Research in Science Education **31:** 499–523, 2001.

Images of Science Linked to Labwork: A Survey of Secondary School and University Students

Marie-Geneviève Séré[1], Manuel Fernandez-Gonzalez[2], Jose A. Gallegos[2],
Francisco Gonzalez-Garcia[2], Esteban De Manuel[2], F. Javier Perales[2]
and John Leach[3]
[1] *Université Paris XI (France)*
[2] *Universidad de Granada (Spain)*
[3] *The University of Leeds (U.K.)*

Abstract

This paper presents findings about the images of science drawn upon in laboratory work, by upper secondary and university students, in academic streams with a science focus. Data were collected through four written questions, administered to a total of 368 students. The questions all required students to comment on laboratory investigations carried out by research scientists or by science students. We show that students' reasoning has an epistemological and an ontological dimension, and that it often differs significantly from accepted perspectives on the nature of science. The issue for teaching appears to be showing students what counts (and what does not count) as appropriate reasoning in actual situations. In other words, explicit teaching about the various relationships that can exist between theory and data would transform labwork towards a more critical process that involves making and justifying decisions.

During their schooling, students acquire images of science and of scientific work. In recent years there have been many studies focused on characterising these images. Several studies have shed light on the images of science actually drawn upon by students, no matter how those images of science were developed. Driver, Leach, Millar and Scott (1996), and Désautels and Larochelle (1998), have made comprehensive reviews of this literature. The findings of studies typically portray students as naïve realists and naïve empiricists. Furthermore, there is some evidence to suggest that many teachers also make naïve statements about the nature of science (e.g., Lederman, 1992), and that teachers' images of science are communicated to students during teaching (e.g., Brickhouse, 1990; Hodson, 1993).

However, the process of drawing out implications for practice from this literature is not straightforward. There is some evidence to suggest that students (e.g., Rowell & Dawson, 1983) and professional scientists (e.g., Samarapungavan, 1992) who make naïve verbal statements about science, demonstrate more sophisticated tacit knowledge about science in action settings. Furthermore, students' reasoning about science is intimately connected to context (e.g., Mortimer, 1995; Leach, Millar, Ryder, & Séré, 2000). Quite often, the diagnostic questions used in paper and pencil surveys to elicit students' and teachers' images of science do not refer to any specific context. We refer to such questions as *decontextualised questions*. For example, Porlan Ariza, Rivero Garcia and Martin del Pozo (1998) used decontextualised questions to discuss possible relationships between teachers' epistemological options on one hand, and their teaching options on the other. Koulaidis and Ogborn (1989) presented teachers with a series of general statements, expressing different philosophical views. In both studies, a predefined epistemological position is attributed to each individual student or teacher. This approach elicits the student's or teacher's espoused images of science. However, the relationship between these espoused positions, and the knowledge drawn upon by students and teachers to inform actions, is open to question.

In this research, we were interested in eliciting the kind of knowledge that science students might draw upon to inform their actions in the setting of a particular classroom activity, namely labwork. Within the European project 'Labwork in Science Education'[1] (LSE) (Séré, 1998), we developed some hypotheses concerning the relationship between science students' images of science and their actions during labwork at the level of upper secondary school, and the beginning of university education (Leach, in press). In order to investigate these hypotheses, we designed questions which require students to make judgements and comments on specific laboratory situations. The design of these questions recognises that the images of science that students develop and use are intimately connected to the scientific activities engaged in through laboratory work: different laboratory activities involve students in making different links between conceptual knowledge and images of scientific activity. The knowledge at stake during labwork is not only conceptual: knowledge about aspects of the situation such as the relationship between ideas and data, the process of measurement and so on is also involved. We were interested in the knowledge students use in making decisions about how much data to collect, which of the data to use, and what can be concluded from the data available. In the physical sciences, it is often necessary for students to make decisions about the validity of knowledge claims from experimental work, given the accuracy and precision of measurements. In the life sciences, students often have to make decisions about design and sampling in order to collect useful data. Data analysis typically involves dealing with large data sets exhibiting a high degree of variability. We were interested in the representations of the functioning of science that underpinned students' decisions during labwork.

In the philosophy of science, questions about the nature and status of scientific knowledge are conventionally discussed around two dimensions:

1. an ontological dimension, which addresses the relationship between scientific models and their empirical referents; and
2. an epistemological dimension, which addresses the warranting of knowledge claims as reliable.

These dimensions have been used in the literature to describe conceptual learning (e.g., Chi, 1992; Tyson, Venville, Harrison, & Treagust, 1997; Vosniadou, 1994). In the study reported in this paper, however, we do not address conceptual learning. Rather, we draw upon ontological and epistemological dimensions to describe the knowledge that students appear to draw upon when dealing with decisions encountered during learning through labwork.

Our initial assumption is that students' reasoning during laboratory work has ontological and epistemological dimensions, because during labwork students have to draw upon implicit knowledge about the nature of scientific research, the analysis of empirical data and the validity and reliability of knowledge claims.

The purpose of this paper is to demonstrate, through the analysis of a questionnaire, the epistemological and ontological dimensions that characterises upper secondary and beginning university students' reasoning about labwork. We show that individual students' reasoning exhibits different characteristics in response to different questions. We go on to consider the implications of our findings about students' reasoning for the practice of teaching science through labwork.

The Survey and the Samples

In order to characterise ontological and epistemological aspects of students' reasoning about labwork, it was necessary to design diagnostic questions which provided students with the opportunity to discuss the practice of labwork. Following piloting, four questions (containing both open and closed-response sections) were administered to a sample of students in upper secondary schools, and at the beginning of their science studies at university.

The survey was administered in France and Spain, notably due to the linguistic competence of the research team. It was administered in either French or Spanish respectively. There were some differences between the forms of questions in the French and Spanish questionnaires; these will be commented upon as relevant to the reporting of findings. In Spain, the questionnaire was given to 89 secondary school students and to 37 university students (mainly in Physics or Chemistry, less in Biology). In France it was given to 76 secondary school students and 166 university students, 60 of whom study Biology or Geology and 106, Physics or Chemistry. This gives a total sample of 368 students (40% female and 60% male). This sample involved randomly selected students in academic streams, in several towns, whose studies focused on science. We do not claim that the sample is nationally representative, though we do not see any specific reasons why the sample differs from the national position in either country. Furthermore, given the similarity of our findings to studies using similar methods and involving a wider range of countries (including an anglophone country) (Leach, Millar, Ryder, & Séré, 2000), we suspect that findings from this study might well be of interest in countries other than France and Spain.

Areas of Focus within the Questions

In this section, we explain the focus of each diagnostic question. The full text of each question, translated into English, can be found in the Appendix.

The Theory Question

This is a decontextualised question, asking whether or not it is necessary to have a theory in mind in order to interpret data. This tells us if students believe that the possession of sufficient data allows people to arrive at a theory, or alternatively, that having a theory in mind is necessary to orient the collection of data. To balance the absence of context, it asks students to describe a situation that exemplifies their answer. This question is relevant to the *nature of empirical enquiry*.

The Curves Question

This question involves two scientists who interpret the same data (a set of points) in radically different ways. One of the scientists represents the relationship between data points by means of a straight line, while the other interprets the relationship as a curve. The open-response questions ask students whether it is legitimate to have two interpretations, and how scientists might justify their respective viewpoints. A secondary objective of this question is to discover the extent to which students believe that computers are capable of solving such problems. Since this question involves modelling, it is informative about students' views of the *nature of empirical enquiry* as well as the *analysis of data*.

The Surprise Question

This question describes a context in which the results of an experiment in a classroom situation do not match canonical knowledge. The teacher, who is aware of the theoretical concepts to be taught, proposes a laboratory experiment that is expected to verify them in some way. The example chosen is a test for starch in plant leaves, which should come out positive for leaves left in the light and negative for leaves kept in the dark. However, when the experiment is carried out, some of the students do not obtain the desired result.

In the French questionnaire, this question is in an open-response format and simply asks what should be done if something like this happens in class. In the Spanish questionnaire, the question is formulated as multiple-choice.

The question elicits information about the links that students make between *theory* and *experimental design*. It also concerns the *analysis of experimental data*.

The Series Question

To discover the students' views about handling measured data, we asked questions about sets of spread measurements. Previous studies (Allie, Buffler, Kaunda, Campbell, & Lubben, 1998; Journeaux & Séré, 1994; Lubben & Millar, 1996) suggest that the way that students treat repeat measurements can be used to characterise their implicit assumptions about the nature of data. Different sets of five measurements (of mass) were presented, together with the mean of each set. The sets had the following characteristics, selected on the basis of findings reported in the literature cited above:

1. The mean may or may not appear as one of the values of the series;
2. Certain values appear up to three times in the same series;
3. The range between the smallest and largest value varies in size.

The question concerns two groups of nutritionists who must measure the mass of one decilitre of oil five times. Students were asked about different situations:

1. the same kind of oil is measured by each group, and each group's sets of measurements have the same mean;
2. the groups have the same oil but their sets of measurements have different means;
3. the groups have different oils and their sets of measurements have different means.

The text of the question suggests that a direct measurement of mass was made. In fact, the values obtained depend on the volume, one decilitre, of the sample taken. Strictly speaking, this signifies that the measurement is not direct, and the design of the experiment thus influences the result.

This question provides information principally about students' views of the *analysis of experimental data*, but also about their views of the *nature of empirical enquiry*.

Categories of Analysis

The categories of analysis[2] of the responses were elaborated by drawing upon our hypotheses, our direct observations of students involved in labwork, and our knowledge of the existing literature on students' images of science.

A: Categories Reflecting Ontological and Epistemological Positions

i) Each Measurement should be considered in itself

This idea that the only values for a quantity that are considered valid are those actually obtained as measured values, has been categorised as *Values*. Similarly, in certain cases when students have to give a unique result from two series of measurements, they keep the value that appears in both series. The same type of reasoning is found for series of measurements and for pairs of values (points to be interpreted as a curve). It places undue prominence on the actual measured value, giving no consideration to issues of probability. This attitude, if systematic, contrasts with the view that considers all measures as bearing information to which statistical analysis can be applied. It has an influence on the choice of a result.

ii) Accuracy or uncertainty should be considered for each measurement, in order to judge each of them

This idea has been categorised either as *Uncertainty*, or as *Deviation*: Uncertainty when quantitative values are in play; or Deviation which corresponds to the same epistemological-ontological position, but generally expressed for non-quantitative results. The latter means, for instance, that there is always a difference between what is really obtained from observation, and the event or value deduced from the theory.

Typically, the decision following this attitude, is frequently to select which measurement ought to be kept and which should be discarded. This attitude contrasts with a statistical view of the treatment of different measured values.

iii) The difference between the biggest and the smallest measurements determines the quality of the data

This category is called *Range*.

iv) It is necessary to take several measurements (or at least more than the number given in the question) in order to arrive at valid results

This idea has been categorised as *N Measurements* (in case of quantitative results) and *Start Again* (in the case of qualitative results: it is necessary to repeat an experiment because one single experiment can never be enough).

v) Statistical analysis is necessary in the processing of data

This attitude is categorised as *Statistics*. It is the exact opposite to *Values*.

vi) There exists one true value, which is knowable and must be known before judgements can be made

This idea is categorised either as *Truth* (belief in the existence of an ideal or theoretical value) or *Official Value* (the idea of an officially recognised value is used, with no reference to how that value might be warranted).

All these categories have ontological/epistemological dimensions, as illustrated in Figure 1. The axes in Figure 1 represent an ontological dimension (identity of the real world with its scientific modelling/distance of the real world from its scientific modelling) and an epistemological dimension (primacy of data/primacy of theory). On the extremities of the axis *primacy of theory/primacy of data*, two categories of the decontextualised closed question appear: *theory first/data first*.

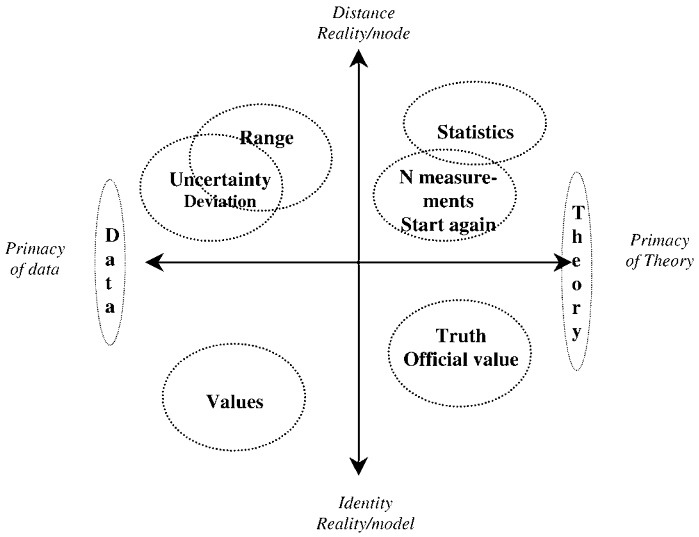


Figure 1: A two-dimensional representation of the categories of analysis of students' answers to questions concerning labwork.

B: Categories Compatible with Any Ontological/Epistemological Positions

i) Existence/inevitability of human error

This is classified in the category *Error*. The terms used are that measurements are imperfect, possibly even 'never perfect,' or 'always with a margin of error.'

ii) The results obtained are explained in terms of the experimental protocol used

This is categorised as *Protocol*.

C: Category of Relativism

Some answers express *Relativism*. All responses stating that nothing can be decided, that subjectivity is in play, or that any conclusion/decision could be equally valid were treated as relativist. Examples of such responses were found on each question in the survey.

Results

We now present the results of the survey, using the categories described in the last section. This means that the majority of responses for each question can be situated in a given quadrant of Figure 1. As will be seen, the answers to the first question we analyse can be interpreted only in terms of underlying epistemology. They are situated on the epistemological axis of Figure 1.

The Theory Question (Focus: the Nature of Research)

The first part of the question is not contextualised. As shown in the appendix, the students are asked to agree or disagree with the 'theory first' statement or the 'data first' statement. They are also asked to give the example they have in mind when making their choice.

In other words, students are asked directly to situate themselves within the plane represented in Figure 1, more precisely on the horizontal axis. The question offers students the choice between two answers that at first seem contradictory. This is due to our assumption that scientists carry out their work in different ways, and the diverse approaches used cannot adequately be described by either of the two statements.

Quantitative results for the whole sample are shown in Table 1:

Table 1
Frequencies of Choices (%) Between General Sentences Describing the Development of Science.

	Agree with statement (%)		Disagree with statement (%)		Not sure (%)	
	Upper secondary	University	Upper secondary	University	Upper secondary	University
Theory first statement	49	39	30	19	20	19
Data first statement	22	11	59	72	12	13

The most frequent response is *Agree with the theory first statement* and *Disagree with the data first statement*. This most frequent response corresponds to the right-hand side of Figure 1. A significant number of students, particularly at the university level, made no response to the question.

By and large, the examples offered by students pertain to physics (almost half of the examples) or biology (approximately a third of the examples). They generally concern historical discoveries or are derived from classroom situations. The following examples were commonly used to justify the *Theory first* position: Biologists who study genetics have been guided by the theory of Mendel and the discovery of DNA; and, a scientist should possess intuition, Pluto (the planet) was discovered theoretically.

Scientists such as Newton, Huygens, Einstein, Darwin, Charpak, and Michelson are also cited as examples to justify this view.

An example cited to justify the *Data first* statement was: "If scientists only started from theoretical knowledge, they would still think that the world was flat."

Students cited Einstein, Le Verrier, Descartes and Bergson to justify statement 2.

About 15% of the students agreed with both statements. The most frequent explanation given is that both are possible because one statement is not sufficient in itself to cover all scientific practice, for example: "The discovery of the bacillus of Koch, as well as the genetic engineering, which searches for one gene or another, show that both methods (i.e. Theory first and Data first) are complementary."

Bohr, Planck, Einstein, Koch, Galileo, Marie Curie, and Lavoisier are cited as examples that prove that both statements are possible, while Planck and Lavoisier were also given as evidence to the contrary. This suggests that a significant proportion of the responses were located on both sides of the horizontal axis of Figure 1, depending on the situation.

In summary, students' responses were spread along the horizontal axis of Figure 1, with a predominance to the right-hand side.

From a methodological point of view, these results mean that one must be careful in the interpretation of decontextualised questions. The answers are heavily dependent on previously learned and memorised examples, and it is impossible to separate the results from their contexts.

The Curves Question (Focus: The Nature of Research and the Processing of Data)

This question asks students what scientists should do to resolve the dilemma when two different interpretations of the same data are available. The question is not contextualised to a specific discipline or situation, and simply offers a set of data points.

Approximately 16% of the students (and as many as 31% of the Spanish secondary school students) stated that the difference is basically due to subjectivity on the part of scientists. Those relativist answers cannot be positioned on Figure 1.

The results for the remaining responses are presented in Table 2:

Table 2

Frequency of Categorised Answers (%) to an Open Question Concerning the Interpretation of Data.

	Categorised answers (%)	
	Upper secondary school	University
VALUES: Students consider each of the data points, its distance from the straight line/curve and its location above or below.	14	10
THEORY FIRST (the word 'model' appears frequently): Students think that the interpretation has been made, and the straight line/curve drawn, on the basis of a model that the scientist has in mind.	36	36
AMBIGUITY (mix of N MEASUREMENTS and UNCERTAINTY): students consider the data points either to be confusing or of insufficient number to make a decision between the two interpretations.	13	23

From these results, it appears that one observes little or no change from secondary school to university.[3] The percentage of *Theory First* responses suggests that over a third of the students are aware of the fact that data interpretation is constrained by a model that the researcher either already has in mind, or must build up. These responses are clearly on the right of the plane of Figure 1, which converge with the results of the Theory question. However, more than 10% of responses, categorised as *Values*, are in the left lower quadrant of Figure 1, for example "The best curve passes through the most points," and "There should be the same number of points above and beneath."

Concerning the discussion between the scientists A and B, 10% of the students thought that no agreement is possible. 3% of the students thought that error bars could be used to reach a consensus. The category *Values* is prevalent at both secondary school and university level (around 19%), while the category *Theory First* has a frequency of 30%. Very few believe that it is necessary to question the experimental protocol, as is evident in the following example, "They (A and B) thought it was enough to examine the experiment and the data, and to try and discover what type of phenomenon it was."

61% of the upper secondary student sample responded that computers can bring closure to the disagreement between the scientists. Such responses may be situated in the lower part of Figure 1 (Identity of real world and model).

The Surprise Question (Focus: The Analysis of Data and Experimental Design)

This question presents a situation where school labwork fails to verify an agreed theory. A possible reason for such deviations is the fact that any modelling always entails simplification, which naturally means that certain phenomena do not fall within the boundary conditions of the model and are not taken into account. Another possibility is that the deviations occur because of random events. This question was administered in an open response format in France, and a closed response format in Spain. Responses to the closed response format are not reported here, as they were difficult to interpret.

Table 3 shows that the responses of students at secondary school and university level are similar. Responses did not generally refer to statistical techniques.

Table 3
Frequency of Responses (%) to the Surprise Question.

	Upper secondary (%)	University (%)
Start again	28	26
Hypotheses	22	17
Deviation	12	12
Statistics	1	2
Other test	2	1

As shown in Table 3, 12% of the responses were classified in the category *Deviation*. Such responses suggest that theory always deviates from reality. This opinion was particularly frequent amongst biology students.[4]

The category *Start Again*, involved a reference to human errors. The category *Statistics* was used rarely in responses to this question.

The responses also allude to the pedagogical problem of the context. 26% of the secondary school students and 33% of the university students affirm that the teacher should give the ‘official value,’ mostly because their primary objective is to prepare students for the course examination by presenting the best theory, for example, “The teacher should repeat the manipulations in front of the students as he explains the results obtained.”

A likely interpretation of this is provided by the statement (chosen by 25% of the Spanish students among a list of statements) of a significant distance between a research laboratory context and the teaching laboratory context.

In summary, it appears that the frequent answers are mainly situated in the two upper quadrants of the plane of Figure 1. The biology students tend towards the left hand side compared to physical scientists.

The Series Question (Focus: The Analysis of Data)

In this question, students were presented with various sets of measured values for the mass of a given volume of oils. Three different situations were presented:

1. Two different sets of measurements of the mass of 5 samples of the same oil, each set having the same mean value.
2. Two sets of measurements of the mass of 5 samples of the same oil, each set having a different mean value.
3. Two sets of measurements of the masses of 5 samples of different oils, each set having a different mean.

Each measurement is represented by decimals with three meaningful digits (except for one value which has four). The mean has four digits and is not the same as any of the values in the set.

i) The analysis of sets of measurements of the same oil: means are the same, means are different

Characteristics of the two sets of measurements A and B are summarised in Table 4:

Table 4
The Main Characteristics of the Sets of Measurements (Same Oil).

	Same oil, same means	Same oil, different means
Range for set A (g)	4 (94.9–98.9)	2 (92.4–94.4)
Range for set B (g)	10.3 (91.9–102.2)	5 (91.9–96.9)
Values repeated in set A (g)	–	–
Values repeated in set B (g)	97.5	–
Values repeated between sets A and B (g)	–	93.3
Interval in which the value has a 95% probability of being (g)	95.4–98.8	93–94.9

Closed-response questions were written to elicit students' views about the quality of measurements. Students had to choose from 4 statements. The frequencies for the whole sample are shown in Tables 5 and 6 (correct statements are emboldened):

Table 5
Percentage of Responses Selecting Preferred Statements About the Quality of Data (Same Oil, Same Mean).

Statement	% selecting as favoured response
(a) The two sets of measurements are equally good because they give the same result.	25
(b) Group A's results are better, because the range between the largest and the smallest measurement is less.	**53**
(c) Group B's results are better, because the measurements cover a wider range of values.	2
(d) From this information, it is not possible to conclude about the quality of measurements.	19

Table 6
Percentage of Responses Selecting Preferred Statements about the Quality of Data (Same Oil, Different Mean).

Statement	% selecting as favoured response
(a) As there are two sets of measurements, it is possible to obtain a value by taking one/some/all of the data from the two.	**29**
(b) Group A's results are better, because the range between the largest and the smallest measurement is less.	**32**
(c) Group B's results are better, because the measurements cover a wider range of values.	5
(d) From this information, it is not possible to conclude about the quality of measurements.	30

As can be seen in Tables 5 and 6, rather a lot of students stated that they could not make a decision (d). Response (c) was not selected by many students, though it is frequently adopted by younger students (Lubben & Millar, 1996). However, responses of the same type are also given by older students when the experimental context is more complex (e.g. manipulation of a pendulum; Larcher, Séré, & Journeaux, 1994).

It is interesting that the frequency of response (b), correct if the quality of measurements is evaluated in terms of their precision (i.e., confidence about the range in which a value lies), decreases in the case where the means are different. Two thirds of the students who answered (b) in the first question, give a different answer in the second. It can thus be surmised that they are referring to something more like the scientific notion of accuracy (i.e., closeness to the 'true' value), rather than something like the scientific notion of precision (i.e., confidence in estimates of the mean).

Consequently, response (b) seems to indicate that the quality is considered as linked to the range, and probably identified to the notion of precision. Fewer students are aware of this link at university level than at secondary level.

The Spanish students that answered (a) were asked to specify the value that ought to be selected. Very few students actually gave an answer, and none of those who did answer gave a completely satisfactory response in that the values given had three meaningful digits when they should only have had two. Apparently the students' calculations were based on common sense or intuition ('mean of means' of the two sets, mode, and so forth). This points to a tendency that has been observed in other contexts. More specifically, students hope to obtain from a set of measurements an interval of confidence that is much smaller than that which is actually obtained.

The types of justification given for these responses fall into the categories that we have previously used for other questions. It was often possible to classify some answers into more than one category. The following results were obtained:

1. When the means are the same, the most frequent response category is *Range* (32% at secondary school level and 24% at university level). Among those students that give this answer, close to half choose the category *Values* when the means are different.
2. More university students in Spain seem to be familiar with statistical analysis (30%) than in France (around 10%). The frequency does not appear to vary between secondary school and university.
3. 20% of the responses belong to the *Values* category (lower left quadrant of Figure 1).
4. 10% of the responses are categorised *Uncertainty*. Some students state that it is necessary to know the degree of uncertainty of each measurement in order to arrive at a valid conclusion. This is also in contrast to the view that a set of measurements provides more information than only one measurement.
5. At secondary school level, about 12% of the students give answers categorised as *Truth* when an authority source is quoted or an *Official Value* or 'A theory', as in the following comment, "All of the values are false. In order to adequately judge, a theoretical value is necessary." These categories are less frequent at the university level.

Students' answers are spread across the quadrants of Figure 1, with a smaller proportion in the right hand sections.

It is also telling to consider the category *Protocol*, which reflects the belief that the quality, and more specifically, the accuracy of the final result is closely linked to the experimental protocol. Around 4% of the secondary school students, and around 13% of the university students, expressed this view.

ii) The analysis of measurements of two different types of oil

The same questions were asked in both the French and Spanish questionnaires, but about different sets of measurements. The main characteristics of the sets of measurements are presented in Table 7.

Table 7
The Main Characteristics of the Sets of Measurements (Different Oils).

	France: different oils	Spain: different oils
Difference between means (g)	0.2	1.2
Range for set A (g)	2.2 (92.2–94.4)	2.2 (92.2–94.4)
Range for set B (g)	1.8 (92.6–94.4)	1.8 (93.6–95.4)
Values repeated in set A (g)	–	–
Values repeated in set B (g)	–	–
Values repeated between sets A and B (g)	92.6; 93.2	94.1
Probability that the two oils are identical (%)	80	8

The closed format question asks if it is possible to identify the two oils on the basis of their density. In both cases the answer 'unsure' was chosen by more than a third of the students (24%–43%, depending on the sample). The frequency of students' responses is shown in Table 8 (correct answers are emboldened).

As can be seen, the most frequent answer is the correct one, in both contexts, though a sizeable minority of students did not answer correctly. Furthermore, the extent to which students' responses were based on a scientific analysis of the situation, as opposed to intuition, is open to question.

The open response to the question regarding the justification of the answer would have been a good opportunity to refer to statistical methods, for instance by using the notions of rate of confidence and of risk. The fact is that none of the responses, either at secondary school or university level in either country, included any such mention. The same categories appear as before, the most frequent being: 1. *Values* (11% to 34% according to the samples), for example: "If a value appears in each of the series, this seems to indicate that it is the same kind of oil with the same density"; and "Either the values of X are smaller than those of Y or vice versa. For this reason,

Table 8
Percentage of Responses Agreeing with Various Conclusions.

	France (means differing by 0.2 g)		Spain (means differing by 1.2 g)	
	Upper secondary school (%)	University (%)	Upper secondary school (%)	University (%)
It is possible to affirm that the measured oils differ in density.	17	18	**55**	**32**
It is possible to affirm that the measured oils do not differ in density.	**49**	**48**	18	19

it is impossible to tell.” 2. Uncertainty (8%–22% according to the samples). 3. As previously, the category *Protocol* is rare (0%–8% according to the samples).

In summary, these results show that: answers interpretable in terms of ontology and epistemology are spread all over the regions of Figure 1, with a predominance in the left-hand side; the importance of the *Values* category, close to the tendency to identify reality and model, depends on the context; and, though difficult to count precisely, tendencies of *Relativism* are expressed occasionally in justifications.

Students are not quick to criticise the experimental procedure itself. They probably are not trained to use this approach.

Synthesis of Results

In the previous section, we showed that students’ responses to the questionnaire could be characterised in terms of epistemological and ontological aspects of their reasoning. We used the four quadrants of Figure 1 to represent these epistemological and ontological aspects. Students did not draw upon a single epistemological/ontological aspect in responding to the questionnaire. The most frequently given answers to different questions were often situated in different quadrants. Generally speaking, more students’ answers were located in the quadrants on the right of the plane, and the quadrants at the top. This means that the tendency *primacy of theory* (on the epistemological axis) and *distance theory/model* (on the ontological axis) were more commonly chosen by students in connection with labwork. However, some responses were to be found in the quadrant characterised by *identity reality-model* and *primacy of data* (though the quantity was generally less than 20%). This leads to posing the following questions: It is possible to situate answers in the quadrants of Figure 1 – is it also possible to situate individuals? In other words, does each individual, at the time he/she answers the questionnaire, draw upon a given position (ontology, epistemology, a tendency to relativism, or another philosophical position)? Does a respondent give answers which all derive from the same position? The method used to address these questions was to carry out a cross-question study[5] of the consistency and inconsistency of responses to different parts of questions, and across questions.

In fact, examples of inconsistency are numerous. The most clear-cut ones are as follows:

1. None of the students answer *Statistics* consistently in the Curves question and the three parts of the Series question, although 49% overall answer this way on one occasion from the four possible occasions.
2. Among six opportunities to answer *Truth,* very few students used this kind of response on three or more occasions.
3. No correlation can be found between the categories *Uncertainty*, *Deviation* and *Range* (each found on the upper left quadrant of the plane).
4. It is possible to use the *Values* category on five occasions in the questionnaire. 36% of the sample used it once, 12% used it twice, very few used it three times.

We tried to identify if some shift from one category to another could be considered as frequent or even systematic. An example is provided by the Series question. Among the students who start with *Range* ("same oil, same mean"), 55% also responded in this way on the "same oil-different means" question. Those who changed (45%) tended to select *Values*. The other examples of frequent types of changes that we noticed led us to the same conclusion: The more complex the situation is, the more frequently *Values* is chosen.

However, certain consistencies exist in responses across questions. This can be seen in the three parts of the Series question. Students remain in the same quadrant, the upper-right one, by answering *N Measurements* and/or *Statistics*. As seen before, the consistency of these students is limited to the Series question and ends as soon as they answer the Curves question. Similarly, when the percentages of the category *Truth* and *Official Value* are added (lower-right quadrant) in the three parts of the Series question, there is considerable consistency in the responses of secondary students. Again the consistency ends when answering the Surprise question. It appears that consistency is mainly the consequence of a similarity in the kind of thing being asked in the question. By contrast, there is no consistency when students are discussing different kinds of labwork (such as that involving quantitative and qualitative data). The same can be said for the category *Protocol*: University students choose it fairly consistently within the three parts of the Series question, drawing upon details of the situation. The consistency is limited to the three parts of the Series question.

From this summary of the cross-question study, it is impossible to conclude that students have unique philosophical positions which account for the way they process data, use theory during an experiment, and consider the respective roles of theory and data in the development of science or personal knowledge. It is impossible to interpret a given student's answers as inspired by the use of a robust mental model. The same persons can possibly change their position on Figure 1 each time a different question is answered, each time they have to judge data or conclude in different situations, for instance in case of passing from qualitative to quantitative results. It means that the shape of the question itself has an influence on students' answers, and that situations themselves are addressed by reasoning with different epistemological and ontological underpinnings.

Conclusion and Perspectives for Teaching

Our survey contributes to literature addressing students' images of science. Our particular contribution is that we have addressed the images of science used by students when talking about the specific situation of labwork. We have proposed a method for classifying students' responses in terms of their underlying epistemology and ontology, and considered whether individuals' responses across a range of questions can be characterised by discrete positions about the respective role of theory and data in experimentation. We present evidence to the contrary, and we give examples of typical ways of thinking concerning labwork, as well as examples of consistencies and inconsistencies.

Some of the consequences are methodological. As it is necessary to take the characteristics of situations into account in the interpretations of responses, it is not valid to characterise ontological or epistemological aspects of reasoning by the use of general, decontextualised questions. Such written questions do have some uses if situations are described. We will suggest below how teachers could fruitfully extend this sort of short questions during labwork.

There are also pedagogical consequences. Though it is not valid to attribute unique philosophical positions to individuals, it cannot be denied that individuals do indeed draw upon philosophical positions in different situations. The results of this survey and our observations in laboratory classes at the same teaching level, converge on this point. These philosophical positions are not stable. Some are appropriate and some are inappropriate. It appears that many students do not learn what counts as appropriate scientific reasoning as a result of current teaching approaches. If this is seen as an important curricular goal for science education, then explicit teaching about the various relationships that can exist between theory and data in different situations is required. This might involve, for example, showing students how theoretical issues influence experimental design in several situations, or showing how the choice of measuring instruments influences findings when a quantity is obtained by means of an indirect measurement. In effect, it is necessary to introduce students to different methods and concepts of data processing.

The relationships between theory and data certainly vary from one discipline to another, between sub-disciplines, and from one situation to another (though not usually from one scientist to another in a given situation). We believe that the study of these relationships in actual situations would have the potential to transform labwork in science education away from a process of the routine application of algorithms, towards a more critical process that involves making and justifying laboratory decisions.

The questionnaire reported in this paper involves a limited number of topics, and the formulation of questions is always open to improvement. However, we believe that teachers could elaborate similar questions with the pedagogical aim of getting students to reflect on the underlying rationale for their laboratory actions. Furthermore, teachers would become more sensitive to the epistemological and ontological commitments that their students were likely to use in specific teaching and learning situations. The questionnaire discussed in this article, as well as the others generated in the LSE project (Leach et al., 1998; Séré, 1998) can be regarded as a databank of teaching resources. They can be adapted to different contexts so as to contribute to establishing the foundations for a method of science teaching which will investigate and provide the means to attain philosophical objectives in real laboratory contexts.

Acknowledgements

We would like to thank Professor Roser Pinto of the Autonomous University of Barcelona (Spain), as well as the members of the DidaScO group of the University of Paris 11 (Orsay, France) for their help in carrying out the survey. We are indebted to the translator, Professor Pamela Faber, from the department of Translation and Interpretation of the University of Granada (Spain). The data set drawn upon in this paper was collected as part of the project Labwork in Science Education, supported by DGXII of the European Commission.

Notes

1. The full title of the project was *Improving Science Education: Issues and research on innovative empirical and computer-based approaches to labwork in Europe*. It was funded in 1996–1998 by the European Commission. It involved research teams from France (Co-ordinator), Denmark, Germany, Great Britain, Greece, Italy, Spain.

2. This sort of categorisation provides results that are very different from an analysis in terms of the use of key words by students, since very different positions may be expressed with the same words (uncertainty, accuracy, precision, etc.).

3. It must be noted that very few respondents (less than ten in all) were so naive as to state that a mathematical expression should always reflect a law. This simplistic view seems not to be learnt from classroom practice.

4. In France, the disciplines of Biology and Geology are taught as one subject in upper secondary schools: *Sciences de la Vie et de la Terre*.

5. To this end, a sample of 250 students has been sufficient to draw conclusions.

Correspondence: Marie-Geneviève Séré, Université Paris XI, DidaScO group,
91405 Orsay, France
E-mail: marie-genevieve.sere@didasco.u-psud.fr

References

Allie, S., Buffler, A., Kaunda, L., Campbell, B., & Lubben, F. (1998). First-year physics students' perceptions of the quality of experimental measurements. *International Journal of Science Education*, *20*, 447–459.

Brickhouse, N. W. (1990). Teachers' beliefs about the nature of science and their relationship to classroom practice. *Journal of Teacher Education*, *41*, 53–62.

Chi, M. T. H. (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* (pp. 129–186). Minneapolis. MN: University of Minnesota Press.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham: Open University Press.

Désautels, J., & Larochelle, M. (1998). *The epistemology of students: The 'thingified' nature of scientific knowledge*. In B. Frazer, & K. Tobin (Eds.), *International handbook of science education* (pp. 115–126). Dordrecht, NL: Kluwer.

Hodson, D. (1993). Philosophic stance of secondary school science teachers, curriculum experiences, and children's understanding of science; some preliminary findings. *Interchange, 24*, 41–52.

Journeaux, R., & Séré, M.-G. (1994). Traitement statistique des incertitudes en Physique: Problèmes scientifiques et didactiques (Statistical treatment of uncertainties in physics: Scientific and didactical problems). *European Journal of Physics, 15*, 286–292.

Koulaidis, V., & Ogborn, J. (1989). Philosophy of Science: An empirical study of a teachers' view. *International Journal of Science Education*, *11*, 173–184.

Larcher, C., Séré, M.-G., & Journeaux, R. (1994). Difficultés dans l'apprentissage du mesurage (Difficulties in learning measurement). *Ensenanza de las Ciencias*, *12*, 27–42.

Leach, J. (in press). The use of secondary data in teaching about data analysis in a first year undergraduate biochemistry course. In D. Psillos, & H. Niedderer (Eds.), *Teaching and learning in the science laboratory.* Dordrecht, NL: Kluwer.

Leach, J., Millar, R., Ryder, J., & Séré, M.-G. (2000). Epistemological understanding in science learning: The consistency of representations across contexts. *Learning and Instruction, 10*, 497–527.

Leach, J., Millar, R., Ryder, J., Séré, M.-G., Hammelev, D., Niedderer, H., & Tselfes, V. (1998). *Survey 2: Students' images of science as they relate to labwork learning*. (Working Paper 4, Labwork in Science Education Project). Leeds: CSSME.

Lederman, N. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*, 331–359.

Lubben, F., & Millar, R. (1996). Chidren's ideas about the reliability of experimental data. *International Journal of Science Education*, *18*, 955–968.

Mortimer, E. F. (1995). Conceptual change or conceptual profile change? *Science and Education, 4*, 267–285.

Porlan Ariza, R., Rivero Garcia, A., & Martin del Pozo, R. (1998). Conocimiento profesional y epistemologia de los profesores, II: Estudios empiricos y conclusiones. *Ensenanza de las Ciencias* (Teachers' professional knowledge and epistemology: Empirical studies and conclusions), *16*, 271–288.

Rowell, J. A., & Dawson, C. J. (1983). Laboratory counter examples and the growth of understanding in science. *International Journal of Science Education, 5*, 203–215.

Samarapungavan, A. (1992, April). *Scientists' conceptions of science: A study of epistemic beliefs.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Séré, M.-G. (1998). *Improving science education: Issues and research on innovative empirical and computer-based approaches to labwork in Europe*. Final report of the project *Labwork in Science Education*, Targeted Socio-Economic Research, Science, Research & Development, European Commission. Brussels, Belgium: The European Commission.

Tyson, L. M., Venville, G. J., Harrison, A. G., & Treagust, D. F. (1997). A multidimensional framework for interpreting conceptual change events in the classroom. *Science Education*, *81*, 387–404.

Vosnadiou, S. (1994). Capturing and modelling the process of conceptual change. *Learning and Instruction, 4*, 45–69.

Appendix: Full Text of the Questionnaire

The Theory Question

The following pair of statements is about how scientists work with data. In each case, please say whether you *agree* with the statement, whether you *disagree* with it, or whether you are *not sure* whether you agree or disagree with it.

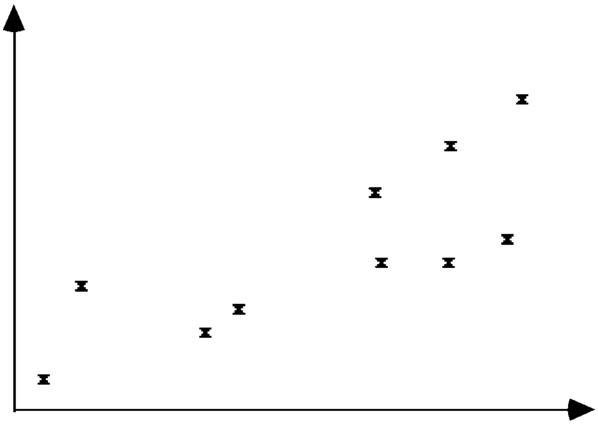
	Agree with statement	Disagree with statement	Not sure
Good scientists ought to have theoretical assumptions which influence their analysis of data.			
Good scientists ought to be neutral and objective: it is not acceptable for scientists' theoretical assumptions to influence their analysis of data.			

Please explain your reasoning: for instance give the example in biology, chemistry, physics etc... that you have in mind when answering.

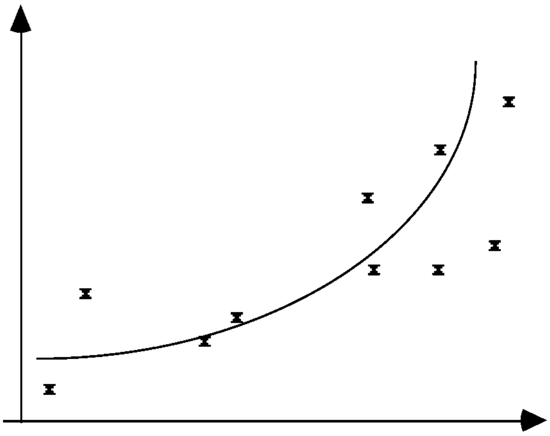
The Curves Question

Two researchers, A and B, have obtained a series of points represented below in the first picture. They give two interpretations shown in the second and third picture.

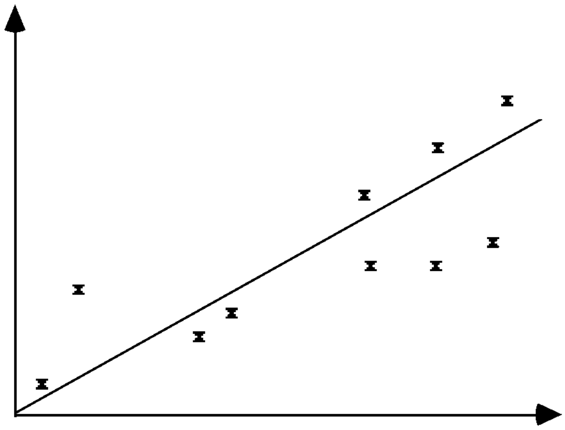
Measurement results:



Interpretation of A:



Interpretation of B:



In your opinion, what could be the reasons for the two different curves?

If the two researchers would meet and discuss, what kind of arguments you would expect them to use?

Which help can be expected from a computer to help them decide?

Would the computer be able to make a decision between the two?

The Surprise Question

A class is carrying out a practical task in the laboratory. They are testing samples of leaves from plants which have been kept for a few days in a dark cupboard, and others which have been in the light. The tests are for the presence of starch in the leaves. The teacher knows that there should be starch in the leaves from plants kept in the light, but none in the leaves of plants kept in the dark. As this is a standard result in all the textbooks, the students will be expected to know it in their examinations. But the students do not yet know what the results should be.

Some groups of students in the class get the results the teacher expects. However, some groups of students do not get a positive starch test for any of their leaves. And some groups get a positive starch test for some of the leaves from plants kept in the light (but not for all of them) and also for some of the leaves from plants kept in the dark (but not for all of them).

What should the teacher do in this situation?

Explain why you think this is the best thing to do:

The Series Question

When the result is different from one measurement to another, how to conclude?

Two groups of nutritionists have been asked to measure the mass of 1decilitre of olive oil. Each group receives one decilitre and makes five measurements. These are their results, after having classified them in increasing order:

Group	Measurements (g)					Mean (g)
A	94.9	96.6	97.1	98.1	98.9	97.12
B	91.9	96.5	97.5	97.5	102.2	97.12

With which of the following statements do you most closely agree?

A	Group A's results are better, because the range between the largest and the smallest measurement is less.	
B	Group B's results are better, because the measurements cover a wider range of values.	
C	Both sets of measurements give the same result	
D	From this information it is not possible to conclude about the quality of measurements.	

Please explain your reasoning:

Another two groups of nutritionists have been asked to measure the mass of 1decilitre of nut oil. Each receives one decilitre and makes five measurements. These are their results:

Group	Measurements (g)					Mean (g)
A	92.4	93.3	93.4	94.0	94.4	93.5
B	91.9	93.3	94.9	95.0	96.9	94.4

With which of the following statements do you most closely agree?

A	Group A's results are better, because the range between the largest and the smallest measurement is less.	
B	Group B's results are better, because the measurements cover a wider range of values.	
C	From the two sets a value can be chosen	
D	From this information, it is not possible to conclude and choose a value	

Please explain your reasoning:

If you chose D, please say what additional information you would need in order to be able to say which group's results are better:

A group of nutritionists wants to compare the masses of two different vegetable oils. They make five measurements, for each decilitre of one sort of oil. These are their results:

	Measurements (g)					Mean (g)
Oil X	92.2	92.6	93.2	94.1	94.4	93.3
Oil Y (France)	92.6	93.2	93.3	94.0	94.4	93.5
Oil Y (Spain)	93.6	94.1	94.3	95.0	95.4	94.5

From these results, the group concludes that "specific mass is greater for olive oil than for nut oil." Do you agree with this conclusion?

Agree	Disagree	Not Sure

Please explain your reasons for agreeing or disagreeing: