of the different strategies mentioned above, confronting and contextual variation, increases the possibility that students will apply their knowledge schemes more thoroughly when answering questions. The degree of regularity in the students' replies can be evaluated by using three criteria:

1. Repetition, or the extent to which the replies remain unaltered despite the modifications of the physical situations introduced using both nonrelevant and relevant variation strategies.
2. Generalization, as observed from analogous replies to the physical situations principally constructed from contextual variation strategies.
3. Adaptation of the replies to the factors intervening in the task (contextual variation and confrontation strategies).

Contextual variation and confrontation strategies allow more information to be gathered from students in such a way that the answers show signs of repetition, generalization, and adaptation. Note that both the strategies followed in the design and construction of the questionnaire used as protocol in the interview are based on the underlying model of cognitive organization, where the organization unit is a knowledge scheme with characteristics of stability and coherence (Case, 1983; Marín, 1994a, 1994b; Pascual-Leone, 1979; Piaget, 1978).

The subsequent analysis of students' replies should reveal two types of response: a first group, which shows certain regularities in so far as the replies are stable, resistant to change and generalizable to different tasks, and a second group, which varies in the face of relevant and nonrelevant factors. The first group shows signs of reflecting students' knowledge and should reveal students' conceptions and present a certain degree of coherence (Pozo & Gómez Crespo, 1998).

The Questionnaire

Below, we detail the questions posed to elicit replies concerning phenomena that can be explained by reference to the particulate nature of matter. The three tasks presented have been taken from a more extensive study (Benarroch, 1998).

For each task, we explicitly describe the confrontation strategy (prediction and confrontation stages). Regarding the contextual variations, the nonrelevant variations are explicitly described, while the relevant variations are implicit in each of the tasks.

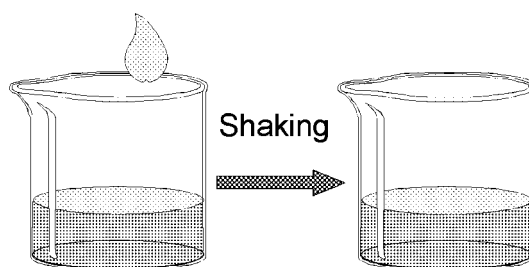


Figure 1. Where is the yellow drop?

Task 1: Solution of Granular Solute (Yellow Watercolor) in Water

Prediction Stage. The replies to this task provide the variable S_ ANTE (S refers to the solution, while ANTE refers to replies given before adding yellow watercolor). The beaker contains a yellow solution. The student is asked to draw what he/she would see through a powerful microscope inside the beaker containing the yellow colored solution.

After the drawing is complete, a small drop of yellow watercolor is added to another beaker of clear water: There is no appreciable change in color and the liquid remains transparent (Figure 1).

Confrontation Stage. The replies to this task provide the variable S_ POST (POST refers to replies given after adding a drop of yellow watercolor to the clear liquid). New drawings are requested for the solution (Questions: Where has the watercolor gone? Why isn't the liquid more yellow? What would you draw if you had a powerful microscope?)

Nonrelevant Variation 1. The investigator adds several drops of paint until the liquid looks yellow (Questions: Why is the water more yellow now? What would you draw if you had a powerful microscope?)

Nonrelevant Variation 2. The investigator puts a drop of green liquid colorant in a beaker containing approximately 50 cm^3 of water at room temperature and another in a glass containing the same quantity of hot water (Questions: Why does the green spread more quickly in the hot water? What would happen if the water was very cold? What would you draw if you had a powerful microscope?)

Task 2: Solution of Two Liquids (Alcohol and Water) Involving a Discernible Reduction in Height

The student is shown a test tube of approximately $80 \text{ cm height} \times 1.5 \text{ cm diameter}$, a bottle of distilled water and another of alcohol. The tube is half filled with water and then, tilting the tube slightly, the alcohol is added in such a way that it runs down the side of the tube. The student marks the level of the liquid by marker pen and the tube is closed with a cork.

Prediction Stage. The replies to this task provide the variable L_ ANTE (L refers to the height and ANTE refers to replies given before vigorously shaking the liquid). The students are

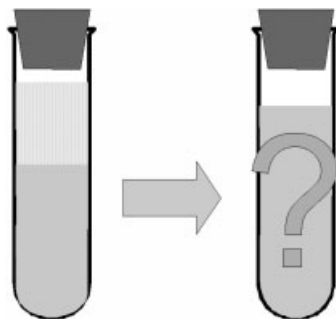


Figure 2. How do you explain that the weight hasn't varied but the height has?

asked to predict what will happen when the tube and contents are shaken vigorously (Questions: Will the quantities of the liquids vary? Will the weight change? Will the height of the liquid change?)

The tube is shaken, inverting it several times until the contents are totally homogenized. The height will be seen to have dropped 1–2 cm, while the principle of mass conservation will be checked (Fig. 2).

Confrontation Stage (L_POST). The student is asked to draw what has happened to the water and alcohol to justify the increased concentration that occurred during mixing (Questions: What has happened? Is it what you thought would happen? How do you explain that the weight hasn't varied while the height has? What would you see inside the tube before and after shaking if you had a powerful microscope?)

Nonrelevant Variation. Repeat the process with alcohol and colored water (same sequence of questions as in prediction and confrontation stage).

Task 3: Compressing Air and Water in a Syringe

The first part of this task involves the material character of gases by weighing a balloon before and after inflating. The student “plays” with the inflated balloon, pulling and squeezing it.

Prediction Stage. The replies to this task provide the variable G_ANTE (G refers to the compressibility of air/water, while ANTE refers to replies given before compressing the air/water in the syringe). The student is asked to draw and to explain how air is made up, bearing in mind its particulate composition and its high degree of compressibility (Questions: What would you see inside the inflated balloon before and after squeezing it if you had a powerful microscope? And in the syringe before and after pushing the piston? Is it possible to reduce the space occupied by the air?)

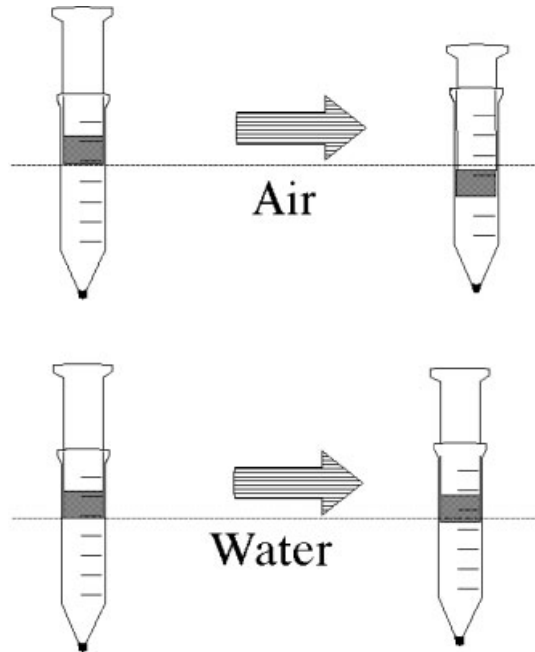


Figure 3. What's the difference between air and water?

Then, water/air is placed in the syringe and its incompressibility/compressibility is demonstrated (Fig. 3).

Confrontation Stage (G_POST). The student is asked for more drawings and explanations about the constitution of air and water after experimenting with their different degrees of compressibility (Questions: What is the difference between air and water? Draw how you would see the air and water before and after pulling the piston if you had a powerful microscope?)

Nonrelevant Variation. After producing a red gas in a Kitasato vessel, part is extracted by syringe and its compressibility is tested (Questions: Draw what you would see inside the Kitasato vessel if you had a powerful microscope before and after extracting part of the gas. Draw how you would see the gas in the syringe before and after pulling the piston.)

Data Gathering

The protocol, which is designed using *confrontation* and *variation* strategies, ensures that each student provides a great amount of information during the personal interviews. Videotape is advisable to extract all the necessary data for the research. We selected 43 students with a broad age range (9–22 years old) in an attempt to obtain a representative sample of possible replies. This number (43) is justified by the fact that above 33 the answers given were repetitions of those previously registered and did not provide any new information. However, the further 10 students interviewed provide a certain degree of security that no new empirical categories or modifications appear (see the fragment of the interview in the Appendix, in which a student is looking at task 2). A similar procedure for selecting the number of subjects was suggested by Taber (2000).

Qualitative Treatment of Data: Identification and Description of Empirical Categories

A transcription of the video recordings made of the interviews permitted us to make a qualitative analysis of the replies given by the subjects. Based on inductive processes of similarities and differences between replies, we were able to distinguish several groupings or categories. These were then arranged hierarchically from the simplest to the most complex, resulting in the construction of *empirical categories*.

One problem with categorization and, especially establishing a hierarchical order, is the different criteria that the investigator may adopt. One criterion may be that the answers given are correct or incorrect, while other criteria may be based on the intuition of the investigator or the historical development of the topic. For example, before the categorization step, the tasks appearing in the questionnaire are interpreted from a historical point of view (Benarroch, 2001), that is, taking into account the ways in which matter has been conceptualized by Aristotle, Democritus, Galileo, Descartes, Boyle, Lavoisier, Dalton, among others. In this way, both the tasks themselves and the students' replies to them, will be interpreted less subjectively than if the interviewer only bears in mind his/her knowledge of present day physics.

For the hierarchization of the categories, we gave most importance to certain criteria deduced from cognitive psychology, among them:

1. The capacity of subjects to transform the direct information provided by the tasks of the questionnaire. Subjects' replies are classified with a low value if they are dependent on the most figurative aspects of the tasks. When the replies go beyond the figurative, they are classified into a higher category.
2. Less importance is given to the terms used by the subjects than to the meaning they assign to them. For example, the *particles* they mention may be denominated *points*, *dust*, *droplets*, *molecules* or *atoms*, etc. The terms are given less importance than the meaning assigned to them.

These criteria permitted us to establish a hierarchical order of the categories, and the new categories are denominated, as mentioned above, *empirical categories*.

In this way, each individual's answer to a question or to a set of questions is given a category number, which represents the position of that answer in the system of hierarchical categories. In Table 1, we shall show the empirical categories obtained for variable L_POSTE; analogous tables were obtained for all the variables.

The establishment of empirical categories allowed us to allot each subject a set of numerical values corresponding to the variables obtained from the questionnaire. For example, the student Cat scored 6 for the variable L_POST (see Table 1). Although we do not reproduce the tables corresponding to the empirical categories of the other variables in the questionnaire, the same subject scored 3, 4, 5, 5, and 6 for the variables S_ANTE, S_POST, L_ANTE, G_ANTE, and G_POST, respectively.

One characteristic of the variables thus constructed is that they do not serve to make comparisons between subjects because the distance between categories does not follow any kind of metric. For example, in the variable L_POST of Table 1, it may well occur that two categories that we have separated "inductively" (e.g., categories 1 and 2) may or may not different levels of knowledge. For this reason, it is necessary to construct new categories, which, based on previous ones, overcome this disadvantage, so that the new groups formed really represent different levels of knowledge. These new categories are the structural categories that we shall describe below.

Table 1

Obtaining empirical categories for students' conceptions related to the variable L-POST

Differentiating Qualitative Features	Empirical Category	Students
Student makes continuous drawing for alcohol and water without explaining the decrease in height.	1	Lor, Ter, Qin, Jun, Mor, Tor, Ril, Vel
Student makes continuous drawings with bubbles or gaps for alcohol and water to explain decrease in height.	2	Gan, Tin, Kar, Nil, Her, Mek, Cas, Par
Student draws particles against continuous background for the water and alcohol. These particles are not used to explain the decrease in height.	3	Pat, Raq, Mon, Nav, Sor
Some drawings are continuous and others include particles. The students see that these is a problem (the decrease in height) and that an explanation is necessary. However, the explanations given are obviously invented.	4	Jos, Sed, Men
Student draws particles between which something ethereal (air/gas) can be seen, the same for the alcohol and water. The particles are sufficiently spaced to explain the drop in height.	5	Fan, Dan, Gos, Gim, Ber, Rom
This is the first time that the existence of a vacuum between the particles is admitted, when the students explain the drop in height.	6	Cat, Mar, Lin, Pel, Kem, San, Win, Cap, Fat, Can
Student explains the experiment by reference to a model (previously constructed or constructed during the interview).	7	Bat, Fol, Fer

Quantitative Treatment: Identification and Description of the Structural Categories

The statistical treatment of variables based on empirical categories (Table 1) makes it possible to obtain others, more precise variables, which we call “structural categories” (Table 2).

Indeed, the system of empirical categories allows us to define a type of ordinal variable, which, although not subject to any metric, can be ordered so that multivariable statistical treatment becomes possible.

Of all the statistical techniques used (*principal components analysis*, *different cluster analyses*, and *correspondence analysis*), the correspondence analysis (CA) from the BMDP statistical package (Dixon, 1990) provides the most information and best reflects the structure of the data. We can now take a closer look at the dependence of variables in terms of closeness or distance between the variables making up the empirical categories (see Figure 4).

In Figure 4, according to the distribution of the categories of variables, the meaning of axis 1 represents the overall evolution in the knowledge of the sample. The empirical categories formed by the six variables (S_ ANTE, S_ POST, L_ ANTE, L_ POST, G_ ANTE, and G_ POST) show a certain degree of continuity in their projection to axis 1, while this continuity is lost along axis 2, probably because this axis does not represent the development of the students' knowledge. Therefore, our analysis is directed more towards axis 1 than axis 2.

Figure 4 shows the empirical categories of the six variables, which are abbreviated for reasons of space. Besides the projections of the categories of the variables analyzed, five areas have been shaded. These are the areas occupied by the maximum number of empirical categories corresponding to the six variables analyzed, where, for example, L_P_7 corresponds to category 7 of the variable L_POST (see Table 1).

Table 2
Description of the structural categories of *L_POST*

Structural Categories	Description of the Structural Categories	Empirical Categories	Description of the Empirical Categories	Students
2	Students make continuous drawings without explaining the drop in height when the water and alcohol are mixed or use bubbles and gaps to explain the drop.	1	Student makes continuous drawing for alcohol and water without explaining the decrease in height.	Lor, Ter, Qin, Jun, Mor, Tor, Ril, Vel
3	Students see the water and alcohol are composed of particles against a transparent background. However, the particles do not serve to explain the drop in height.	2	Student makes continuous drawings with bubbles or gaps for alcohol and water to explain decrease in height.	Gan, Tin, Kar, Nil, Her, Mek, Cas, Par
4	Students see the water and the alcohol as being composed of particles, between which "something ethereal" can be seen. The particles are sufficiently spaced to explain the drop in height.	3	Student draws particles against continuous background for the water and alcohol. These particles are not used to explain the decrease in height.	Pat, Raq, Mon, Nav, Sor
5	Students see the alcohol and water as composed of particles with the necessary vacuum in between.	4	Some drawings are continuous and others include particles. The students see that these is a problem (the decrease in height) and that an explanation is necessary. However, the explanations given are obviously invented.	Jos, Sed, Men
6	Students see the water and alcohol as being composed of particles, between which "something ethereal" can be seen. The particles are sufficiently spaced to explain the drop in height.	5	Student draws particles between which something ethereal (air/gas) can be seen, the same for the alcohol and water. The particles are sufficiently spaced to explain the drop in height.	Fan, Dan, Gos, Gim, Ber, Rom
7	Students see the alcohol and water as composed of particles with the necessary vacuum in between.	6	This is the first time that the existence of a vacuum between the particles is admitted, when the students explain the drop in height.	Cat, Mar, Lin, Pel, Kem, San, Win, Cap, Fat, Can
		7	Student explains the experiment by reference to a model (previously constructed or constructed during the interview).	Bat, Fol, Fer

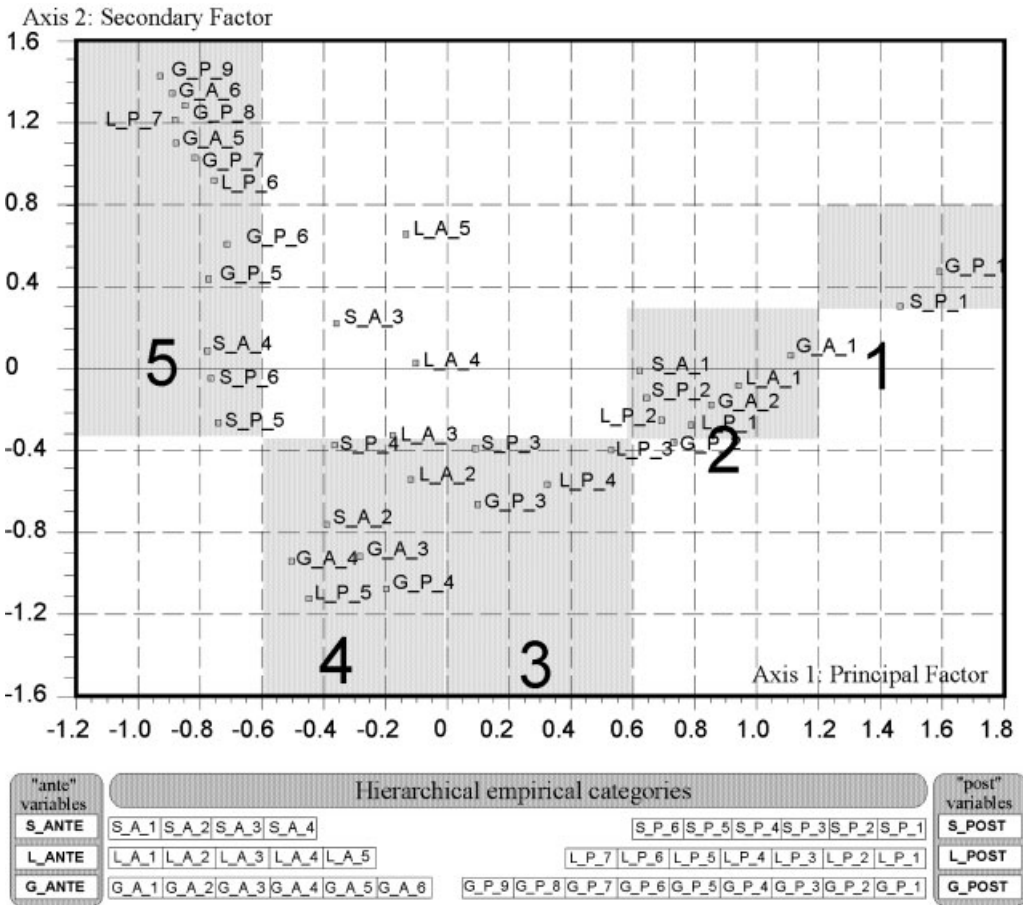


Figure 4. Categories of the six variables projected (Correspondence Analysis).

Figure 4 is the starting point for obtaining what we have termed structural categorization (Table 2), which regroups the empirical categories (Table 1) as a function of their distance from other categories. Intervals into which the different categories can be grouped were established on the two axes of the correspondence analysis using a double criterion:

1. To use the minimum number of intervals (otherwise the grouping effect would be lost).
2. To obtain the maximum number of empty areas and the minimum number of areas that contain the highest possible number of categories. In this way, the areas are closely representative of the data.

Substituting the value of the empirical category of each subject by the value taken within the graphical space of the correspondence analysis, new structural categories can be formed.

If, for example, we choose the variable G_POST, with nine categories denominated G_P_X (with values of X from 1–9), we see that the distance between the first four categories is greater than that between the five last, which are situated in the fifth rectangle. The structural categorization results from grouping the empirical categories that are found within the same rectangle.

Continuing with the same variable, then, the fifth structural category will contain the five last original categories. In this way, it was possible to effectively characterize the six variables selected (S_ANTE, S_POST, L_ANTE, L_POST, G_ANTE, and G_POST). Table 2 shows the structural categorization of the variable L_POST. Every variable would have its own table of structural categories, but are omitted for obvious reasons of space.

Validity of the Variables

Using the structural variables, it is possible to analyze the validity of the different variables defined in the questionnaire. It can be seen that the variables constructed based on the confrontation stage (S_POST, L_POST, and G_POST) are better than those corresponding to the prediction stage (S_ANTE, L_ANTE, and G_ANTE) for investigating students' knowledge.

The relation between the empirical categories of each variable and the structural variables are analyzed in Table 3.

The nomenclature and meaning of the symbols used in Table 3 are as follows:

1. The names of the variables and the tasks giving rise to them occupy the first and second columns, respectively.
2. The third column shows the number of students in each structural category (1, 2, 3, 4, and 5, structural categories which are described in Table 2 for the variable L_POST). Immediately below, the numbers correspond to the empirical categories into which the same students were classified according to the difficulty and complexity deduced empirically from the students' replies. For example, for the variable L_POST, 16 students are classified in structural category 2, while the same 16 students are classified in empirical categories 1 and 2 described in Table 2.
3. The data in the penultimate column (Discr.) express the intensity of the discrimination (+ indicating little and +++ a substantial degree of discrimination). Thus, L-ANTE only has subjects in two intervals, which means it does not discriminate well, while G_POST has all the intervals occupied and also has a good number of subjects in each interval, conferring on it a high degree of discrimination.
4. The data in the last column, Proj.: projection to axis 1 (principal factor), indicate the degree to which the original order of categories is maintained (+ indicating little and ++++ a high degree of continuity). For example, L-ANTE is not in anyway parallel to the principle axis and it has four of its five categories on the same vertical. The categories of S_POST and G_POST, on the other hand, are well distributed along the principal axis.

From the characteristics of the six variables (see Table 3), the following can be appreciated:

1. S_ANTE (variable referring to the students' replies for task 1 in prediction stage) has categories in intervals of distribution 2, 4, and 5 (see Table 3), which means its discriminative power is slight. In addition, its degree of continuity in the projection is slight because two categories (S_A_2 and S_A_3) are very close on the vertical axis in Figure 4. This, then, is a poor variable, providing information of little significance (nonrelevant).
2. S_POST (variable referring to replies to task 1 in confrontation stage), on the other hand, divides the subjects into five intervals (see Table 3) and its order of projection to the principle axis is maintained. It is, then, a good variable and much better than S_ANTE.
3. L_ANTE (prediction variable in task 2) is similar to S_ANTE in that it has little discriminative power (only occupying the second and fourth interval of distribution with a very unbalanced number of subjects: 5 and 38, respectively). Furthermore, because four of its five categories are very close (L_A_2, L_A_3, L_A_4, L_A_5), this is a poor variable.

Table 3
Characteristics of the variables on the particulate nature of matter

Task	Variable	Tasks Producing the Variables	Distribution					Discr.	Proj.
			1	2	3	4	5		
1	S_ANTE	Drawings and explanations provided for the yellow watercolor in water when the resulting color is yellow	0	16	0	23	4	+	+
	S_POST	Distribution of the empirical categories	1	13	3	2,3	4	+++	+++
		Drawings and explanations provided for the yellow watercolor in water when the resulting liquid is transparent	1	2	3	4	5,6	0	+
2	L_ANTE	Predictions about weight and height of mixture of water and alcohol in tube after shaking	0	5	0	38	0	+	+
	L_POST	Distribution of the empirical categories Representation of water and alcohol together to explain increase in concentration after shaking	0	16	8	2-5	13	+++	++
3	G_ANTE	Distribution of the empirical categories Predictions concerning constitution of air after playing with syringe and testing its compressibility	0	1,2	3,4	5	6,7	+++	+
	G_POST	Distribution of the empirical categories Explanations of constitution of air and water after observing their different degrees of compressibility	8	1,2	5	3,4	5,6	++++	++++
		Distribution of the empirical categories	1	2	3	4	5-9		

4. L_POST (confrontation variable of task 2) is present in four of the shaded rectangles (Figure 4) or four intervals of distribution (see Table 3). Its degree of continuity is poor (note that the pair of categories, L_P_2 and L_P_3 are very close on axis 1). This is neither a good nor bad variable, but it does provide more meaningful information than L_ANTE.
5. G_ANTE (prediction variable of task 3) distributes the subjects into only three intervals, although quite evenly. It shows a poor degree of continuity, and the pairs of categories (G_A_1 and G_A_2; G_A_3 and G_A_4; G_A_5 and G_A_6) are very close on the principal axis. This, too, is neither a good nor bad variable.
6. Last, G_POST (confrontation variable of task 3) not only distributes the subjects through all the intervals, it does so with good continuity throughout the graphical space defined by the first five categories. The fact that its categories with the highest distribution level are grouped in the last interval lessens the value of its discriminative information, for which reason this variable is not the best of the six.

It can be concluded, then, that the variables constructed after the confrontation between the initial replies given by the students and the empirical evidence *have a greater discriminative capacity*. This can be clearly seen from comparing S_POST with S_ANTE, L_POST with L_ANTE, and G_POST with G_ANTE. The variables _ANTE are characterized by their presence in the initial replies to the questions, while the subsequent replies (variables _POST) are enriched by the contrast between the students' predictions and the empirical evidence.

*Students' Conceptions Concerning the Particulate Nature of Matter:
Evaluation of the Degree of Coherence*

So far, we have carried out an analysis of the six variables selected (S_ANTE, S_POST, L_ANTE, L_POST, G_ANTE, and G_POST) for the three tasks used (see Table 3) and demonstrated the suitability of the methodology for differentiating types of reply. It has been seen that the strategies of confrontation, contextual variation, and asking questions of an increasing level of difficulty permit us to identify empirical (Table 1) and structural categories (Table 2), the basic characteristic of which is the regularity of the replies given by the subjects when interviewed. We can now turn our attention to the degree of coherence existing in the replies when all the replies obtained in the different tasks are presented.

From this point onwards, we shall only use the variables S-POST, L-POST, and G-POST, because these are the most useful for representing students' knowledge. The cluster analysis of cases and variables, module 7M of the BMDP (Dixon, 1990), processes the matrix of data from the three variables mentioned and the 43 subjects interviewed. This analysis makes the greatest possible number of groupings of similar cases and variables. The results can be seen in Table 4. The first column of the same table shows the different conceptual levels concerning the particulate nature of matter and the second column the most frequent explanations provided for each level.

The fact that it is possible to characterize each subject according to a given type of explanation that underlies all the tasks and which corresponds to a certain way of understanding the particular nature of matter is due to the fact that the replies of each individual are much more coherent than the investigator might at first appreciate. However, it must be said that the methodological strategy used—confrontation (prediction and confrontation stage) and contextual variation (relevant and nonrelevant) together with the increasing level of difficulty of the questions and the qualitative and quantitative analysis of the replies—was decisive in this study, because it enabled the different conceptions that the students had concerning the particulate nature of matter to be identified and ordered according to several conceptual levels.

Table 4

Description of the different conceptions of students concerning the particulate nature of matter

Level of Conceptualization	Type of Explanation	Students (age)
I. Matter is viewed as being continuous and static, except that macroscopically the opposite is observed	No explanation given. Only macroscopic description	Vel (9,6); Ril (11,4); Tor (10,8); Tin(12,9); Qin (9,7); Her (9,4); Nil (9,2)
II. Matter is viewed as continuous and stuffed with particles, or continuous with gaps	Elements perceived (bubbles, gaps, etc.) are translated into pseudo-microscopic explanations	Men (13,8); Mor (13,2); Cas (9,2); Kar (11,0); Nav (11,1); Lor (15,3); Sed (14,8); Pat (11,3); Mon (16,9)
III. Matter is formed of particles and gaps in between them. There is no need for a vacuum between particles	Microscopic explanations based on particles or ethereal gaps, to which macroscopic properties are transferred	Gan (17,0); Par (19,7); Sor (11,6); Mek (9,8); Raq (15,5); Gos (12,8); Jun (13,2); Ter (13,5); Gim (20,1)
IV. Matter is formed of particles and a vacuum is necessary between them	Microscopic explanations based on the arrangement of particles (separated to a greater or lesser degree)	Ber (17,8); Jos (14,8); Kem (9,3); Rom (19,8); Fan (15,8); Can (17,1); Mar (15,0); Dan (17,3); Lin (13,1)
V. Movement is necessary and there is a causal coordination with the vacuum	Academic or quasi-academic microscopic explanations	Win (19,3); Pel (11,8); Cap(19,6); Fat (20,3); Cat (15,7); San (22,7); Fol (17,6); Bat (17,3); Fer (17,3)

To identify regularities, the values of the structural categories for the three variables mentioned (S-POST, L-POST, and G-POST) were compared with the levels established by the cluster analysis (see Table 4). Table 4 shows that, for some subjects, the value of the structural variable differs from the level assigned by the cluster analysis. For example, the student Can, placed in level IV (Table 4) scores 5 in L_POST (Table 2). This can be seen in Table 5.

Table 5 shows the matrix of cases and variables grouped according to level of conceptualization. The first column of each level shows the subjects and the next three columns the values obtained for the three variables (S-POST, L-POST, and G-POST). The letter N (normality) indicates that the difference between the value of the variable of the structural category (Table 2) and the level assigned by the cluster analysis (Table 4) is between -1 and $+1$. The numbers depicted in Table 5 are outside this range. For example, student Can, who has 5 in the structural category of L_POST (Table 2) and is in level conceptual IV (Table 4), has an N (normality) in the corresponding box. The same subject, on the other hand, has 2 in S_POST, because he is outside the established limit because he shows a conceptual level of IV (Table 4), while the category value of S_POST is 2.

It can be seen from Table 5 that only seven subjects show irregularities with respect to the established margin and that any irregularity only affects one variable of the three considered. These irregularities mean that the conceptual level assigned (Table 4) to the interviewer for one of the three variables (S-POST, L-POST, and G-POST) differs substantially. Except in these seven cases, the level of the replies to the tasks is similar.

The high degree of regularity between the structural categories of the replies underlines the high degree of coherence in these replies. It should be borne in mind that the replies were given for widely varying tasks, that is, they refer to different states of aggregation (see questionnaire).

Table 5
Coherence in the variables S_POST, L_POST and G_POST (S_P, L_P, G_P)

Stu.	Level V			Level IV			Level III			Level II			Level I		
	S_P	L_P	G_P	S_P	L_P	G_P	S_P	L_P	G_P	S_P	L_P	G_P	S_P	L_P	G_P
Fer	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Bat	N	N	N	N	N	N	N	N	N	Mon	N	N	N	N	N
Fol	N	N	N	N	N	N	N	N	N	Pat	N	N	N	N	N
San	N	N	N	2	N	N	N	N	N	Sed	N	N	N	N	N
Cat	N	N	N	N	N	N	N	N	N	Lor	N	N	N	N	N
Fat	N	N	N	N	N	N	N	N	N	Nav	4	N	N	N	N
Cap	N	N	N	N	N	N	N	N	N	Kar	4	N	N	N	N
Pel	N	N	N	N	N	N	N	N	N	Cas	4	N	N	N	N
Win	3	N	N	N	N	N	N	N	N	Mor	N	N	N	N	N
										Men	N	N	N	N	4

It also revealed that for each kind of explanation or way of conceptualizing matter, a group of subjects existed whose replies were coherent for all the tasks presented.

Results and Suggestions

We have attempted to throw some light on the degree of coherence of students' conception of the *particulate nature of matter*, which some authors have regarded as high and others as low. We observed a high degree of coherence in the replies given by the interviewees when confronted with tasks referring to different aggregation states, with relevant and nonrelevant variations, and questions of varying degrees of difficulty. The high degree of coherence in the replies was also made explicit from the level of conceptualization of each subject.

Knowledge of the level of students' conceptualization concerning a given topic should help educators decide the best approach, knowing which conceptual elements to include and which to exclude. The same knowledge also provides a guideline as to how explanations should be ordered to help students in the development of their mental capacity.

The method used enabled us to identify the different conceptual levels of students concerning the particulate nature of matter. We outline below the most important steps followed.

First of all, the fact that the information obtained after the confrontation stage showed a greater degree of regularity than the information obtained during the prediction stage emphasizes the need for contrasting any information obtained. For this reason, an interview protocol should follow certain strategies, as detailed below:

1. Order the questions from easiest to most difficult (vary the degree of complexity). The academic content can provide the most complex questions, while formulation of the easiest questions is the most difficult part for the investigator because they must be the most meaningful for the student.
2. Identify different contexts for which different tasks can be designed, but with the same underlying topic (context variation).
3. Formulate a variety of questions for the same context, modifying the relevant factors (relevant variations).
4. Formulate a variety of questions for the same context, modifying the nonrelevant factors (nonrelevant variations).
5. Take a sample student population covering a wide and representative age-range (age variations). This is done to ensure that the most representative levels of knowledge on a given topic are covered.
6. Plan the tasks using confrontation strategies, whereby the interviewee can predict the possible outcome of a task or physical event (prediction stage) and then, after carrying out the experimental demonstration, confront the empirical results with the predictions. These strategies, together with an individualized interview protocol, set in motion an interactive dynamic that favors the elaboration of reflexive replies, were the subject is fully involved cognitively.

Second, the data treatment, first qualitative and then quantitative, permitted us to obtain the conceptual levels of the students concerning the particulate nature of matter and to observe the degree of coherence in the students' replies. For this, the following steps are followed:

1. The replies of the subjects (interviewed using the questionnaire constructed according to the above described strategies) are characterized inductively according to their similarities and differences and deductively in accordance with cognitive psychology. This permits us to break down the above process and provides a hierarchical order which we refer to as "empirical categories."

2. The empirical categories enable us to construct ordinal variables susceptible to multi-variable treatment. The results of the statistical modules, such as principal components analysis, different cluster analyses and, especially, correspondence analysis provide a new type of category called “structural.”
3. Analysis of the structural categories permits them to be evaluated. In this way, it was seen that variable constructed after the confrontation between the initial replies given by the students and the empirical evidence have a greater discriminative capacity and, therefore, are more valid for describing students’ knowledge.
4. Cluster analysis of cases and variables (interviewees and structural categories) groups the subjects into levels of knowledge independently of the tasks.

The conceptual levels established concerning the particulate nature of matter, the observed regularity of the students’ replies for each level and the structural categories for the most significant variables of the tasks lead us to affirm that the interviewees’ knowledge of the topic was greater than we at first supposed.

To summarize:

1. A greater degree of coherence was evident in the replies when the interviewees were replying to confrontation questions than when predicting events. In general, pen-and-paper questionnaires collect information from students based on *prediction replies* and offer a more heterogeneous image of students’ knowledge than when personal interviews are used. In such interviews, which involve confrontation and context variation strategies, students are encouraged to activate their knowledge schemes to a much greater extent.
2. Confrontation replies imply that a certain degree of learning has taken place since the *prediction stage* because students will have had the opportunity to activate their knowledge schemes and also to acquire new data or reorder those held, thus generating new conceptualizations.

Appendix: Interview with TER (13,5 years old), Task 2

TER: pseudonym of student interviewed; P: questions and notes; R: interviewee’s replies

- P *How do you explain that the weight is the same but the height is less? (See Figure 2)*
- R I’ve got here a test tube. I’ve put some water in and now I’m going to pour some alcohol on top. Can you see the division between them? Look there’s like a break.
- P *What do you think will happen if we shake it well?*
- R It will mix and the alcohol will go next to the water...it doesn’t dissolve, or let’s say the water doesn’t dissolve with the alcohol. I suppose there are two different things, which can’t dissolve. So, when you mix them, they don’t go into each other, form different parts.
- P *Won’t they mix, really?*
- R That’s right. There’ll be a part that’s water and another of water... an area surrounded by another...like oil...a circle surrounded by water. Something like that.
- P *Will the height change?*
- R No.

P *What about the weight?*

R No...unless it escapes.

P *The interviewer marks the top of the column with a marker pen....127....Now, let's weigh it...You look at the weight.*

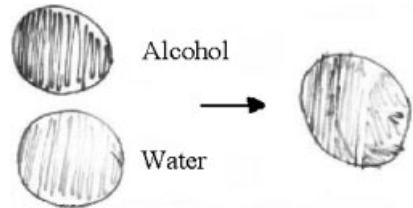
R 14.9 grams. Shall we shake it now?

P *Yes, OK. What's happened?*

R Is it mixed? Looks like it. It's not dissolved. Some parts of the water have mixed with other parts of the water.

P *Can you draw it?*

R Can I use two different colors? Just to show the difference. The water's yellow. Just to differentiate it .and the alcohol's green...so, the mixture will.....without mixing, in other words, some parts next to the others, but not mixed....like so.



P *Does that explain the drop in height?*

R I don't know...I've never thought of it...I don't think so....(The student doesn't know how to continue).

P *How much do you think it weighs now?*

R The same, I think.....nothing's coming out.....and nothing's gone in.

P *Let's do the same experiment but in color. Do you think the same thing will happen?*

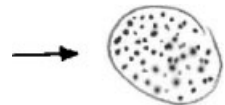
R Yes. It'll do the same...the height will fall. (The student is worried about the height).

P *[Shake tube]. What's left?*

R The same as before. It looks as if it's mixed, but if we blew it up in size, we'd see blue and red areas, it wouldn't open up.....But I don't know if water and alcohol can dissolve, maybe yes, maybe no. I don't know. Judging from the color, it looks as if they can.

P *If we suppose they can dissolve, how would you draw it?*

R The bits that are separated join up...you'd have to color them like before, with dots, making like small molecules, perhaps lilac colored.

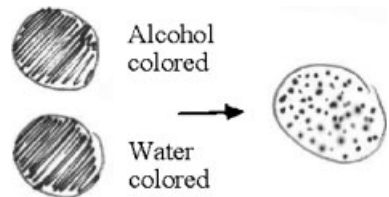


P *Is that how you want to draw it?*

R With dots?

P *I don't know.*

R Maybe. The alcohol breaks into very small parts, tiny, and the water breaks up, too...they mix with each other...making a compact mass, I think..., maybe like this



- P *Before mixing, the alcohol is also small blue dots?*
- R No. I don't think so. . . . it opens up. . . . divides into tiny pieces, dissolve with each other, making a uniform mauve mass
-

References

- Benson, D.L., Wittrock, M.C., & Baur, M. (1993). Students' preconceptions of the nature of gases. *Journal of Research in Science Teaching*, 30, 587–597.
- Benarroch, A. (1998). Las explicaciones de los estudiantes sobre las manifestaciones corpusculares de la materia. Tesis inédita. Granada: Facultad de Educación.
- Benarroch, A. (2001). El desarrollo cognoscitivo de los estudiantes en el rea de la naturaleza corpuscular de la materia. *Enseñanza de las Ciencias*, 18, 235–246.
- Brook, A., Briggs, H., & Driver, R. (1984). Aspects of secondary students understanding of the particulate nature of matter. *Children's Learning in Science Project*, Center for Studies in Science and Mathematics Education, University of Leeds.
- Brown, D.E. (1989). Students' concept of force: The importance of understanding third Newton's law. *Physics Education*, 24, 353–358.
- Brumby, M. (1979). Problems in learning the concept of natural selection. *Journal of Biological Education*, 13, 119–122.
- Carmichael, P., Driver, R., Holding, B., Phillips, I., Twigger, D., & Watts, M. (1990). Research on students' conceptions in science: A bibliography. University of Leeds: Children's Learning in Science.
- Case, R. (1983). El desarrollo intelectual: Una reinterpretación sistemática. En Carretero & Madruga. *Lecturas de psicología del pensamiento. Razonamiento, solución de problemas y desarrollo cognitivo* (pp. 339–362). Madrid: Alianza Editorial.
- diSessa, A.A., & Sherin, B.L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20, 1155–1191.
- diSessa, A.A., Gillespie, N.M., & Esterly, J.B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, 28, 843–900.
- Dixon, W.J. (1990). *BMDP statistical software manual* (vols. I and II). Berkeley: University of California Press.
- Driver, R. (1986). Psicología cognoscitiva y esquemas conceptuales de los alumnos. *Enseñanza de las Ciencias*, 4, 3–15.
- Driver, R. (1988). Un enfoque constructivista para el desarrollo del curriculum en Ciencias. *Enseñanza de las Ciencias*, 2, 109–120.
- Driver, R., Guesne, E., & Tiberghien, A. (1985). *Children's ideas in science*. London: Open University Press.
- Eilam, B. (2004). Drops of water and of soap solution: Students' constraining mental models of the nature of matter. *Journal of Research in Science Teaching*. Published online in Wiley InterScience (www.interscience.wiley.com).
- Gamble, R. (1989). Force. *Physics Education*, 24, 79–82.
- Gómez, M.A., Pozo, J.I., & Sanz, A. (1995). Students' ideas on conservation of matter: Effects of expertise and context variables. *Science Education*, 79, 77–93.
- Griffiths, A., & Preston, K. (1992). Grade 12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching*, 29, 611–628.

Haidar, A.F., & Abraham, M.R. (1991). A comparison of applied and theoretical knowledge of concepts based on the particulate nature of matter. *Journal of Research in Science Teaching*, 28, 919–938.

Halldén, O. (1999). Conceptual change and contextualization. In W. Schnotz, S. Vosniadou & M. Carretero (Eds.), *New perspectives on conceptual change* (pp. 53–65). London: Elsevier.

Hewson, P., & Thorley, R. (1989). The conditions of conceptual change in the classroom. *International Journal of Science Education*, 11, 541–553.

Hierrezuelo, J., & Montero, A. (1991). La ciencia de los alumnos, “Su utilización en la didáctica de la Física y Química.” Vélez Málaga: Elzevir.

Ioannides, C., & Vosniadou, C. (2002). The changing meaning of force. *Cognitive Science Quarterly*, 2, 5–61.

Kuiper, J., & Mondlane, E. (1994). Student ideas of science concepts: Alternatives framework? *International Journal of Science Education*, 16, 279–292.

Lee, O., Eichinger, D.C., Anderson, C.W., Berkheimer, G.D., & Blakeslee, T.D. (1993). Changing middle school students’ conceptions of matter and molecules. *Journal of Research in Science Teaching*, 30, 247–270.

Marín, N. (1994a). Elementos cognoscitivos dependientes del contenido. *Revista interuniversitaria de formación del profesorado*, 20, 195–208.

Marín, N. (1994b). Evolución de los esquemas explicativo en situaciones de equilibrio mecánico. Tesis inédita, Granada: Facultad de Educación.

Marín, N., Benarroch, A., & Jiménez Gómez, E. (2000). What is the relationship between social constructivism and piagetian constructivism? An analysis of the characteristics of the ideas within both theories. *International Journal of Science Education*, 22, 225–238.

Marín, N., Jiménez Gómez, E., & Benarroch, A. (2004). How to identify replies that accurately reflect students’ knowledge? A methodological proposal. *International Journal of Science Education*, 26, 425–445.

McClelland, J.A.G. (1984). Alternative frameworks: Interpretación of evidence. *European Journal of Science Education*, 6, 1–6.

Meheut, M., Larcher, C., & Chomat, A. (1988). Modelos de partículas en la iniciación a las ciencias físicas. *Enseñanza de las Ciencias*, 6, 231–239.

Nakhlah, M.B., & Samarapungavan, A. (1999). Elementary school children’s beliefs about matter. *Journal of Research in Science Teaching*, 36, 777–805.

Novick, S., & Nussbaum, J. (1978). Junior high school pupils understanding of the particulate nature of matter: An interview study. *Science Education*, 63, 273–281.

Novick, S., & Nussbaum, J. (1981). Pupils’ understanding of the particulate nature of matter: A cross-age study. *Science Education*, 65, 187–196.

Nussbaum, J., & Sharoni-Dagan, N. (1983). Changes in second grade children’s preconceptions about the Earth as a cosmic body resulting from a short series of audio-tutorial lessons. *Science Education*, 67, 99–114.

Oliva, J.M. (1996). Estudios sobre consistencia de las ideas de los alumnos en Ciencias. *Enseñanza de las Ciencias*, 14, 87–92.

Osborne, R.J., & Freyberg, P. (1985). *Learning in science. The implications of children’s science*. London: Heinemann.

Pascual-Leone, J. (1979). La teoría de los operadores constructivos. En Juan Delval. *Lecturas de psicología del niño* (pp. 208–228). Madrid: Alianza Universitaria.

Piaget, J. (1946). *Les notions de mouvement et de vitesse chez l’enfant*. Paris: P.U.F.

Piaget, J. (1975). *La composición de la fuerzas y el problema de los vectores*. Madrid: Morata.

Piaget, J. (1978). "La equilibración de las estructuras cognitivas, "Problema central del desarrollo." Madrid: Siglo XXI.

Pfundt, H., & Duit, R. (1994). *Bibliography: Students' alternative frameworks and science education* (4th ed.). Kiel: IPN.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211–227.

Pozo, J.I. (2003). *Adquisición de conocimiento*. Madrid: Morata.

Pozo, J.I., Gómez Crespo, M.A., Limón, M., & Serrano Sanz, A. (1991). *Procesos cognitivos en la comprensión de la ciencia: Las ideas de los adolescentes sobre la Química*. Madrid: CIDE (MEC).

Pozo, J.I., & Gómez Crespo, M.A. (1998). *Aprender y enseñar ciencia*. Madrid: Morata.

Samarapungavan, A., & Wiers, R. (1977). Children's thoughts on the origin of species: A study of explanatory coherence. *Cognitive Science*, 21, 147–177.

Snir, J., Smith, C., & Raz, G. (2003). Linking phenomena with competing underlying models: A software tool for introducing students to the particulate model of matter. *Science Education*, 87, 794–830.

Stavy, R. (1988). Children's understanding of gas. *International Journal of Science Education*, 10, 553–560.

Stavy, R. (1990). Pupils' problems in understanding conservation of matter. *International Journal of Science Education*, 12, 501–512.

Stavy, R., & Tirosh, D. (1993). When analogy is perceived as such. *Journal of Research in Science Teaching*, 30, 1229–1239.

Stavy, R., & Tirosh, D. (1993). When analogy is perceived as such. *Journal of Research in Science Teaching*, 30, 1.229–1.239.

Taber, K.S. (2000). Case studies and generalizability: Grounded theory and research in science education. *International Journal of Science Education*, 22, 469–487.

Terry, C., & Jones, G. (1986). Alternative frameworks: Newton's third law and conceptual change. *European Journal of Science Education*, 8, 291–298.

Vosniadou, S., & Brewer, W.F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535–585.

Watts, D.M., & Zylbersztajn, A. (1981). A survey of some children's ideas about force. *Physics Education*, 16, 360–365.

Yair, Y. & Yair, Y. (2004). Everything comes to an end: An intuitive rule in physics and mathematics. *Science Education*, 88, 594–609.