International Journal of Science Education
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tsed20

Characteristics of the methodology used to describe students' conceptions

N. Marín Martínez , I. Solano & E. Jiménez Gómez

Published online: 20 Jul 2010.

INT. J. SCI. EDUC., 2001, VOL. 23, NO. 7, 663-690

RESEARCH REPORT

Characteristics of the methodology used to describe students' conceptions

N. Marín Martínez, Dpto. de Didáctica de la Matemática y de las Ciencias Experimentales, I. Solano, Universidad de Almería, Spain and E. Jiménez Gómez, Dpto. de Didáctica de las Ciencias Experimentales, Facultad de Educación, Universidad de Murcia, Spain.

In a paper published elsewhere (Jiménez Gómez *et al.* 1997), we compared research carried out on students' conceptions of force in the decade 1975-85 with later research in the period 1985-95. Only a slight progression was found in the results offered by different authors.
In this contribution we analyse the research objectives of selected investigations into students' conceptions of force, their methodological characteristics and their most widely used theoretical foundations, in an attempt to explain the lack of progress made in this research line. Some suggestions are made which may help the science teaching profession to understand students' conceptions better.

Introduction

More than two decades have passed since the need was seen for more attention to be paid to the conceptions held by students concerning science content (see Driver and Easley 1978). During these 20 years, a large number of studies have been published and a new language has come into being among educators and researchers to describe the problems associated with the teaching and learning of science. The studies, most of them concerned with science teaching, have been accepted by most educators and researchers to make up a 'theory' in the postmodernist sense (Solomon 1994), giving rise to the so-called 'alternative conceptions movement' (ACM) (Gilbert and Swift 1985) or 'social constructivism' (Gilbert 1995; Marín *et al.* 1999).

In previous papers (e.g. Marín and Jiménez Gómez 1992; Jiménez Gómez *et al.* 1994; Marín and Benarroch 1994; Jiménez Gómez *et al.* 1997; Marín *et al.* 1997) we have attempted to show that achievements in research into students' conceptions of the subject matter need to be looked at afresh in a theoretical context and using a different research methodology. In this paper, we analyse the most noteworthy methodological characteristics of research into students' conceptions. Based on the findings of such an analysis we make suggestions designed to avoid the methodological deficiencies found in this kind of study.

International Journal of Science Education ISSN 0950-0693 print/ISSN 1464-5289 online © 2001 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/09500690010006509

Selection and classification of a sample of studies on conceptions

Mechanics was chosen since it is one of the subjects where students' conceptions have been most studied (Moreira 1994). However, the enormous number of articles found led us to restrict our investigations to the principles of dynamics, more specifically the concept of force and laws related with it.

Another problem was to choose representative samples of research into students' conceptions concerning force (Jiménez Gómez *et al.* 1997). Most research along these lines is published in the proceedings of congresses and meetings, so they are frequently difficult to obtain. For this reason, we opted only for those articles published in journals specializing in science teaching. However, this still meant a very large number of articles and so we decided to concentrate on articles published in the following journals, which, besides being easily available (in Spain) cover the object of our research: *Enseñanza de las Ciencias, European Journal of Science Education* (now: *International Journal of Science Education*), *Physics Education* and *Science Education*.

We excluded articles dedicated almost exclusively to theoretical discussion, even though they may occasionally make reference to students 'conceptions' on force since they do not offer methodological proposals or results. In total, 29 articles were chosen and classified into three categories.

Type 1: works that categorize students' replies according to their similarities and differences usually expressed as percentages. The results and conclusions, which are basically inductive, are linked with the peculiarities of the physical facts presented in the questionnaire. We have called this group of studies *descriptive-physical fact dependent*.

Type 2: works that, besides using the above described procedures, attempt to establish relationships between the different groups of answers corresponding to the questions on different physical facts. The general applicability of the conclusions reached in this type of study is greater than that of the first type, to such an extent that the conclusions may sometimes reach a degree of generality that is not implied in the facts contained in the test. Some authors make predictions as to the possible answers of students to situations not presented in the questions. As in the first group of studies, the data are interpreted from the point of view of the logic of the content. This category of study is referred to as *descriptive-physical fact independent*.

Type 3: works that establish relationships between the categories system used to classify the students' answers and aspects or variables shown to be relevant in the cognitive structure of the student, such as cognitive style, short- and long-term memory, mental operations, etc. We refer to this type of study as *descriptive-relational* since they only use descriptive techniques to relate variables, but do not establish relationships between the students' answers and those basic cognitive variables contained in a proper theoretical cognitive context.

Table 1 shows the authors classified according to the above three categories.

Note that small number of studies classified as descriptive-relational makes this group unrepresentative. The methodological characteristics used in the descriptive-physical facts dependent group are weaker than those contained in the independent group. For this reason we consider it sufficient for our analysis to concentrate on the 18 articles as *descriptive-physical fact independent*.

Table 1. Classification of the articles selected.

Descriptive dependent	*Descriptive independent*	*Descriptive relational*
Helm 1980; Gunstone and White 1981; Ivowi 1984; Terry and Jones 1986; Brown 1989; Clement *et al.* 1989; Kruger *et al.* 1992; Galili 1993.	Viennot 1979; Watts and Zylbersztajn 1981; Watts 1982; Watts 1983; Maloney 1984; Ruggiero *et al.* 1985; Terry *et al.* 1985; Noce *et al.* 1988; Boeha 1990; Villani and Pacca 1990; Finegold and Gorsky 1991; Thijs 1992; Galili and Bar 1992; Reynoso *et al.* 1993; Bar *et al.* 1994; Kuiper and Mondlane 1994; Twigger *et al.* 1994; Montanero *et al.* 1995.	Bar 1989; Acevedo *et al.* 1989.

Strategies and categories used in abstracting data from the selected sample

It is of fundamental importance to respect the original data when evaluating the methodological procedures used by other authors, although there is always the temptation to offer one's own interpretation when recording and categorizing bibliographical data, a temptation that must be resisted in the interest of improving methodological studies. Data were recorded using a double procedure: first, inductive, recording the most relevant data offered by each author and, by using criteria of similarity and difference, analysing the data common to several authors, the data particular to one work, the relationships between studies, etc; second, deductive, based on methodological considerations described in more detail in previous works (e.g. Marín and Jiménez Gómez 1992; Jiménez Gómez *et al.* 1994; Marín 1995), but briefly:

- The need to base the search for, and definition of, the students' knowledge on a theoretical platform which avoids, to the greatest extent possible, bias and distortion in the interpretation of the data.
- The need to operate with a model of the subjects' cognitive activity, which avoids the inductive and descriptive character of investigation, in which the search is not directed by any type of hypothesis, expectations or foresight.
- The need to use methodological strategies that discriminate between the responses that reflect some sort of scheme in the student's knowledge and those arising from compromise, inventions, or simple chance. Such methodology should:

  (i) present a variety of situations in which the topic is involved;
  (ii) vary the different factors intervening in the several physical situations presented;
  (iii) ask for likely outcomes (prediction), based on a given initial configuration of a physical fact, in order to evaluate the subject's transformational capacity; and

(iv) weigh up the operational capacity of the subjects using situations, in which different figurative factors are varied while the variable in question is kept constant.

Thus, far from rushing forward to the bibliographical data to see what could be found, we established a line of study in order to collect data. However, this line of study was not rigid and interacted continually with fresh bibliographical data. This permitted us to modify the line of study to include the data (an inductive process), at the same time as deductive processes showed up certain absences or shortfalls which might otherwise have been overlooked.

The application of the above strategies to bibliographical data made it possible to draw up tables 2 and 3 whose content we shall briefly describe:

- Table 2, which is of a descriptive nature, summarizes the bibliographical data of the selected studies (descriptive-physical fact independent).
- Table 3, which is obtained from the data of table 2, is a synthesis which reveals the most important methodological data.

Both tables (2 and 3) contain the same columns; which describe:

- what the authors are looking for;
- the procedure used and theoretical framework;
- the techniques used in obtaining information from the student;
- data treatment. In table 3, the single column of table 2 has been subdivided in order to show: (i) the criteria for grouping and relating data; and (ii) 'test-tuning' and studies on the coherence and reliability of the data.

We describe each of these columns in more detail below.

What are the authors looking for?

This column shows the name of the author(s), the year of publication and the subject matter of the research. The data recorded answer the questions: what are the authors looking for? What do they intend to show? We pay special attention to the terminology used. We also record the authors' views on the possible persistence, coherence, consistency, etc. of what is being looked for.

In the first column of table 3 we categorize the search for regularities such as : A.1 laws, A.2 models, A.3 strategies, A.4 beliefs, A.5 theories, A.6 common ideas, A.7 conceptions, A.8 schemes or alternative frameworks, A.9 ideas which were accepted at certain times during the history of science.

Procedures and references

We attempt to give an overview of the methodology used in the study concerned and analyse whether it follows a clearly inductive procedure, hypothetical-deductive or whether both procedures are used, depending on the phase of the research.

First, we describe the suppositions on which the research is founded, mentioning: (i) the authors most cited or considered most relevant to the work; (ii) the preoccupation of the authors or aspects on which they put most emphasis; and (iii) the suppositions or foundations presented by the authors as a starting point for their research.

We pay special attention to the coherence of such initial suppositions and their use in subsequent steps of the research in the formulation of hypotheses, construction of questionnaires, data treatment, interpretation and categorization. In this way, it is possible to detect any inconsistencies between the various parts of the work in question. In particular, if the authors set out any hypothesis, we attempt to see whether these are constantly adhered to throughout the paper.

For the corresponding column of table 3, three aspects are emphasized:

A. The general *sequence* that the investigation followed: 1. inductive; 2. hypothetical-deductive; 3. a combination of the two.
B. The explicitly *expressed framework*, on which the research is based: 1. the academic content; 2. rules and mental models; 3. historical conceptions.
C. *Other works* are used to: 1. confirm data; 2. use techniques; 3. evaluate data; 4. continue the research.

Techniques used to gather information from students

This section describes the sample and the structure of the test. The scheme for homogeneous data taking is as follows:

- The make up of the sample: number of subjects, level of study and ages, type of school or university studies.
- Type of technique used for obtaining information (data acquisition):
  - (i) The questions, are they problem situations which must be solved? Are they questions on the definition of concepts? Do they refer to everyday situations? Can the student physically handle materials or are drawings presented? Are predictions and explanations asked for?
  - (ii) As regards possible variations in factors: do the factors (whether relevant or irrelevant) intervening in the task change gradually? Are the questions repeated by changing the context or situation?
  - (iii) Concerning the sought for content: is it explicitly shown up in the question or is it implicitly involved?
- Are parallel tests given which have previously been shown suitable for evaluating an important aspect of what the student is supposed to know? If so, which tests?
- It is important to know whether several tests are given, in which the initial questions are gradually modified according to the students' replies? If so, the reason for using this technique should be considered.
- Any other peculiarities of the study.

To complete column 3 of table 3 ('Techniques used for gathering information from the students'), it has been divided into five sections, each with its corresponding categories:

A. Paper and pencil: 1. open; 2. closed with no explanation; 3. closed with explanation.
B. Interview: 1. individual; 2. in pairs.
C. Type of question: 1. direct question; 2. explanations; 3. predictions.
D. Situations: 1. in drawings; 2. real.

Table 2. Methodological analysis of the descriptive-physical fact independent.

1. What are the authors looking for?	*2. Procedures and references*	*3. Technique used in obtaining information from the student*	*4. Data treatment*
Viennot (1979) explores and analyses the *spontaneous reasoning of students* concerning elementary dynamics, considering it a kind of everyday thinking that can be formalized in terms of its own laws. An *intuitive explanatory scheme, general tendencies in spontaneous reasoning and intuitive laws and notions* are described.	All students share a common explanatory scheme of 'intuitive physics' which, although not taught at school, represents a common and self-consistent stock of concepts that resist attempts to change or modify them, however wrong they may be. Thus, inductively, there exist an *intuitive explanatory scheme, general tendencies in spontaneous reasoning, intuitive laws and notions* and it is possible to set up and roughly classify the types of questions which give rise to each notion in spontaneous reasoning.	The research involved several hundred students (mainly French, but also British and Belgian) from the last year at secondary school to the third year at University using *pencil-and-paper tests*, each taking 20 to 30 minutes. *The questions asked were open and direct*, (e.g. 'Is there a force?') about some specific physical situation and provided an opportunity for the students to make *predictions* and give *explanations. Occasionally, some variables involved in the situation were modified.*	Different *intuitive laws, notions and/or general tendencies in spontaneous reasoning* were hypothesized to take incorrect answers or mistakes into account. No percentages are given, perhaps because the author is looking for tendencies in reasoning which, it is supposed, are common to groups of subjects.
Watts and Zylbersztajin (1981) assessed the popularity of some *particular alternative frameworks* (Driver and Easley 1978 put forward the term 'alternative frameworks' to represent students' personal conceptualizations) found in the literature. The authors were keen to ask two questions: how prevalent are non-Newtonian views and to what extent are teachers aware of the prevalence?	The authors act in the belief that children construct coherent ideas of the world based on their own experiences. These ideas or alternative frameworks are seen as individual idiosyncratic hypotheses and speculations concerning events and conversations. Words, meanings and experiences are construed and developed into personal constructions of science. The authors *inductively* look for the *alternative frameworks* of students described in other studies (Warren 1971, Driver 1983, Driver 1979, Viennot 1979, Watts 1980). The reference for this search is the academic content itself, starting with the supposition that a 'gulf of understanding' exists between teachers and pupils, the extent of which the authors wish to measure.	The participants were 125 end-of-third-year pupils and their five teachers from four comprehensive secondary schools in Reading and London *The paper-and-pencil questionnaire consisted of multiple choice questions*. At the end of each question, there was a space and the pupils were asked to include a *reason for their answer*. In each question, one possible response was always the 'scientifically correct' answer, whilst at least one of the others was inspired by an alternative framework already described by research workers. Questions included stick-people drawings and *forces were indicated through arrows*.	Answers were analysed by both quantitative and qualitative techniques. *Generally, each alternative framework was derived from a particular physical situation (except one)* and no relation was established between different alternative frameworks. The results were presented numerically in the paper, generally, for the most frequent choices and content analysis of the pupils' written explanation alongside the numerical predictions and comments of the teachers were presented.

Watts (1982) attempts *to describe children's alternative frameworks of gravity* and provide physics teachers with quick and useful insights into children's thinking.	Rather than dismiss students' ideas as 'wrong' or 'misconceived', Watts prefers the term 'alternative framework' used by Driver and Easley (1978) to describe children's imaginative efforts to exlain the world around them. A pupil's own view can often be very different from the traditional physics view and can persist even through A-level and under-graduate work. *Inductively*, the authors look for *children's alternative frameworks of gravity:* the youngsters use the frameworks both implicitly and explicitly during the course of the interviews, such frameworks being drawn from the comments that children make. This report follows up other studies: e.g. Nussbaum and Novak (1976), Gunstone and White (1980) etc.	The participants were 20 pupils from secondary schools (science/physics group) in and around the Greater London area, who ranged in age from 12 to 17 years and who formed part of a larger sample. The students took part in *forty-minute interviews* about force. During these, the youngsters were asked to *interpret some simple line drawings* depicting various situations where the concept of force may or may not be applicable. During the course of the discussions, all 20 youngsters felt the need of use the term gravity and were often asked *to expand on their meaning* of that word.	The eight *frameworks* that were described belonged to no one pupil, but were composite pictures drawn from many discussions. No attempt was made to quantify these results, nor to make generalizations about children's understanding from what was quite a small sample. There were many subtle variations within each one student, leading to the conclusion that for any group of pupils there might be as many qualitatively different frameworks as there are children, although sufficient similarities exists to warrant their grouping into frameworks. No hierarchies between alternative frameworks are established. *The theoretical framework used to evaluate the data is the academic content.*
Watts (1983) attempts to provide science educators with a repertoire of *common frameworks and to describe some ways in which they differ from orthodox physics*. He describes eight different frameworks that pupils use when describing common phenomena in terms of their own concept of force. The author sees the need to develop *models of student understanding* that are powerful enough *to distinguish important individual differences*, yet *not so specific* that the end product is the same number of models as there are pupils.	*Inductively*, the author describes *children's alternative frameworks of force*. In analysing individual interview transcripts, an attempt is made to construct frameworks that can account for statements by a pupil in such a way that all the statements are mutually compatible. Throughout the study, children's alternative frameworks are respectfully given a status that reflects their widespread use, internal coherence and their tenacity in the face of classroom teaching. The assumption that all of a person's statements are logically consistent to a listener (or a reader), is difficult to maintain. However, it is one that has to be made as a working hypothesis, otherwise it would be easy to discount sections of a student's discourse that seem inconsistent with other understandable sections.	The participants in this study were 12 schoolchildren ranging in age from 11 to 18 years, from schools (junior science classes and advanced level physics groups) in the Greater London urban area. Pupils were interviewed individually using a method called the *Interview-about-instances approach* (Gilbert and Osborne 1980, Gilbert, Watts and Osborne 1982). The interviews centred on a collection of simple line drawings, which the pupils were required to interpret as examples or non-examples of their concept of force. They were asked for their *reasons for each choice*. Direct questions about force were asked and the interviews were audio-taped and then fully transcribed with a detailed notation.	The *frameworks* described do not come from one pupil but are pieced together from the implicit and explicit conceptions used by the children during the course of the interview. Each framework was used by at least two pupils and used at least twice during the interviews. The frameworks form a composite picture based on ideas shared by a number of pupils, an approach termed the 'mosaic' method. *The theoretical framework for evaluating data is the academic content.*

Continued

Table 2—*continued*

1. What are the authors looking for?	*2. Procedures and references*	*3. Technique used in obtaining information from the student*	*4. Data treatment*
Maloney (1984) investigates *students' conceptualizations of situations involving two objects exerting forces on each other (Newton's third law)*. The author thinks it is important to *identify students' existing strategies and conceptions* and uses Siegler's '*rule assessment*' *technique* to do that.	Siegler's rule assessment technique enables the author to make certain predictions *(hypothetical-deductive)* and requires a task analysis to: (i) identify the problem types and (ii) determine the strategies (correct or incorrect) that can be applied to the tasks (problems). The incorrect strategies revealed by the task analysis were used along with the correct strategy to generate 'ideal' response sequences. This was done by taking the particular strategy (e.g. making all decisions on the basis of the greater mass) and applying it to each item to determine the ideal sequence of 24 answers that strategies would produce. This was done for all the strategies and the subjects' responses were then compared with all of the ideal sequences until a match was found.	The participants were 112 high school students. The question paper involved the simplest possible situation, i.e. two blocks on a horizontal surface and the masses of the blocks and the state of motion of the system were considered. There were two possibilities for the first feature: the masses were equal or different. The system could be in one of the three states of motion: at rest, moving with constant velocity or moving with constant acceleration. The two sets of tasks consisted of 24 *multiple choice items* where the possible answers were always the same three. The questions directly concerned the content and the student was required to *explain the procedure* used to answer them.	Each *rule (category) explains a given sequence of replies to the test questions*, so that the *categories can be put into hierarchies*. An individual rule results from a particular combination of decisions about the importance of the different features. Possible combinations range from global, where all decisions are made on the basis of mass only, to the very particular, where the different states of motion are treated differently. *The five rules described accounted for almost two-thirds of the responses* (63%) and the results are shown as a percentage of students using a particular rule against high school background. A match between subjects' responses and the ideal sequences consisted of agreement on not less than 21 of the 24 items, i.e. only three deviations from the ideal sequence were allowed. Finally, another check was made on the rule being used. Subjects were instructed to write an explanation of their procedure after they had completed the 24 items. These explanations were double checked with rule assignments to be sure they were consistent.
Terry *et al.* (1985) attempt to find out what children understand about forces and equilibrium in the static state. The authors are interested in *the intuitive preconceptions of children about forces and equilibrium* and any changes in their conceptual framework that might take place as a result of maturity and instruction. In addition, they are concerned with teaching strategies that might enhance children's conceptual development. They use the terms *intuitive, pre- and misconceptions, conceptual framework and children's conceptions* without providing any definition.	The replies are grouped *inductively* (by percentages) and a conceptual framework (linked to an particular physical situation). The academic content is used as theoretical framework. References are made to studies by other authors (e.g. Watts and Zylbersztajn 1981). Pupils bring to the physics laboratory a view of the world that is derived from their own experiences in that world and from what they have been previously taught. Many of the views that children hold are based upon intuitive pre- and misconceptions of physical phenomena. There is evidence to suggest that children's misconceptions are extremely tenacious.	The participants were pupils from two comprehensive secondary schools, who were at the mid-point of their third, fourth and fifth years of studying physics. A pencil and paper questionnaire with drawings to represent physical situations was used. *Direct and open questions* on the content being sought were posed, *modifying perceptive aspects*. All the students were asked to provide an explanation about different physical situations (all static equilibrium situations) or to draw arrows representing forces on a diagram.	Only one *conceptual framework* was detected, the rest being descriptions of *the most common erroneous replies* for the physical situation presented. Percentages of correct and incorrect answers are provided. No percentages are given for the conceptual framework. *The theoretical framework of the study is the academic content*. No strategy is provided for considering the reliability or validity of the data.

Ruggiero *et al*. (1985) are concerned with *schemes of commonsense knowledge* connecting the three elements, weight, air and gravity to the phenomenon of free fall.	The research is based on the assumption that the knowledge of facts and phenomena used by a child, or any person without specific competence in a given field, is organized into schemes or networks of relationships which are 'scientific' in a broad sense of the word, even though they may differ from the theories or models accepted by the scientific community. This is a *hypothetical-deductive* approach since it involves a hypothesis on the possible concepts, phenomena and relations which are to be investigated. Three steps were followed in the research: a) The historical development was considered; b) The need for the researchers to follow an exercise of commonsense epistemology; (c) Each child was interviewed using a number of different questions concerning facts which could be explained by the same scheme. *The theoretical framework is the academic content.*	Three samples were used ($N_1 = 22$, $N_2 = 8$ (8 N_1), $N_3 = 40$). All were Italian middle school pupils (12-13 years of age). Some questions were chosen from the tests developed previously by Viennot (1979), Gunstone and White (1981) or Watts and Zylbersztajn (1981) but reformulated to better suit to the samples. Questions were presented in different forms: some of them in *written form (paper-and-pencil test) with drawings*, some *orally following a practical demonstration of the physical situation*. Some of them were *open questions*, some *multiple choice*, where pupils were required *to explain their choice*, and some asked pupils for *definitions or to make predictions*.	Typical answers allowed four *schemes* to 16 inferred (one was the physically correct answer), three of them being subdivided into two or three *subschemes*. Possible consistencies among the whole set of answers were sought, for which purpose was presented a first series of questions to the children. From an analysis of the answers four schemes were inferred where the effect of air did not explicitly appear. The same children were then presented with a further series of questions in which explicitly posed the problem of air, while the results of the first series were checked. Finally, a *quantitative reproducibility* study was made, in which a wider sample of children from the same school was interviewed to confirm (or to disprove) any inferences.
Noce *et al*. (1988) proposed extending Ruggiero's (Ruggiero *et al*. 1985) investigation (schemes for interpreting free fall) to different age levels (adulthood and primary school children). They wanted to ascertain if all children made predictions about an observable fact according to some theory (scientific Newtonian theory or an alternative 'commonsense' one) and, therefore, to check the question about the possibility of generalizing the Ruggiero' schemes as the only possible ones for interpreting free fall.	Ruggiero's paper (Ruggiero *et al*. 1985) was used as a starting point, because it was thought that a large number of answers would fit some of the explanatory schemes (*hypothetical-deductive treatment*), but when the answers of the primary school children did not, they described and explained the answers inductively. *The subject's knowledge was described in a theoretical Newtonian framework*.	The authors used a *paper-and-penceil test* in two fifth grade classes in an elementary school. In addition, children were also asked to provide a *definition/description* of 'force of gravity'. This was done in order to bring about a discussion in the classroom about gravity. Then, they extended the investigation to different levels up to adulthood using the same instrument. In addition they performed the same *test* and also *clinical interviews* on a small sample (n = 10) of primary school children (8-12 years old). Questions in the *paper-and-penceil test* were chosen from Ruggiero *et al*. (1985) (*open questions* about physical situations where *predictions and justification* of the answer were required). The *clinical interviews* were followed by an exploration of the pupils' views through the description/interpretation of the events depicted in a series of photographs of the first moon landings.	When the sample was made up of secondary school and university students and adults, the authors used the categories of Ruggiero *et al*. (1985), that is, *explanatory or alternative schemes*, but when the participants were primary school childen, they distinguished between *descriptive* (replies which describe the objects' behaviour) and *explanatory* (replies with a consistent set of relations between notions and facts) *approaches* to account for their answers. The percentages of subjects in each of the categories are presented.

Continued

Table 2—*continued*

1. What are the authors looking for?	*2. Procedures and references*	*3. Technique used in obtaining information from the student*	*4. Data treatment*
Villani and Pacca (1990) proposed to show the *ways of thinking* of Physics Graduate students about the phenomenon of collisions and the physical quantities which are conserved in this process. They found *spontaneous models*, which would be the global interpretations realized by the investigator from the predictions or spontaneous ideas of the students when confronted with a physical situation, in this case involving collisions. The spontaneous ideas are expressed as models.	The authors deduce *spontaneous models inductively* and using the *academic content as theoretical framework*. They consider that the formal solutions given by the students confronted with a situation (in this case, collisions) frequently mask ways of thinking that are not compatible with the mathematical formalism used, although the results are correct. However, certain simple problems which include conditions or factors for their understanding can reveal alternative modes of thinking (McDermot 1984).	The sample was constituted by Physics Graduate students (56 in the first part of the research and 59 in the second with 29 being common to both groups) (Sao Paulo University). *Written and direct questions* (with drawings) and *interviews* about physical situations required the students to make *predictions and explanations* of the effect of simple and multiple collisions. *Variables (and, implicitly, perceptive aspects) were modified.*	The categories were the spontaneous models deduced by the authors from an analysis of the students' answers (students' expectations and explanations). The modles are not presented as percentages. *A certain degree of data validation is attempted* by *by presenting the students with a summary to confirm (or not) the record of their ideas. The study is based on the students' final reply which, according to the authors, present greater internal coherence than the first reply.*
Boeha (1990) aims to evoke *students' Aristotelian-like views* about natural phenomena that they bring with them to the science classroom and provide science educationalists with a repertoire of *common Aristotelian-like beliefs* which have persisted in students. To do that, he reports some Aristotelian-type views in situations involving the concept of force held by students in Papua New Guinea. The author affirms that these views differ in some ways from the orthodox views of physics.	Since it uses a Science History context, the research takes on a *certain hypothetical-deductive character*. It takes into account the contributions made by other studies although *the theoretical framework is the academic content*. Throughout the study, students' Aristotelian-like views are given a respected status that reflect their widespread use, their internal coherence and their tenacity in the face of classroom teaching. It is assumed that all of a student's statements are logically compatible to a listener or reader.	Participants were 126 randomly chosen 12th grade physics students (17 or 18 years old) in a National High School or senior secondary school in Papua New Guinea. The author used a mixture of *pencil and paper questions and interviews* similar to the interview-about-instances of Osborne and Gilbert (1979, 1980). This instrument took on the form of diagrams (line drawings) on individual cards. *Direct questions* were used and the two situations were similar to those used elsewhere (Osborne and Gilbert 1979, 1980, Osborne 1980).	A qualitative study where the *categories* refer to different aspects of the same topic and *are not mutually exclusive*, since *they refer to do to different aspects of the Aristotelian view of force*. Attempts are made by the researcher to construct ideas/beliefs that can account for statements by each student in a manner that the statements are consistent with each other. Students' views have been compiled to give a composite picture of Aristotelian-like ideas.
Finegold and Gorsky (1991) investigate the percentage of students who are *consistent in their beliefs about the forces acting on objects at rest and objects in motion*, i.e. those who consistently apply the correct Newtonian laws and those who consistently apply alternative frameworks. For those students who appear to be inconsistent in their beliefs, the authors attempt to identify their categories of thought and *to look for evidence of an underlying logic in their apparently contradictory beliefs*.	They consider the contributions of other authors not so much to look for support as to throw doubt on them. They agree with Clough and Driver (1986) that outcomes of a study can be dependent on the tool of inquiry used, from which stems the importance of *investigating the consistency*, if any, with which concepts of force are used by individual students in different but closely related contexts. The theoretical context is the academic content itself.	The sample consisted of 534 university students and high school pupils, with different levels of instruction, 35 of whom were interviewed when the had completed the written test. The test was administered as a *written task* or as a *written task accompanied by an interview*. For every item, the students or pupils were presented with a drawing of the system and asked to use arrows to show each of the forces acting on the objects in question and to name each force. *Variables were changed sometimes.* Those interviewed were also asked to explain their answers. *The authors used Ebel's criteria (1972) for evaluating content validity.*	The answers were grouped according to whether they represented *students' beliefs* (groups of similar although not necessarily consistent answers), *students' conceptual categories* (replies obeying an underlying logic although they may seem contradictory) or *conceptual frameworks* (groups of replies consistently applied to different physical situations). *The authors attempt to relate the greatest number of data, looking for consistency or coherence.* They do not establish hierarchies.

Thijs (1992) evaluates the effectiveness of the constructivist teaching approach in reducing *misconceptions* on force and attempts to identify which students benefit from that approach. The study attempts to answer the following questions: What are the specific learning effects of the course with regard to forces on objects at rest, and forces on objects in motion? What are the differences in results between bright and weak students, and between boys and girls, as regards the appreciation of the constructivist approach and the learning effects of the course?	A basically *inductive* study, although it takes into consideration the contributions of some previous studies about students' intuitive ideas on 'force' (the essential misconceptions in the subject area of force and motion lead to some misinterpretations) and these are taken into account when the question papers are written. The author adopts a constructivist approach to teaching and learning: 'learning involves the generation and restructuring of students' conceptions (Driver and Oldham 1986)'. Firmly placed in the current of conceptual change (Posner *et al.* 1982).	The sample consisted of 190 students in a Dutch secondary school, of whom 99 were from the VWO stream and 91 from HAYO; 92 of the total were girls. Initially, *the paper and pencil questionnaire was a type of two-tier multiple choice test* (Treagust 1988), of which the first tier of a question required a content response, and the second tier required a reason for the response (although the students were also given the opportunity to give their own arguments). Subsequently, *in light of the results obtained the format was changed* and students were given a simple *multiple choice questionnaire* in which they had the opportunity of making a *written comment. Direct questions* were used in both cases.	The answers to the questions were placed in categories and the authors differentiated between answering categories according to rest situations and to moving object situations. Behind each category lay a *misconception* which led to the given type of answer. *Evaluation of the replies took the academic content as the theoretical framework.* The percentages of correct answers before and after instruction were compared. The authors did not look for possible relationships between groups of replies and did not establish hierarchies.
Galili and Bar (1992) wanted to test the hypothesis that *the conceptual misunderstanding* of motion might survive after the instructional process. They also wanted to check how genuine and stable is the process of conceptual change that takes place when students study the main dictates of classical dynamics, which factors might be of importance for the success of this process, and how it can best be tested. They distinguish between *misconceptions* (drawn from the replies given) and *intuitive beliefs* (based on experiences of the senses).	The authors maintain that the obvious discrepancies between formal principles of classical mechanics for simple idealized systems and the actual behaviour of real, more complicated systems in the world around us create many of the *wrong preconceptions* which 'naive' (pre-instructional) students usually posses about nature. Furthermore, correct answers are not always proof of correct reasoning. *Misconceptions are* sought inductively.	The research sample included pre-and post-instructional high school students (33 of 10th-grade, 60 of 11th-grade, 36 of 12-grade), students of a University Pre-academic Study Department (27) and preservice teachers in a Technology Teacher Training (19). An *open paper and pencil test* about physical situations with *direct questions* was administered, where students were asked to draw arrows (to represent forces) and provide an *explanation.*	*Misconceptions* or categories of the given answers are not mutually exclusive and one reply may belong to more than one category. All seem to reflect pre-Newtonian views and some the same *intuitive belief.* Moreover, the general results of the research are presented in three complementary mutally exclusive categories (as percentages): success, pre-Newtonian views and no-reply.
Reynoso *et al.* (1993) analyse the different *frameworks* concerning free fall presented by Mexican students and teachers in primary, secondary and pre-university schools.	The investigation procedure developed at the University of Rome (Ruggierio *et al*. 1985, Noce *et al.* 1988) served as a reference for the study reported in this paper. The authors carry out a more detailed *inductive* analysis of the content of some ideas mentioned in other studies, different frameworks and their evaluation as a consequence of the incorporation of academic knowledge. *The academic content serves as theoretical framework.*	The sample included 33 primary school students, 158 secondary school students, 111 pre-university school students, 19 primary teachers and 20 pre-university teachers. *Paper and pencil tests (with direct and open questions*, where pupils were required to make a *definition*, an *explanation* or make a *prediction* relating to different physical situations where *the context was changed*) and *personal interviews* were used. In the case of primary school children, only the personal interviews were taken into consideration. Some questions were taken from an article by Ruggiero *et al.* (1985).	The replies were classified into *frameworks* and *subframeworks* (as percentages). Without making hierarchies in the categories, the different frameworks and subframeworks are shown for different age groups. Any relation between frameworks is determined by the different ways in which the fall of objects on earth and the moon are treated. *The data are evaluated by reference to the academic content.*

Continued

Table 2—*continued*

1. What are the authors looking for?	*2. Procedures and references*	*3. Technique used in obtaining information from the student*	*4. Data treatment*
Bar *et al*. (1994) wanted to expand previous researches to include children from a broader range of ages to be able to draw stronger inferences about *children's ideas* about weight and free fall. Their intention was to: 1) identify children's common ideas, at various age levels, 2) determine how these ideas influenced their predictions, 3) correlate the development of children's conceptions of weight of deformable bodies, 4) test the consistency of children's ideas 5) verify children's ideas about the role of the earth in the process of free fall.	*A hypothetical-deductive procedure* is used to ascertain the usefulness of the 'common sense theory of motion' suggested by Ogborn (1985). The notions taken from this theory were: support, heaviness and lightness, weight as a pressing force and the need of a medium to transmit forces. *The academic content formed the theoretical framework*.	The sample comprised 400 children (between the ages of 4 and 13) from a kindergarten, an elementary school and a junior high school in Jerusalem (20 boys and 20 girls. These children had not been formally taught). Each child was *interviewed* for about a half hour, during which the tasks were presented with demonstrations and pupils were required to make an *explanation* or *prediction*. *This involved a Piagetian task, that is, the conservation of weight of deformable bodies (plasticine)*.	The categrories were made using the '*common sense theory of motion*' suggested by Ogborn (1985) and also in age groups so that *certain hierarchies appeared in the groupings*. This made it possible to make a proposal about the possible development of children's ideas with age. Even so, *the data were evaluated by taking the academic content as theoretical framework*.
Twigger *et al*. (1994) attempt to identify *common prior conceptions* in the reasoning of students aged 10-15 about aspects of horizontal and vertical motion and explore the extent to which these prior conceptions are age-dependent.	They describe the students' conception *inductively*. This interview study was conducted as an initial diagnostic phase in the Conceptual Change in Science (CCIS) Project (Driver and Scanlon 1988, Twigger *et al*. 1991). The project team recognized that students have prior conceptions about aspects of force and motion which differ from Newtonian theory (Gunstone and Watts 1985, McDermott 1984). *The theoretical framework is the academic content*.	A total of 36 students aged 10-15 were selected for interview. The *interviews were conducted in pairs* to encourage discussion. Each individual student, however, was asked to give his/her own response to each question and, where required, to write or draw answers. Questions were related to real physical situations, drawings or videos. Pupils were required to give their meaning of force, make *predictions*, *check or give explanations*.	The categories described were the *students' prior conceptions*, which reflected a number of features of students' conceptualization of forces and motion as identified from the responses to the tasks. No hierarchies are established. *Data are evaluated using the academic content as theoretical framework*.

Kuiper and Mondlane (1994) aim to study the *framework* character of student understanding of force.	They use the term '*student idea*' which is more neutral. Ideas are seen in this study as 'units' of a student's understanding of a concept; each student idea refers to one particular problem situation, context and aspect of force. The term '*Framework*' can be understood to mean that a particular student has a set of student ideas concerning one and the same concept which appear to be logically coherent and ordered (at least to some extent). The search for ideas and frameworks is *inductive*, although once the authors define the characteristics of how a framework is manifested, they make some hypothetical deductions. *As theoretical framework, the academic content is used, since student ideas are classified by their distance from this content.*	The sample from Zimbabwe (which is the one mainly discussed in the article) consists of 143 students from Forms 1, 2, 3, 4, 5 and 6. A variable indicating the amount of science instruction received ('ASIR') is created. The other samples are 266 secondary school students (Form 1 to Form 6) from The Netherlands (used mainly as a reference), and 29, 50, 50, and 83 students from pre-university science courses in other countries. A *paper and pencil test* was used which consisted of *four different problem areas, in each of which the same problem was put four times, only placed in different contexts* (no examples were given). The test consisted of *multiple choice items and open questions* to account for the answer chosen. Physical situations in drawings were presented to students and they were asked *direct questions* and for *explanations* and *predictions*.	In all the problem areas two main types of incorrect idea were used besides the correct answer: intuitive and intermediate ideas (*categories of reply based on the distance from the academic content*). There appears to be a progression in the answering patterns of students: intuitive-intermediate-correct ideas. *The data are evaluated according to the academic content.* A factor analysis of the 16 questions of the test was made in order to see whether they could be grouped into four factors representing the four problem areas. *To further study consistencies in answering patterns, a cluster analysis was performed*, in order to group students into clusters with similar answering patterns. Each of these patterns was studied more closely *to check for logical coherence* and to see to what extent the physics contents of the sets of ideas used by a particular cluster of students formed an alternative framework or not.
Montanero *et al.* (1995) analyse how subjects explain interactions between two bodies and look for similarities and differences with the official explanations of classical dynamics. The authors attempt to go further than the descriptive and search for the roots of these conceptions, investigating the possible *laws* the subjects may devise to form their *implicit theories*.	*A hypothetical-deductive treatment* since a preliminary study provides the data for elaborating a series of hypotheses to be taken int oaccount when designing the final test. The authors consider that *students' spontaneous conceptions* can be grouped into *implicit theories* with a certain degree of consistency (Claxton 1984). Comparisons with classical dynamics means that *the theoretical framework is the academic content*.	A *preliminary test* was given to three groups of forty 16-17 years old pupils. *Oral interviews* and *group discussions* were also used. The *definite test* was given to 414 students (between 13 and 18 years old), 47 from the 1st year of chemistry and 21 of an undergraduate physics course (3rd year). The test was also given to 28 teachers. This test was futher studied by means of 18 '*clinical*' *interviews* given to 14-16 years old students. Both the preliminary and definitive test were *paper and pencil test* with *multiple choice* and *direct questions* where *explanations* were required. *Contexts were modified but which exactly is not explained.*	*Emphasis is laid on coherent replies* and these are used to propose *implicit theories*, from which the laws used are defined. Each law reflects a different aspect of collisions. No hierarchies are established. Percentages are given for the educational levels of the number of students replying consistently.

Table 3. Schematic categorization of the studies

1. *What are the authors looking for?*	2. *Procedures and references*	3. *Techniques used to gather information from students*		4. *Criteria for grouping or or relating data*	5. *Test-tuning and studies on the coherence of the data*
A. Searching for regularities such as: 1. laws 2. models 3. strategies 4. beliefs 5. theories 6. common ideas 7. conceptions 8. schemes or frameworks 9. ideas accepted during the history of science	A. Sequence: 1. inductive 2. hyp.-deductive 3. combination B. Framework: 1. academic content 2. historical conceptions 3. rules and models C. Other works are used to: 1. confirm data 2. use techniques 3. evaluate data 4. continue	A. Paper and pencil: 1. open 2. closed with no explanation 3. closed with explanation B. Interview: 1. individual 2. in pairs C. Type of question: 1. direct question 2. explanations 3. predictions	D. Situations: 1. in drawings 2. real E. Variations in the test: 1. situational 2. relevant 3. irrelevant 4. age 0. no variation	A. Level of data abstraction: 1. grouping according to similarities of the replies 2. grouping from A1 3. grouping from A2 B. Filtering data using standards: 1. rules and mental models 2. ideas during the history of science	A. Test-tuning: 1. previous tasks 2. double take (according to answers) 3. double take (according to validity) 0. no adjustment B. is the coherence of the replies studied?: 1. test given several times 2. comparison between groups 3. cluster 4. models 5. reproduction 6. confirmation replies 0. no such studies made
1. Viennot (1979) A1(A8) $\rightarrow$ A3	A1 B1	A12 C12	D1 E1	A3	A0 + B0
2. Watts and Zylbersztajn (1981) A8	A1 B1 C4	A3 C12	D1 E1	A2	A1 + B0
3. Watts (1982) A8	A1 B1 C4	B1 C2	D1 E1	A2	A1 + B0
4. Watts (1983) A8	A1 B1 C4	B1 C12	D1 E1	A2	A1 + B0
5. Maloney (1984) A23	A2 B13	A3 C1	D1 E23	A2 + B1	A0 + B4
6. Terry *et al.* (1985) A7	A1 B1	A1 C12	D1 E14	A12	A0 + B0

7. Ruggiero *et al.* (1985) A8	A3 B12 C24	M_1:A1	C123	D1	E1	A2	A2 + B5
		M_2:A13	C123	D1			
8. Noce *et al.* (1988) A8	A3 B1 C124	M_1:A1	C123	D1	E14	A1	A1 + B6
		M_2:A1	C23			A2	
		M_3:A1 B1	C23			A1	
9. Villani and Pacca (1990) A2	A1 B1	B1	C123	D1	E2	A2	A0 + B6
10. Boeha (1990) A9	A3 B12 C2	A1 B1	C123	D1	E0	A2 B2	A1 + B0
11. Finegold and Gorsky (1991) A4 (A8)	A3 B1 C3	A1 B1	C12	D1	E12	A1	A3 + B2
12. Thijs (1992) A7	A1 B1	A3	C1	D1	E0	A2	A2 + B1
13. Gallili and Bar (1992) A7 + A4	A1 B1	A1	C12	D1	E14	A2	A0 + B0
14. Reynoso *et al.* (1993) A8	A1 B1 C12	M_1:B1	C123	D1	E124	A2	A1 + B0
		M_2:A1 B1					
15. Bar *et al.* (1993) A6	A2 B13 C24	B1	C23	D2	E134	A2 B1	A0 + B0
16. Twigger *et al.* (1994) A7	A1 B1	B2	C123	D12	E14	A1	A0 + B0
17. Kuiper and Mondlane (1994) A8	A3 B1	A13	C123	D1	E14	A2	A0 + B35
18. Montanero *et al.* (1995) A5	A3 B1	A2 B1	C12	D1	E1234	A2	A2 + B2

E. Variations in the test: 1. situational. The problem situations are changed, while the concept in question remains the same; 2. relevant. Different questions are posed for the same situation by varying factors that are relevant for the result (effect); 3. irrelevant. For the same situation, different questions are posed by varying factors that are irrelevant for the result; 4. age. The same questionnaire is presented to subjects of different ages; 0. no variation.

Data treatment

In the last column of table 2 we describe the way in which the data are treated, both at the descriptive (analysing the procedures used to group them or to establish relationships between groups) and interpretative level.

The schema for acquiring data in a homogeneous way could be:

- Is any attempt made to group the replies which show some regularity? If so, what do the percentages refer to? What, if any, statistical treatment is applied?
- Are the groups or categories related among themselves? Are they related hierarchically? If so, what criterion is used?
- Is some procedure used to study the quality, validity and/or reliability of the data? If so, which?
- Other details of interest.

The most significant data have been highlighted, in italics those which are common to various authors and define some sort of regularity, and underlined those which are particular or diverge from the norm. This column (column 4) of table 2 has been divided into two columns in table 3 (columns 4 and 5) as follows:

(1) Criteria for grouping or relating data. The following categories have been considered:

A. Level of data abstraction.
   A.1 Grouping according to similarities of the replies: students' conceptions are obtained by grouping their replies according to criteria of similarity and difference.
   A.2 Grouping from A1: from A1, new groupings are made with a greater level of generality by comparing the types of reply obtained for a situation according to different criteria of variation, such as: (i) changes of situations with the same underlying content (variations in context); (ii) modifications of the variables which change the described situations (relevant variations); and (iii) variation of factors which do not alter the situation (irrelevant variations).
   A.3 Grouping from A2: from A2, the authors make greater abstractions and look for ways of reasoning, schemes, laws, explicative models, etc. underlying the A2 groupings.
B. Filtering data using standards.
   B.1 Rules and mental models.
   B.2 Ideas corresponding to those defended by researchers during the history of science.

(2). Test-tuning and studies on the coherence of the data:

A. Test-tuning.

A.1 Previous tasks: the questions and results of other authors are taken into consideration.

A.2 Double-take (according to answers): this is a technique whereby the data obtained from the questionnaire are used to discriminate between the questions that provide significant information from those that do not. These last questions are eliminated or remodelled before the test is presented again.

A.3 Double-take (according to validity): in this case statistical modules are used to analyse the degree of validity of the questions so that some may be eliminated or remodelled.

A.0 No adjustment.

B. Is the coherence of the replies studied?

B.1 The test is given several times: each time it is improved and made more precise.

B.2 Comparision between groups: the results given by several groups which have taken the test are compared to ascertain any coherence.

B.3 Cluster: cluster analysis to ascertain the coherence of the data.

B.4 Models: models from the history of science or from modern-day science are used to analyse the degree to correspondence of the data with these models.

B.5 Reproduction: the same test is given to other subjects to see if the answers are reproduced.

B.6 Confirmation of the replies: the students are interviewed individually to see if their replies are confirmed.

B.0 No such studies are made.

Both categorical groups have the same objective: to ascertain the degree of validity and reliability of the data obtained. In the first case, because the questions are adjusted to the students' supposed knowledge and in the second case in an attempt to evaluate the data obtained in the test.

Data analysis

The relations between the data of the columns of table 3 mean that an analysis of each column is inadvisable, therefore we decided to make a more overall analysis.

Investigations are mainly carried out inductively and descriptively. The data contained in columns 1 and 2 of table 3 are mainly used for this analysis. Most authors are clearly seeking a certain degree of regularity in the students' replies, which leads them to find:

- Schemes (alternative or spontaneous), such as coherent ideas concerning the world constructed from the students' own experience, which would permit predictions, explanations and other consistent or coherent manifestations when confronted with changing tasks (Watts and Zylbersztajn 1981; Watts 1982, 1983; Ruggiero *et al.* 1985; Noce *et al.* 1988; Finegold and Gorsky 1991; Reynoso *et al.* 1993; Kuiper and Mondlane 1994).
- Laws, models, strategies, beliefs, implicit theories, etc. that govern the students' reasoning and that underlie the students' answers and are as

important as their conceptions or spontaneous ideas (Viennot 1979; Maloney 1984; Villani and Pacca 1990; Finegold and Gorsky 1991; Montanero *et al.* 1995).
- Wrong answers or misinterpretations that appear with a significant frequency. The 'distance' of the replies from the correct answers is emphasized and the most frequent errors are categorized (Terry *et al.* 1985; Thijs 1992), as are the replies which would represent prenewtonian thinking (Boeha 1990; Galili and Bar 1992). Another common thread in the four mentioned studies, besides their search for erroneous conceptions, is that they look for constructivist strategies of teaching to overcome such cognitive shortfalls.

Such a preoccupation in looking for 'schemes', 'laws', 'models', etc. might suggest that constructs belonging to some cognitive model concerning the individual's mental organization are being referred to, although the explicit or implicit absence of such models leads us to discard such an interpretation.

The terms 'schemes', 'laws', 'models', etc. when used to draw up the tests or to evaluate the replies, reflect the need to find a certain regularity in the students' replies rather than an unobservable element in the subject's cognition. In this sense, the 'schemes' in these studies are more 'visible' and 'tangible' than those referred to in cognitive theories of learning (e.g. Pascual-Leone 1979; Rumelhart and Ortony 1982; Case 1983; Pozo 1989).

It seems that the signifier rather than the true meaning of the terms was taken from the theoretical framework referring to 'schema', 'model', 'strategy', etc. (presumably cognitive theories), and that, according to bibliographical references, Driver and Easley (1978) were mainly responsible for this transfer (see also Solomon 1994). Thus, there seems to be a tendency in most of the studies consulted *to describe the students' knowledge in an inductive manner*, except in those cases where the use of certain initial suppositions point to certain expectations. For example:

- The 'technique of rule assessment' makes it possible to contrast ideal strategies with those really used by students, so that each rule would explain a given sequence of replies to the questions contained in the test (Maloney 1984).
- In an attempt to prove the usefulness of 'Ogborn's common sense theory' in explaining why things fall, using notions taken from this theory such as support, heaviness and lightness, weight as a pressing, force and the need of a medium to transmit forces, certain suppositions are made at the outset which orientate the investigation (Bar *et al.* 1994).
- Supposedly Aristotelian viewpoints held by students create certain expectations (Boeha 1990).
- In most cases, the hypotheses are established from the concepts, phenomena and relationships regarding the content being investigated (Ruggerio *et al.* 1985; Noce *et al.* 1988). Without being so explicit, other authors also create expectations of what must be looked for, but they are always led by the teaching content which is the object of the research.

As the academic knowledge is different from the student's knowledge, the expectations arising from the academic content are not as good as those inferred from a student's cognition model. Without such expectations the studies are predominantly inductive and, as regards data treatment and evaluation, are merely descriptive and cannot explain or relate the replies of students (Piaget and García 1973; Bunge 1981; Moreno *et al*. 1993), except in very few cases.

Decision-making is influenced by the content about which conceptions are being sought. This group of analyses is based on data from columns 3 and 4 of table 3 and, to a lesser extent, of table 2.

During a study to determine students' conceptions certain decisions must be taken, and these are best illustrated in the design of the test and the treatment of the data thus obtained. An author's underlying criteria can normally be glimpsed in the decisions she/he takes as regards: (i) a certain type of question, problem situation or strategy followed in collecting data; (ii) the way in which data are grouped and the related and the categories to which replies are assigned.

a) *As regards the test questions and techniques used to obtain information from the student*

The most usual way of obtaining such information is a written exam on physical situations presented in drawings. This may consist of an open-ended or multiple choice question, in which an explanation is usually demanded for the why a particular answer has been chosen. Individual interviews or in pairs, in which the situations are presented in the form of drawings or photographs (Watts 1982, 1983; Noce *et al*. 1988; Boeha 1990; Villani and Pacca 1990; Finegold and Gorsky 1991; Reynoso *et al*. 1993; Montanero *et al*. 1995) or using real objects (Bar *et al*. 1994; Twigger *et al*. 1994) are less frequent. Half of the studies use a mixture of these techniques.

Most studies share a common characteristic, which consists of asking a high percentage of questions in which the signifier of the concepts or laws that are the object of the research are directly referred to. This involves using expessions and situations suggested by or taken directly from the 'academic' context. For example, is there any resulting force on the trolley? Is the resultant force constant or increasing? Which is exercising the greater force, the block or the water? What forces act on the moon? etc. Some questions ask for definitions of such terms (Ruggiero *et al.* 1985; Noce *et al*. 1988; Reynoso *et al*. 1993; Twigger *et al*. 1994).

Predictions are sometimes asked for (Viennot 1979; Ruggiero *et al*. 1985; Noce *et al*. 1988; Villani and Pacca 1990; Reynoso *et al*. 1993; Bar *et al*. 1994; Twigger *et al*. 1994; Kuiper and Mondlane 1994) and, sometimes, such predictions must be confirmed by experience (Twigger *et al*. 1994).

It is clear that the questions and situations making up the test are principally taken from the academic world in which the authors move. However, perhaps we should ask ourselves if a more suitable context exists than the strictly academic, a question we shall attempt to answer in the next section.

b) *The procedure used to group and relate the students' replies*

Judging by the data contained in the fourth column, no study loses sight of the *academic content* when grouping students' replies. For this reason, the groupings present a certain dichotomy (correct/incorrect), within which the following categories may be found:

- most frequent errors (Viennot 1979; Watts 1982, 1983; Maloney 1984; Terry *et al.* 1985; Villani and Pacca 1990; Thijs 1992; Twigger *et al.* 1994);
- consistent or coherent (Finegold and Gorsky 1991; Kuiper and Mondlane 1994; Montanero *et al.* 1995);
- prenewtonian (Boeha 1990; Galili and Bar 1992);
- incorrect in multiple choice test (Watts and Zylbersztajn 1981);
- most frequent in a longitudinal study (Bar *et al.* 1994); or
- correct or nearly correct (Ruggerio *et al.* 1985; Noce *et al.* 1988; Reynoso *et al.* 1993).

Usually, the categorization process finishes when the replies have been grouped and, perhaps, expressed as a percentage. However, some studies attempt to relate groups (or categories) of replies either amongst themselves or to aspects external to the test. For example:

- analogies are established between the types of reply and prenewtonian science (Boeha 1990; Galili and Bar 1992);
- a certain hierarchy between groups is created. This may be based on the logical or semantic structure of the concepts being sought (Viennot 1979; Maloney 1984; Ruggiero *et al.* 1985; Noce *et al.* 1988), the degree of complexity of the replies (Noce *et al.* 1988; Bar *et al.* 1994) or the 'distance' of the replies from the academically acceptable answers (Kuiper and Mondlane 1994).

In other words, 'the academic content' is again the criterion used to establish such relations. However, perhaps this is not the only option available for grouping the replies or for establishing relations between the groups.

Among such alternatives which, furthermore, would complement the criteria based on the academic content under examination, one might include general cognitive limitations, difficulties in assimilating new information or the transforming capacity of schemes, all of which comes under general tendencies in students' cognition (Marín 1995). As long as there are no arguments to the contrary, there is no reason to opt for one or other of the two poles, the content and the students' cognition.

c) *The validity and reliability of the information obtained from the student is not usually analysed*

In this analysis, we have mainly used the data from column 5 of table 3 and, to a lesser extent, from columns 3 and 4 of the same table.

In column 5 we now analyse the preoccupation of authors to evaluate the validity and reliability of the information gathered. Other columns also illustrate such a preoccupation. For example:

- Category/group A of column 4 shows whether all the data are treated in the same way during grouping or whether only those which fulfil the prerequisite regularity criteria are used.
- Category/group E of column 3 illustrates some of the methods used by authors to collect data in order to ascertain the existence of any regularity in the replies.

From the above data we can state:

- Two thirds of the studies do not analyse the reliability and validity of the data.
- A third of the studies that make such an analysis do so only partially. Ruggiero *et al.* (1985), for example, make quantitative studies in which the test is applied to a bigger group; Villani and Pacca (1990) attempt to weigh up the reproducibility of the replies in subjects of the same group; Finegold and Gorsky (1991) compare groups of answers in different samples; Kuiper and Mondlane (1994) consider the coherence of the results by cluster analysis and, finally, Montanero *et al.* (1995) make comparisons between groups to emphasize the replies showing a certain coherence.
- A regards the use of changed situations or variation in relevant and irrelevant factors, the first strategy is the most widely used and a third use the second. Only two studies (Maloney 1984; Montanero *et al.* 1995) were observed to use irrelevant factors. If these three strategies are not involved simultaneously in a task, it is very difficult to weigh up the three requirements of regularity for a reply to be seen as reflecting some scheme of knowledge: repetition, generalization and differentiation (Piaget 1980; Marín 1995).
- The most recent studies (since 1992) use samples of different ages, although this technique is not accompanied by what would seem to be a necessary complement, the use of hierarchies of categories to reflect the cognitive development of the specific knowledge that is being analysed. Only Bar *et al.* (1994) provide this hierarchy, while Reynoso *et al.* (1993) do so to a certain extent (schemes and subschemes). Valid techniques for internal evaluation cannot be applied if this hierarchical order is not established.

When attempting to identify and interpret data of such a complex phenomenology as students' knowledge, a minimum of measures must be taken, without which it is not possible to ascertain the source students' replies or separate the fundamental from the incidental.

Most relevant characteristics of the methodology used to describe student knowledge

The most relevant methodological characteristics of the studies on conceptions follow.

First, the replies of the students in response to the situations and questions presented are more or less implicitly catalogued and interpreted by taking *the academic content as reference*. This is the most general characteristic and one which conditions the rest.

Second, *in most cases conceptions are sought inductively* (no models are used to explain their presence or their possible relation with other constructs of the student's cognitive structure). These characteristics are closely linked with the above since the absence of models to explain how the student organizes, develops and realizes his/her mental acitivity, means that the researcher does not normally have any expectations or hypothesis as to what might be found or how the information is best extracted from the student.

Third, *it is not common practice to analyse the validity and reliability of the information obtained from the students*. In the search for and intepretation of a phenomenology so complex as cognitive development, certain measures should be taken simultaneously and systematically. These include the need to:

- Differentiate the replies which reflect some scheme of knowledge from those that are 'automatic' or simply inventions, in response to the type of question posed by the researcher or the student's lack of interest.
- Analyse the coherence of the data categorized, treating such data as a matrix of cases (subjects interviewed) and variables (categories), making correlation studies, grouping cases according to the similarity of replies, correspondence analyses, etc. All this would make it possible to verify whether the orders established in the categorization were more or less correct with respect to the general order of the data (internal validation).
- Make studies involving samples of different ages, making it possible to discriminate groups of data which are best adapted to a given stage of cognitive evolution from others which present anomalies with respect to the evolution (Marín 1995; Benarroch 1998).
- Compare the results concerning a given unknown aspect with others related with the student and which are better known such as: the cognitive level, field dependence-independence, replies to a test of proven reliability, etc.

Fourth, *the information taken from 'what a student knows' is very restricted*. From the bibliographical data, we cannot evaluate the quality and quantity of the information obtained from the students if only inductive procedures are used. However, when the three previous conclusions are taken into account, together with external factors linked to the types of knowledge which a student may contain in his/her cognitive baggage, we observe that:

- At the declarative knowledge level, only those ideas directly related with the content are taken into account. However, there are two cases in which it is not usual to detect student conceptions:
  — Ideas which are not linked to the content in an adult scientific logical sense, are linked for the student to such an extent that they can be used to understand or assimilate the content (Piaget 1977). Put another way, the academic content refers to objects, situations, phenomenologies, which make it possible to prepare certain situations and questions which can be answered adequately if one possesses this content. The student possesses a knowledge in this respect which permits him to infer different replies or ones far from 'correct' ones.
  — Ideas which are still in a developmental stage 'far' from the content to be taught (Piaget and García 1973), that is, at the moment the test is taken, the student possesses ideas which are little developed, intuitive or preoperative. These are sufficient for him/her to reply to the questions although after a certain (perhaps quite long) time during which such ideas would develop and be enriched, they would have permitted a correct answer or, at least, a more appropriate one. As the student's knowledge of the situations described evolve, he/she may be able to provide answers which are more correct or, at least, more analogous, while at the cognitive level at which his/her ideas are found they are only capable of providing

answers 'far from the academic content'.

The fact that in both cases the students' ideas are different from the academic does not mean they are less important since they can be used to understand or assimilate the teacher's explanation of a given content. Ideas that are unlinked to the content in an adult scientific logical sense may be linked for the student (Piaget 1977). This type of idea is very difficult to detect if only the academic content is taken as reference in the search for information.

- The students' procedural capacity (to classify empirical data, infer inductively, reason proportionally, to control variables, formulate hypotheses, etc.) and motor skills (the ability to construct educational material or experimental set-ups) are almost completely absent from these studies despite the fact that such capacities are determining factors in student performance in science (e.g. Shayer and Adey 1984; Niaz 1991; Lawson 1993).

Methodological suggestions

If the object of the research is only to understand the scholastic knowledge that students have gained from previous years' science classes, the methodological characteristics observed in the studies we have analysed would have a certain validity since they are perfectly suitable for comparing and contrasting this knowledge with a standard (in this case the academic content itself). In such a case, a questionnaire could closely resemble the tests that are frequently used to evaluate a student's progress at the end of an academic year.

The problem is that students spontaneously develop many conceptions through interaction with the social, cultural and physical environment. In such a situation, the academic content is not the most suitable reference for assessing this knowledge and the methodologies described above are inadequate. Characteristics of these methodologies are:

(1) There are more suitable contexts than the academic which may help avoid skewing observed in the obtaining and interpretation of data about student knowledge.
(2) No attempt is made to weigh up information to avoid such skewing.
(3) The information obtained from the students is only a small part of that which could be considered useful for teaching purposes.

Deficiency (1) is a problem that has much to do with the professional training of the investigator. 'Normal' scientific training is not necessarily well-suited to understanding students' conceptions since we are dealing with a cognitive problem which a cognitive theoretical context is best suited to describing and interpreting.

Research into conceptions would be better served by a combination of two interpretative frameworks:

- The teaching content and, by extension, the scientific knowledge which acts as a base to this content, which would serve, for example, to structure the research content.
- A theoretical context taken from the field of cognitive psychology, which has tackled the development and organization of knowledge in children and adolescents with some degree of success. This framework could help, for

example, design the most suitable way of making the student interact with physical situations presented in the interview in order to obtain the most reliable information possible. It would also help decide the method best suited to describing the student's spontaneous schemes avoiding to the greatest extent possible any skewing and distortion.

A good cognitive model would help avoid the inductive/descriptive studies which are so common in the publications consulted and enable the design of hypothetical-deductive strategies to identify, describe, intepret and relate the data obtained.

Suggestions for overcoming deficiency (2) are based on the idea that the replies given by students do not always reflect their knowledge either because they are 'tainted' to a greater or lesser degree by the 'investigator's knowledge' or because they are given randomly, invented or constructed 'in situ'. If such is the case we must put into pratice methodological mechanisms which diminish to the greatest extent possible skewing and distortion. This can be done by using a cognitive model which permits the researcher to get close to the student's cognition. The use of a good individual interview dynamic (following Piaget), which avoids academically 'marked' questions, is a prerequisite.

Such mechanisms will also separate the relevant from the non-relevant information. This is a complex area since it will be necessary to combine qualitative and quantitative strategies before and after data collection. Before, protocols must be designed to bring into play as much as possible the students' schemes of knowledge by using, for example, the above mentioned variation strategies (see the paragraph on the criteria for grouping or relating data, above). After, the data must be submitted to a statistical treatment which evaluates their coherence.

As for deficiency (3), which deals with the restricted information collected from the student, the findings of research into students' conceptions will be very different if: (i) we do without a model and simply attempt to ascertain what a student knows on a given academic content; (ii) the model can be assimilated into a conceptual framework (for example, Ausubel 1982); or (iii) the model contemplates the existence of an important procedural component associated to declarative knowledge or that there is a strong component of implicit knowledge (for example, see Karmiloff-Smith 1994; Vosniadou 1999).

From our review of papers published on student conceptions, deficiency (1) is the most common but if we take into account deficiency (3) and the fact that a large part of students' knowledge will have been generated spontaneously in their daily environment, much more than that found in the studies should be expected of the students, for example: (i) previous ideas unrelated with the academic, (ii) sensorimeter or graphic manifestations of implicit knowledge; or (iii) a knowledge of procedures (see the paragraph on the most relevant characteristics of the methodology used to describe student knowledge). To uncover this new kind of information, the type of questionnaire seen in the publications examined will have to be redesigned in order to obtain more and better information of didactic interest (for example, see Marín 1995). However, this exceeds the limitations of the present article.

We should not finish this list of suggestions without referring to perhaps the most important for the science class: any questionnaire must be given to a wide age range in order to avoid the static image of conceptions that is normally given in the articles analysed, and another type of dynamic must be sought which describes the different levels of knowledge through which students pass in accordance with their age.

Final thoughts

A simple comparison of the papers published on conceptions in ACM and the experiments of Piaget in the 1950, 1960 and 1970s (see Marín and Benarroch 1994) leads one to question the supposed progress which the former is supposed to represent for science teaching (see Gilbert and Swift 1985). It also shows the need to look for a more suitable theoretical context than is normally the case in such studies.

For several years now there seems to have been a change taking place as regards thinking into student conceptions. There is a certain tendency, by no means accepted by the whole science teaching community, to raise critical questions about the subject of conceptions:

- A change from an interest in alternative conceptions, previous ideas, etc. as regularities in student replies towards constructs as mental models, implicit schemes and theories to explain conceptions (Vosniadou 1994; Pozo and Gómez Crespo 1998; Marín 1999; Oliva 1999).
- A questioning of the basic methods used (Gutiérrez and Ogborn 1992; Vosniadou and Brewer 1992; Benarroch 1998).
- A significant increase in studies involving samples of different ages (e.g. Galili and Bar 1992; Reynoso *et al.* 1993; Bar *et al.* 1994; Kuiper and Mondlane 1994; Twigger *et al.* 1994; Montanero *et al.* 1995).
- A determination to look for a theoretical basis beyond the pragmatism and inductivism that characterized the first investigations into the field (Chi 1992; Lawson 1994; Vosniadou 1994; Glynn and Duit 1995; Pozo and Gómez Crespo 1998; Vosniadou and Ionnaides 1998; Marín *et al.* 2000).

References

AUSUBEL, D. P. (1982) *Psicología educativa, 'Un punto de vista cognoscitivo'* (Mexico: Trillas).

ACEVEDO, J. A., BOLIVAR, J. P., LÓPEZ MOLINA, E. J. and TRUJILLO, M. (1989) Sobre las concepciones en dinámica elemental de los adolescentes formales y concretos y el cambio metodológico. *Enseñanza de las Ciencias*, 7, 27-34.

BAR, V. (1989) Introducing mechanics at the elementary school. *Physics Education*, 24, 348-352.

BAR, V., ZINN, B., GOLDMUNTZ, R. and SNEIDER, C. (1994) Children's concepts about weight and free fall. *Science Education*, 78, 149-169.

BENARROCH, A., (1998) Las explicaciones de los estudiantes sobre las manifestaciones corpusculares de la materia. (tesis inédita), Facultad de Educación, Universidad de Granada.

BOEHA, B. B. (1990) Aristotle, alive and well in Papua New Guinea science classrooms. *Physics Education*, 25, 280-283.

BROWN, D. E. (1989) Students' concept of force: the importance of understanding third Newton's law. *Physics Education*, 24, 353-358.

BUNGE, M. (1981) *La investigación científica* (Barcelona: Ariel).

CASE, R. (1983) *El desarrollo intelectual: una reinterpretación sistemática*. In M. Carretero and J. A. Madruga (eds), Lecturas de psicología del pensamiento. Razonamiento, solución de problemas and desarrollo cognitivo (Madrid: Alianza Editorial) 339-362.

CHI, M. (1992) Conceptual change within and across ontological categories: examples from learning and discovery in science. In R. Giere (ed.), *Cognitive models of Science: Minnesota studies in the philosophy of science* (Minneapolis, MN: University of Minnesota Press).

CLAXTON, G. (1984) Teaching and acquiring scientific knowledge. In T. Keen and M. Pope (eds.) *Kelly in the classroom: Educational Applications of Personal Construct Psychology* (Montreal: Cybersystem).

CLEMENT, J., BROWN, D. E. and ZIETSMAN, A. (1989) Not all preconceptions are misconceptions: finding 'anchoring conceptions' for grounding instruction on students' intuitions. *International Journal of Science Education*, 11, 554-565.

CLOUGH, E. and DRIVER, R. (1986) A study of consistency in the use of students' conceptual frameworks across different task contexts. *Science Education*, 70, 473-496.

DRIVER, R. (1973) The representation of conceptual frameworks in young adolescent science students'. (tesis inédita), University of Illinois.

DRIVER, R. (1979) The pupil as a scientific. Paper presented at the GIREP conference, Rehovot, Israel, August.

DRIVER, R. and EASLEY, J. (1978) Pupils and paradigms: a review of the literature related to concept development in adolescent science students. *Studies in Science Education*, 5, 61-84.

DRIVER, R. and OLDHAM, V. (1986) A constructivist approach to curriculum development in science. *Studies in Science Education*, 13, 105-122.

DRIVER, R. and SCANLON, E. (1988) Conceptual change in science. *Journal of Computer Assisted Learning*, 5, 25-35.

EBEL, R. (1972) *Essentials of educational measurement* (New Jersey: Prentice-Hall) 2nd edn.

FINEGOLD, M. and GORSKY, P. (1991) Students' concepts of force as applied to related physical systems: a search for consistency. *International Journal of Science Education*, 13, 97-113.

GALILI, I. (1993) Weight and gravity: teachers' ambiguity and students' confusion about the concepts. *International Journal of Science Education*, 15, 149-162.

GALILI, I. and BAR, V. (1992) Motion implies force: where to expect vestiges of the misconception? *International Journal of Science Education*, 14, 63-81.

GILBERT, J. K. (1995) Studies and fields: directions of research in science education. *Studies in Science Education*, 25, 173-197.

GILBERT, J. K. and OSBORNE, R. J. (1980) I understand, but I don't get it: some problems of learning science. *School Science Review*, 61, 664-678.

GILBERT, J. K., WATTS, D. M. and OSBORNE, R. J. (1982) Eliciting student views using an Interview-about-Instances technique. *Physics Education*, 17, 62-65.

GILBERT, J. K. and SWIFT, D. J. (1985) Towards a lakatosian analysis of the piagetian and alternative conceptions research programs. *Science Education*, 69, 681-696.

GLYNN, G. M. and DUIT, R. (eds) (1995) *Learning science in schools* (Hillsdale, NJ: Erlbaum).

GUNSTONE, R. F. and WHITE, R. T. (1980) *A matter of gravity*. Paper given at the meeting of the Australian Science Education Research Association, Melbourne, May.

GUNSTONE, R. F. and WHITE, R. T. (1981) Understanding of gravity. *Science Education*, 65, 291-299.

GUNSTONE, R. and WATTS, M. (1985) *Force and motion*. In R. Driver, E. Guesne and A. Tiberghien (eds), *Children's ideas in Science* (Milton Keynes, Open University Press).

GUTIÉRREZ, R. and OGBORN, J. (1992) A causal framework for analysing alternative conceptions. *International Journal of Science Education*, 14, 201-220.

HELM, H. (1980) Misconceptions in physics amongst South African students. *Physics Education*, 15, 92-105.

IVOWI, U .M. O. (1984) Misconceptions in physics amongst Nigerian secondary school students. *Physics Education*, 19, 279-285.

JIMÉNEZ GÓMEZ, E., SOLANO, I. and MARÍN, N. (1994) Problemas de terminología en estudios realizados sobre 'lo que el alumno sabe' en Ciensias. *Enseñanza de las Ciencias*, 12, 235-245.

Jiménez Gómez, E., Solano, I. and Marín, N. (1997) Evolución de la progresión de la delimitación de las 'ideas' de alumno sobre fuerza. *Enseñanza de las Ciencias*, 15, 309-328.

Karmiloff-Smith, A. (1994) *Más allá de la modularidad* (Madrid: Alianza).

Kruger, C., Palacio, D. and Summers, M. (1992) Surveys of English primary teachers' conceptions of force, energy and materials. *Science Education*, 76, 339-351.

Kuiper, J. and Mondlane, E. (1994) Student ideas of science concepts: alternative frameworks? *International Journal Science Education*, 16, 279-292.

Lawson, A. E. (1993) Deductive reasoning, brain maduration, and science concept acquisition: are they linked? *Journal of Research in Science Teaching*, 30, 1029-1051.

Lawson, A. E. (1994) Uso de los ciclos de aprendizaje para la enseñanza de destrezas de razonamiento científico and de sistemas conceptuales. *Enseñanza de las Ciencias*. 12, 165-187.

Maloney, D. P. (1984) Rule-governed approaches to physics: Newton's third law. *Physics Education*, 19, 37-42.

Marín, N. (1995) *Metodología para obtener información del alumno de interés didáctico* (Almería: Servicio de Publicaciones de la Universidad de Almería).

Marín, N. (1999) Delimitando el campo de aplicación del cambio conceptual. *Enseñanza de las Ciencias*, 17, 79-92.

Marín, N. and Benarroch, A. (1994) A comparative study of Piagetian and constructivist work on conceptions in science. *International Journal of Science Education*, 16, 1-15.

Marín, N. and Jiménez Gómez, E. (1992) Problemas metodológicos en el tratamiento de las concepciones de los alumno en el contexto de la filosofía e historia de la Ciencia. *Enseñanza de las Ciencias*, 10, 335-339.

Marín, N., Jiménez Gómez, E. and Benarroch, A. (1997) Delimitación de 'lo que el alumno sabe' a partir de objetivos and modelos de enseñanza. *Enseñanza de las Ciencias*, 15, 215-224.

Marín, N., Solano, I. and Jiménez Gómez, E. (1999) Tirando del hilo de la madeja constructivista. *Enseñanza de las Ciencias*, 17, 479-492.

Marín, N., Benarroch, A. and 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.

McDermott, L. (1984) Research in conceptual understanding in mechanics. *Physics Today*, 37, 24-37.

Montanero, M. and Perez, A. L. (1995) A survey of students' understanding of colliding bodies. *Physics Education*, 30, 277-283.

Moreira, M. A. (1994) Diez años de la revista *Enseñanza de las Ciencias*: de una ilusión a una realidad. *Enseñanza de las Ciencias*, 12, 147-153.

Moreno, M., Sastre, G., Bovet, M. and Leal, A. (1993) *Papel de los modelos representacionales en la construcción de conocimientos. Un ejemplo concreto*. IV Congreso Internacional sobre investigación en la Didáctica de las Ciencias and de las Matemáticas, Barcelona, 105-106.

Niaz, M. (1991) Correlates of formal operational reasoning: a neo-piagetian analysis. *Journal of Research in Science Teaching*, 28, 19-40.

Noce, G., Torosantucci, G. and Vicentini, M. (1988) The floating of objects on the moon: prediction from a theory or experimental facts? *International Journal of Science Education*, 10, 61-70.

Nussbaum, J. and Novak, J. D. (1976) An assessment of children's concepts of the earth utilising structured interviews. *Science Education*, 60, 535-550.

Ogborn, J. (1985) Understanding students' understanding: an example from dynamics. *European Journal of Science Education*, 7, 141-150.

Oliva, J. M. (1999) Algunas reflexiones sobre las concepciones alternativas and el cambio conceptual. *Enseñanza de las Ciencias*, 17, 93-108.

Osborne, R. J. and Gilbert, J. K. (1980) A technique for exploring students' views of the world. *Physics Education*, 15, 376-379.

Pascual-Leone, J. (1979) *Le teoría de los operadores constructivos*. In J. Delval (ed.), *Lecturas de psicología del niño* (Madrid: Alianza Universitaria) 208-228.

PIAGET, J. (1977) *Epistemología genética* (Argentina: Solpin) Original (1970) *L'epistemologie génetique* (París: Presses Universitaires de France).

PIAGET, J. (1980) *Biología Conocimiento* (México: Siglo XXI).

PIAGET, J. and GARCÍA, R. (1973) *Las explicaciones causales* (Barcelona: Barral).

POSNER, G. J., STRIKE, K. A., HEWSON, P. W. and GERTZOG, W./A. (1982) Accommodation of a scientific conception: toward a theory of conceptual change. *Science Education*, 66, 211-227.

POZO, J. I.(1989) *Teorías cognitivas del aprendizaje* (Madrid: Morata).

POZO, I. and GÓMEZ CREPO, M. A. (1998) *Aprender and enseñar ciencia* (Madrid: Morata).

REYNOSO, E., FIERRO, E., GERDO TORRES, O., VICENTINI /MISSONI, M. and PÉREZ DE CELIS, J. (1993) The alternative frameworks presented by Mexican students and teachers concerning the free fall of bodies. *International Journal Science Education*, 15, 127-138.

RUGGIERO, S., CARTELLI, A., DUPRE, F. and VICENTINI, M. (1985) Weight, gravity and air pressure: mental representations by Italian middle school pupils. *European Journal of Science Education*, 7, 181-194.

RUMELHART, D. E. and ORTONY, A. (1982) The representation of knowledge in memory. *Infancia y apendizaje*, 20, 115-158.

SOLOMON, J. (1994) The rise and fall of constructivism. *Studies in Science Education*, 23, 1-19.

TERRY, C., JONES, G. and HUNFORD, W. (1985) Children's conceptual understanding of forces and equilibrium. *Physics Education*, 20, 162-185.

TERRY, C. and JONES, G. (1986) Alternative frameworks: Newton's third law and conceptual change. *European Journal of Science Education*, 8, 291-298.

THIJS, G. D. (1992) Evaluation of an introductory course on 'force' considering students' preconceptions. *Science Education*, 76, 155-174.

TWIGGER, D. *et al.* (1991) The conceptual change in Science project. *Journal of Computer Assisted Learning*, 7, 144-155.

TWIGGER, D. *et al.* (1994) The conception of force and motion of students aged between 10 and 15 years: 'an interview study designed to guide instruction'. *International Journal of Science Education*, 16, 215-229.

VIENNOT, L. (1979) *Le raisonnement spontané en dynamique élémentaire* (París: Hermann).

VILLANI, A. and PACCA, J. (1990) Conceptos espontáneos sobre colisiones. *Enseñanza de las Ciencias*, 8, 238-243.

VOSNIADOU, S. (1994) Capturing and modelling the process of conceptual change. *Learning and Instruction*, 41, 45-69.

VOSNIADOU, S. and BREWER, W. F. (1992) Mental models of the earth: a study of conceptual change in childhood. *Cognitive Psychology*, 24, 535-585.

VOSNIADOU, S. and IOANNIDES, C. (1998) From conceptual development to science education: a psychological point of view. *International Journal of Science Education*, 20, 1213-1230.

VOSNIADOU, S. (1999) Conceptual change research: state of the art future directions. In W. Schnotz, S. Vosniadou and M. Carretero (eds), *New perspectives on conceptual change* (Londres: Elsevier) 3-14.

WARREN, J. (1971) Circular motion. *Physics Education*, 6, 74-77.

WATTS, D. M. (1980) *An exploration of students' understanding of the concepts 'force' and 'energy'*. Paper presented at the Conference on Education for Physics Teaching. Trieste.

WATTS, D. M. (1982) Gravity—don't take it for granted. *Physics Education*, 17, 116-121.

WATTS, D. M. (1983) A study of schoolchildren's alternative frameworks of the concept of force. *European Journal of Science Education*, 5, 217-230.

WATTS, D. M. and ZYLBERSZTAJN, A. (1981) A survey of some children's ideas about force. *Physics Education*, 16, 360-365.