STUDENT VIEWS ABOUT SCIENCE
A Comparative Survey

Ibrahim A. Halloun

Professor of physics and education
Science education consultant

A monograph published, in 2001, by

Educational Research Center, Lebanese University, Beirut, Lebanon

Lebanese students' views about knowing and learning science are surveyed and correlated with course achievement and teaching practice. The Views about Sciences Survey (VASS) was revised to this end, and administered to over two thousand Lebanese students enrolled in physics courses of different levels. The revised survey includes a set of dichotomous items and a larger set of items devised following the Contrasting Alternatives Design (CAD). By comparison to traditional assessment formats like open-ended, multiple choice and rating scales, CAD is shown to be significantly more reliable for assessing VASS type of views. Data further show that Lebanese students share with their international peers, particularly those in USA, common distorted views about science, and that these views correlate significantly with course achievement. Student views are not affected by traditional teaching practices, irrespective of how explicitly teachers might explain the nature of science or spell out specific guidelines for studying physics in conventional settings.

April 2001

Note that VASS P204

supersedes VASS P20

and that the two versions

have different profile cutoffs

www.inco.com.lb/halloun

P. O. Box 2882, Jounieh, Lebanon

Table of Contents

Students' views about the nature of science and their attitudes towards it have drawn a particular attention within the educational community in the second half of the twentieth century, and especially in the last two decades. Two major reasons were behind this concern. First, many researchers had suspected, and found, that student achievement in science courses may be affected by these views and attitudes (Aikenhead, 1987; Baker & Piburn, 1991; Cobern, 1993; Edmonson & Novak, 1993; Halloun & Hestenes, 1998; Lederman, 1992; Meichtry, 1992, 1993; Redish & Saul, 1995; Reif & Larkin 1991; Schibeci & Riley, 1986; Songer & Linn, 1991). Second, science and technology have become the main driving forces of development in our globe. Governments and concerned educational bodies around the world have consequently been working to ensure scientific literacy for all citizens, with an emphasis on the interplay between subject-matter on the one hand, and learning styles, beliefs and attitudes toward science and technology on the other (AAAS, 1990, 1993; NCEE, 1983; NRC, 1996; NSTA, 1995; UNESCO, 1993).

Despite all calls for scientific literacy and efforts put by teachers and concerned educators for improving students' ideas about knowing and learning science, research around the world has been constantly showing the following:

- Most people have unsatisfactory understanding of the most elementary conceptions in science and mathematics, irrespective of their educational background (Halloun, 1993; NSB, 1996; TIMSS, 1994).
- Students of all levels are encumbered with folk views about the nature of science and science education that are at odds with the views of scientists and educators (first list of references above).
- Introductory science courses do little to change student views and, more often than not, when changes occur, students' views shift closer to folk views than to views commonly accepted within the scientific community (*ibid.*).
- Students' achievement in science courses may be negatively affected by their folk views (*ibid.*).

With the reconstruction of its infrastructure and its commitment to long-term developmental projects in all sectors, Lebanon is in particular need for expertise in science and technology (S&T) and for public awareness of the role of these fields in development. The new pre-college curricula that took effect in the 98-99 academic year put an emphasis on some sort of S&T literacy and acknowledge that "it is important that students become lifelong learners of science", that they "understand the nature of science and technology", "abide by…scientific values" and develop positive attitudes toward S&T (Lebanese Government Decree 10227, 1997, p. 458).

Lebanon has participated so far in no international S&T survey. Except for research conducted by this author and bearing on the first point (Halloun, 1986, 1988, and 1993), no national survey has yet been instituted to assess regularly where Lebanese students stand on any of the issues mentioned in the four points above. In an attempt to prepare the way in this direction, this author has taken the initiative of assessing these students' views about knowing and learning science, using updated forms of VASS, the *Views About Science Survey*.

VASS is a paper-and-pencil instrument designed to survey student views in question (Halloun, 1996, 1997; Halloun & Hestenes, 1998). Its items are written following the *Contrasting Alternatives Design* (CAD), a novel assessment format devised by this author (Halloun, 1996, 2001a). VASS and CAD were revised recently, and the instrument was subsequently administered, in the span of over one year, to more than two thousand Lebanese students enrolled in secondary school and university physics courses nationwide. Five major questions were addressed in the process:

1. How can VASS be made accessible to Lebanese students (or students of any nationality to that matter), and easy for educators to administer and interpret?
2. What are the advantages of the Contrasting Alternatives Design (CAD) used in VASS over other assessment formats, especially Likert rating scale that is most commonly used in major instruments assessing student views about science?
3. How do Lebanese students' views about science compare to those of their international peers especially in USA?
4. How do student views about science relate to course achievement?
5. How does teacher practice affect students' views?

The answers to all these questions follow in this monograph. The monograph takes the form of technical report that would benefit all those who would like to use VASS and assist them in the interpretation of their data. In order to put things into perspective, let us start our report with a brief review of how VASS and CAD originally came about.

VASS and CAD

1. VASS history

The Views about Science Survey (VASS) was originally developed by the author, in collaboration with the Modeling research team at Arizona State University. The instrument development started in 1993. Over three academic years, it has evolved gradually from an open-ended questionnaire to a closed type with *Contrasting Alternatives Design* (CAD, §2), and from an instrument with a single form for physics education to one with four forms, for physics, chemistry, biology and mathematics education respectively (Figure 1). At an intermediary stage (Spring 94 through Spring 95 in Figure 1), two complementary forms of VASS were developed for each discipline, one for students' learning styles and their attitudes toward science education, and another for their views about the nature of science. The two forms were subsequently refined and merged in a single one. By the spring semester of 1996, the instrument was administered to over ten thousand US high school and university students, and validated for surveying student views about knowing and learning science (and mathematics) and assessing the impact of these views on course achievement (§3). Since then, the instrument has been widely used in many countries around the world for evaluating science or mathematics instruction and related reform projects.

Figure 1

1.1 Objectives

VASS was designed to meet the following objectives:

1. To ascertain significant *differences* between the views of students, teachers and

scientists.

2. To identify patterns in student views and classify them in general *profiles*.
3. To measure the *effectiveness of instruction* in *changing* student views and profiles.
4. To compare student views/profiles at various grade levels (8-16).
5. To assess the relation between student views/profiles and achievement.
6. To compare student views/profiles in different science courses (physics, chemistry, biology...) and across various demographic strata.

1.2 Taxonomy

To identify major issues to be addressed in VASS, we reviewed related works in the relevant literature, including the following:

1. Scholarly views on the philosophy of science (e.g., Bernard, 1865; Bunge, 1973; Giere, 1988; Harré, 1959; Johnson-Laird, 1983; Kuhn, 1970; Lakatos & Musgrave, 1974; Lakoff, 1986; Lecourt, 1974; Popper, 1983; Tobin, 1993; Ullmo, 1969).
2. Major works in cognition (e.g., Changeux, 1983; Dewey, 1933; Ericsson & Charness, 1994; Gardner, 1985; Gilovich, 1991; Ginsburg & Opper, 1979; Glass, Holyoak & Santa, 1979; Grossberg, 1982; Jones, 1986; Klahr, 1976; Lochhead & Clement, 1979; Margolis, 1987; Newell & Simon, 1972; Perry, 1970; Piaget, 1972; Resnick, 1989; Simon, 1979; Squire, 1986; Stewart, 1985; Tobin, 1993)
3. US national science standards (AAAS, 1990 & 1993; NCEE, 1983; NRC, 1996; NSTA, 1993 & 1995).
4. Research on student views about science (e.g., Aikenhead, 1987, 1988; Aikenhead, Fleming & Ryan, 1987; Baker & Piburn, 1991; Barrington & Hendricks, 1988; Cobern, 1993; Cooley & Klopfer, 1961; Ebenezer & Zoller, 1993; Edmonson & Novak, 1993; Fleming, 1987; Germann, 1988; Gilbert, 1991; Kimball, 1968; Klopfer, 1969; Lederman, 1992; Lederman & O'Malley, 1990; Mackay, 1971; Meichtry, 1992 & 1993; Redish & Saul, 1995; Roth & Rychoudhury, 1993; Rubba & Andersen, 1978; Ryan, 1987; Schibeci & Riley, 1986; Schmidt, 1967; Simpson & Oliver, 1985; Songer & Linn, 1991; Welch & Pella, 1967).

In constructing a taxonomy of the issues we identified, we sought to avoid: (a) arcane and problematic questions about the epistemology of science, and (b) bias toward our own position (Halloun, 2001b, 2000; Hestenes, 1992). We devised one VASS instrument after another to assess student views on the targeted issues, and we kept refining our taxonomy and VASS items based on:

1. Peer review done by university professors and experienced high school teachers who are versed in science education literature (Halloun, 1996).
2. Students' answers on VASS items, and their relation to course achievement and performance on content-based conceptual surveys like the Force Concept Inventory (Halloun & Hestenes, 1985a, 1985b; Hestenes et al., 1992).
3. Interviews with respondents.

We finally settled on three scientific dimensions and three cognitive dimensions. The *scientific* (or, actually, metaphysical) dimensions pertain to the structure and validity of scientific knowledge, and to scientific methodology. The *cognitive* (or pedagogical) dimensions pertain to learnability of science, reflective thinking, and personal relevance of science. To assess variability in student views in different disciplines, we constructed parallel forms of VASS along these dimensions for physics, chemistry, biology, and mathematics. This report is strictly devoted to the discussion of the physics form.

Each of the six dimensions is framed in Table I in the form of pairs of contrasting views about science or science education that our analysis revealed to be the most prevalent. The primary view, hereafter referred to as the *expert* view, is the one we found to be most common among philosophers of science, scientists and educators. The opposing view, hereafter referred to as the *folk* view, is often held by the lay community and science students at all grade levels.

Table I

The taxonomy shown in Table I refers to VASS–Form P20 (§1.3) and is an update of the taxonomy from which ensued Form P12 that was made available in the last phase of Figure 1 (Halloun, 1997; Halloun & Hestenes, 1998). Five new items have subsequently been added in the new form to accommodate the changes. These are items 4, 5, 6, 12, and 21 (cf. appendix).

Aside from the diagnosis of student views along the six dimensions of VASS (Table I), one objective in the current project was to assess the relationship between these views and course achievement. Given the difficulties we expected –and actually encountered in Lebanon– in getting participants' grades in their physics courses, we had to find a reliable alternative. To this end, we added in VASS Form P20 a *course achievement* dimension consisting of three new questions asking students to rate their standing in their physics classes (questions 1, 2, 3). This new dimension does not figure in Table I because it does not actually reflect student views about science, which VASS is all about. It serves a different purpose discussed in section 9 of this report: to provide us with a reliable indicator of actual course performance every time we are unable to get formal course grades.

1.3 Item format

The latest form of VASS developed to meet physics education needs in USA was Form P12. This form consisted of 31 items, all designed on an 8-point CAD scale in the manner shown in Figure 2. In a pilot study conducted in the spring of 1997, VASS–Form P12 was administered to a sample of secondary school and university students in Lebanon. Data analysis, including US data, revealed that the form needs to be revised to improve its efficiency in giving a valid and reliable picture of the state of Lebanese students, or students of any nationality to that matter, regarding the original VASS dimensions. Hence, and during the fall of 1998 and early spring of 1999, VASS Form P12 was revised in collaboration with Lebanese teachers, and VASS Form P99LB came about, both in English and French. This new form is an upgrade of Form P12 with respect to taxonomy (§1.2), item format and wording, and it is supplemented with a teacher survey for assessing the impact of teaching practice on student views about science. The new form was administered in 1999 to about two thousand Lebanese students (§4), and slightly revised subsequently to take the shape of Form P20 provided in the appendix. Form P20 is distinguished from Form P12 (and P99LB) in the following respects:

1. Up to Form P12, we allowed participants to chose neither of the two contrasting alternatives provided with each question (Alternative 8 in Figure 2). Except for two

items, no more than 3% of participating students had ever chosen alternative 8 on any given item. We thus dropped out alternative 8 altogether after revising the wording of the two exception items (No. 13 and 15 in Form P20), and adding item 12 to consolidate the changes made heretofore.

Figure 2
----

2. Lebanese students expressed some unease with the 7-point scale (excluding alternative 8). This was due to these students' unfamiliarity with rating scales, and, most importantly, to the inability they expressed during the interviews conducted during the pilot study to make the expected distinction between choices 2 and 3 or choices 5 and 6 in Figure 2. Moreover, analysis of US data revealed insignificant differences between the proportions of students choosing answers 2 and 3 (or 5 and 6) in virtually every item. We thus combined each of the two couples of choices in one choice (actually dropping out choices 2 and 6 in Figure 2), and we ended up with a 5-point scale CAD format. In Form P99LB, this 5-point scale included choices 1, 3, 4, 5 and 7 in Figure 2.
3. Analysis of US data and Lebanese data coming from the pilot study revealed that student answers on some questions in Form P12 were practically polarized toward one alternative or the other, i.e. toward one end or the other of the 7-point scale. These items were thus behaving more like dichotomous than rating-scale items. We thus converted them into the former format in Form P99LB, and we asked students to chose exclusively one or the other of the two alternatives provided with each of these items instead of expressing their position on a 5-point CAD scale. The items in question are items 6 through 15 in section II (Form P20 in the appendix).

Analysis of data pertaining to form P99LB and subsequent peer review implied some minor revisions in the new VASS form, mainly with regard to the CAD format. The wording of some items have been slightly revised in the subsequent form P20, but the list of items and their taxonomy remained the same as in form P99LB LB. In Form P20, CAD alternatives have been modified so that the polar ends (Only "a" or Only "b") could better accommodate students who would like to account for choices 2 and 6 in Figure 2 (Mostly "a", Rarely "b", or the opposite), especially US and international students who are used to rating scales. The modification was made possible in the direction shown in Form P20 (appendix) especially because items that were actually bipolar in the original P12 form have already been transformed into dichotomous items in section II of the new form, and the remaining items were no longer truly bipolar. All these changes in Form P20 did not affect students' answers on the various items; answer distribution remained practically the same as in Form P99LB (§3).

2. Contrasting Alternatives Design

Traditional assessment instruments present items in one of two formats: (a) open–ended (constructed, or free–response), or (b) closed (objective, or selected–response). Open formats like interviews and essays can be valuable and informative means of assessment for purposes like ours. However, they are not feasible for large samples. Objective formats like multiple-choice and Likert scale are more practical and cost-efficient. However, research indicates that they encounter insuperable validity and reliability problems when used in surveying students' views about science (Halloun, 1994, 2001a; Krynowsky, 1988;

McComas, 1998; Munby, 1983; Rennie & Parker, 1987; Symington & Spurling, 1990).

For VASS, we needed a valid and reliable testing format that could be used to survey large samples efficiently. Since no traditional format meets all three criteria: validity, reliability and feasibility, the author of this report devised a new item format, the *Contrasting Alternatives Design* (CAD), which requires respondents to balance between two contrasting alternatives in the manner shown in Figure 2. Each item consists of a statement followed by two alternatives which respondents were originally asked to balance on an eight-point scale (now, on a 5-point scale). They could pick either alternative exclusively (options 1 or 7), a weighted combination of the two (options 2, 3, 4, 5, or 6), or neither one (option 8, now excluded).

CAD advantages and other features of VASS are discussed elsewhere in ample details (Halloun, 2001a; Halloun, Hestenes & Osborn Popp, 2001). In this section, we compare CAD to the three most popular assessment formats, open-ended, multiple choice and Likert scale, with an emphasis on the latter. We pay a special attention to comparing CAD to the Likert rating scale in this report because it is most commonly used in other instruments assessing student views about science.

Before we compare CAD to other survey formats, let us first point out some precautions that we took originally in designing VASS and that actually set this instrument further apart from all other instruments so far developed for assessing student views about science.

In the wording of each item (stem and alternatives), we tried, as much as we could, to use colloquial terms and avoid the use of formal scientific terminology that students may not be familiar with. Furthermore, and as much as we avoided arcane issues in our taxonomy, we addressed each included issue in a familiar context, knowing that people, and especially secondary school students, have difficulty thinking about any issue in the abstract.

In the latter respect, and contrary to common practice in the design of traditional instruments for assessing student views about science, VASS: (a) asks questions about specific disciplines, (b) narrows issues in a given question down to a single factor in a given dimension, and (c) is restricted to issues that are within the scope of target populations (Halloun, 1996; 2001a).

As a rule, surveys that we have examined ask questions about "science" in general. We suspected that student opinions would differ according to discipline, so we designed different VASS forms for different disciplines (biology, chemistry, physics, and mathematics, so far), in accordance with the same taxonomy. Our suspicion has been actually confirmed by the results we kept getting in the last seven years.

Student views about science vary not only between disciplines, but also within a given discipline. Student epistemological views often vary from one theory to another within the same science or even from one law to another within the same theory. Where appropriate, VASS accounts for students' sensitivity to content by asking the same question in more than one context within the same science. Figure 3 compares our approach to that of others in this regard. As we shall see in section 5 of this report, student answers varied significantly on the two VASS items shown in this figure (and others), which supports our position to frame each item in the context of a single familiar situation.

Figure 3

Traditional instruments often address several factors in a single question (Figure 4). In

VASS, each question concentrates on a single factor within a given dimension as can be seen in the VASS forms presented in the appendix and in the list of items associated with each issue in the taxonomy of Table I. Moreover, these instruments often address issues that are beyond students' purview and experience. The test called TOUS in Figure 4 has been administered to fifth graders! Notwithstanding apparent misconceptions that TOUS authors have with the nature of scientific theory, our research has shown that such ambiguous and debatable questions have little utility, and are thus avoided in VASS.

Figure 4

2.1 Open-ended questions and CAD

Open-ended questions are often open to a wide variety of interpretations by respondents as well as by researchers. They can be misleading, especially when respondents' priorities or value judgments are not the same as those of the concerned researcher, which is often the case. When VASS was first administered in open-ended format (Figure 1), students were asked in one of the questions to state the first thing they do in solving a physics problem. The student in Figure 5 replied by writing that he starts by looking for the appropriate *formula*. When interviewed, it became evident that the first thing this student actually does in solving a physics problem is *draw diagrams*, but this procedure seemed so trivial to him that he thought it was not worth mentioning in his written response (Halloun, 1996, 2001a).

Figure 5

The risk of interpretation mismatch between surveyor and respondent that we run with open-ended questions (as well as with those in other traditional formats) is resolved in CAD where the surveyor's position can be easily and explicitly contrasted with a popular position distinguished from his. This latter position should of course come from open-ended questionnaires coupled with interviews to clear up any possible mismatch in the manner described in Figure 5. This is actually what we have done with VASS (Figure 1), and this is how we ended, for example, asking question 28 (cf. appendix) in the manner we did to resolve the issue emanating from the protocol of Figure 5.

2.2 Multiple-choice questions and CAD

Multiple-choice questions could be valid and reliable means of assessment when they are adequately constructed, and when only one answer could be exclusively acceptable to a given question. This format defeats its purpose in the event a respondent feels that s/he could pick more than one of the provided alternatives, whether or not this is acceptable by the surveyor, and especially when the choice can be of ranked alternatives. This dilemma presents itself in the case of VASS where there are a number of issues where even experts do not favor exclusively one alternative or another, but prefer some ranked combinations as it has been shown during VASS development (Halloun, 1996, 2001a).

CAD resolves this problem by allowing respondents to choose either alternative exclusively (options 1 or 5 in section III of VASS presented in the appendix), should they wish so, or a ranked combination of the two, either favoring one or the other (alternatives 2 or 4) or without any preference for one more than the other (alternative 3).

2.3 Likert rating scale and CAD

The interpretation mismatch we discussed in §2.1 above is common to a wide range of formats including, especially, Likert rating scales. Two respondents may express opposite positions on a Likert item for the same reason, or the same position for contradictory reasons (Aikenhead, 1988). When presented in a CAD format, respondents are focused on the context within which they need to answer a given question, and so are researchers in interpreting responses.

Had the question addressed in Figure 5 been asked in a Likert format:

The first thing I do when solving a physics problem is to search for formulas that relate givens to unknowns,

the student implicated in the figure would have undoubtedly agreed with the statement. Had the same question been asked differently in the same format, such as:

The first thing I do when solving a physics problem is to represent the situation with sketches and drawings,

the same student would have also agreed with the statement. Thus, contradictory results would be obtained with two Likert items that are supposedly intended to measure the same thing.

CAD rectifies the situation by formulating the discussed question in the form (cf. appendix):

The first thing I do when solving a physics problem is:

(a) represent the situation with sketches and drawings.

(b) search for formulas that relate givens to unknowns.

and asking respondents to contrast the two alternatives on 5-point rating scale going from "mostly (a)" to "mostly (b)", with a middle position corresponding to "equally" both or "(a) as often as (b)" (cf. appendix).

We devised a paper-and-pencil instrument to compare Likert and CAD formats and assess our position. The instrument consisted of two sets of questions bearing on six CAD items picked at random in VASS Form P20 (or P99LB). These are items 16, 18, 20, 28 (corresponding to Figure 5 above), 32, and 34 (cf. appendix).

The first set of questions consisted of 12 questions asked in a Likert format. The second set consisted of the six corresponding questions in VASS asked in CAD format. Respondents were asked to express their position on a 5-point rating scale in either a Likert question or a CAD one. Matching questions were asked in the same order in the two questionnaires. The two alternatives associated with each CAD question were though provided in two separate questions in the Likert questionnaire. In order to diminish the possibility of student figuring out our purpose and matching their answers on two Likert items corresponding to the same VASS item, we had such couples of items spaced apart in the Likert test. In this test, all six items were first asked with the corresponding CAD alternative (a), and then later followed with the corresponding CAD alternative (b).

The two sets of questions were administered concurrently to 230 secondary school students who had never taken VASS before. Students were first presented with the set of the 12 Likert questions. When finished answering these questions, the Likert questionnaire was collected, and student were presented immediately afterwards with the set of the six CAD questions. At no time, students had both questionnaires at hand simultaneously, and they were not allowed to reconsider their answers on either questionnaire once they have turned in their sheets.

Interviews conducted with some participating students revealed that they all understood the questions in the manner asked and that their answers actually reflected what they thought about the issues addressed. Most had realized our purpose after taking both questionnaires, and all had argued that their actual positions are better reflected in their answers on the CAD questionnaire than in those on the Likert questionnaire.

Student answers on the two questionnaires are compared in Table II. Three consecutive rows in this figure correspond to a particular VASS item. Percentages are given for answers on the two corresponding Likert items first, then on the corresponding CAD item. For example, Likert 1a and Likert 1b correspond to questions 1 and 7 in the Likert questionnaire, and match the first item in the CAD questionnaire (CAD1 in Table II).

Table II

Contradictions between Likert choices and discrepancies with CAD choices can be easily noticed in Table II for every single item. Take for example the first item, which deals with whether learning physics requires a serious effort and/or a special talent. In Likert format, 83% of respondents affirmed, by picking choices 1 or 2, that "learning physics takes a serious effort" (11% denied it) and 46% affirmed that it takes "a special talent" (36% denied it). Meanwhile in CAD format, 47% of the same respondents affirmed that it takes "mostly" a serious effort or "more of" that than of a special talent, and 38% affirmed that it takes both equally.

A similar situation presents itself with item 4 that deals with the issue of Figure 5 that we have been discussing all along this section. When asked about the first thing they do in solving a physics problem in Likert format, 74% of respondents sided with "sketches and drawings" (18% did not) and 76% sided with "formulas" (14% did not). When asked the same question in CAD format, 31% sided "mostly" or "more" with the former than the latter, and 39% equally with both (Table II).

Hence, our data show that Likert rating scale can be misleading, even when all possible alternatives to a given issue are asked about in separate questions. For example, Likert 4a and 4b in Table II tell us that about three quarters of respondents may resort to either "sketches and drawings" or to "formulas" in starting a problem, but they do not tell us how often they resort to one or the other. The problem would have been worse had we given only one of the two Likert items, which is normally the case with traditional Likert tests. For instance, had we restricted the test to Likert 1b exclusively, we would have been left with the wrong impression that 76% of respondents start solving physics problems with a quest for the right formulas. Yet, CAD4 data in Table II show that this is far from being the case!

The rest of the data in Table II shows practically the same thing for every single item, which leads us to the conclusion that the CAD format is significantly more reliable than the Likert format (or any other traditional format to that matter). In the CAD format, a person's response is internally normalized; one's position has to be stated for one alternative by comparison to the contrasting one and not in the absolute sense. This normalized comparison is not made possible with the Likert format (or other formats) where there is no point of reference relative to which one can state his/her position (Halloun, 2001a).

3. Validity and reliability of VASS

Validity and reliability of VASS have been constantly assessed as we developed the various forms. In our assessment, we resorted to classical statistical models and to Item Response Theory (IRT). We hereby limit our discussion to classical assessment whose results were practically the same as those of IRT based assessment (Halloun et al., 2001).

VASS dimensions are part of a broad spectrum of dimensions pertaining to what is often referred to in the literature as "nature of science" (scientific dimensions), "learning styles" and "attitudes" toward science or science education (cognitive dimensions). Moreover, the spectrum of issues within each dimension is so broad that whoever is constructing an instrument like VASS is forced to restrict the sampling of dimensions and issues to specific needs. *Content validity*, and more specifically *sampling validity*, would thus be concerned with those particular issues and not the total content area of the broad spectrum. In this respect, and as we mentioned in §1.2, our choice of VASS dimensions and items was based on what the literature, peer review and especially our own analysis of earlier forms of VASS have revealed to provide a meaningful snapshot of student views that significantly affect achievement in science courses. In the latter respect, the *predictive validity* of VASS was established in the manner discussed in §9.

Item validity of VASS was assessed in three respects. First, a number of university professors and experienced high school teachers who are versed in educational research pertaining to our work verified the validity of VASS items to assess what we intended to measure (Table I). Second, the same professors and teachers virtually all agreed on what we consider as expert answers to all questions, thus corroborating the *face validity* of the instrument. Third, we did exit interviews with some participating students in all stages of the project to ensure that students have understood the questions and the nature of the anticipated answers. No flaws were detected, especially after we accounted for instructor and student remarks given in the early stages of the development of VASS.

Reliability of VASS was essentially established in getting consistent measurements with parallel VASS forms and with a given form administered repeatedly in different semesters to similar samples. Because of the nature of our instrument, *internal consistency reliability* was assessed indirectly, and not with commonly used coefficients like Cronbach's alpha (or Kuder-Richardson's for dichotomous items).

Classical reliability coefficients like Cronbach's alpha are underlined by many assumptions especially those of unidimensionality (or the existence of no sub-scales) and linearity of relationship among various items. However:

- VASS items are distributed in six dimensions that can be grouped in at least two subscales; these are the broad scientific and cognitive domains in Table I. These subscales are of different nature, both from an ontological and from an epistemological perspective.
- Because of the relative importance of each dimension and its impact on learning and achievement, the number of items varies from one dimension to another, and from one subscale to another. This further prevents us from resorting to split-half reliability assessment between the two subscales.
- The loading of items within subscales or of individual dimensions on the entire instrument is not uniform, partly because of the two points above. A rough estimate of such loading can be inferred from Pearson's correlation coefficient between a given dimension and the corresponding subscale, or between each subscale and the entire instrument. In one administration of VASS, this coefficient took the respective values of .40, .61 and .78 when correlating the dimensions of structure, validity and methodology

with the scientific subscale. It took the respective values of .43, .56 and .91 when correlating the dimensions of learnability, personal relevance and reflective thinking with the cognitive subscale. Finally, it took the respective values of .64 and .92 when correlating the scientific and the cognitive broad domains with the entire instrument.

- ♦ The assumption of linearity is not quite verified, at least between items belonging to different subscales.
- ♦ Different item formats are used in the new VASS forms (P99LB and P20). Items 6 through 15 in section II are dichotomous, while items 16 through 39 in section III are CAD items.

All this undermine the two assumptions in question and rules out the viability of classical coefficients for assessing VASS internal consistency; notwithstanding the fact that, in the case of VASS, such coefficients do actually take values that are comparable to those commonly reported in the literature for "reliable" instruments.

VASS internal consistency can be assessed indirectly instead, in terms of the relative difficulty of the six dimensions. A detailed scoring scheme of VASS is discussed in §6 where we show that student average scores are about the same on all six dimensions (Table VII, §6). This lends some support to the reliability measure in question.

Now, let us turn to test-retest and stability reliability of our instrument.

In one form of *test-retest reliability* assessment, some of the interviewed students were asked to orally answer specific VASS questions, a few days after they had filled the written survey, without reminding them of their written answers. Virtually all these students reiterated the same answers they had indicated previously. In a related aspect, we included an item at the end of the instrument asking students how seriously they took the test (item 39). At least 88% of respondents expressed a positive position in this respect every time VASS was administered (Table IV, §5).

For *stability* (and *equivalence*) assessment, we compared results obtained with parallel forms of VASS, mainly forms P12 and P99LB, as well as P99LB and P20, and results obtained with a given form administered to similar samples of students taking the same course in different semesters.

Figure 6 shows boxplots for similar samples of secondary school students who took form P99LB in the spring of 1999 and form P20 in the spring of 2000. The 1999 sample consists of 167 students, and the 2000 sample consists of 169 students. Students in both samples were enrolled in the same class (1st secondary), at the same Lebanese secondary schools, but obviously in two different years. Notwithstanding the insignificant differences in VASS data among schools and grades discussed in §8, these two samples were expressly picked for comparing the two VASS forms in order to minimize sample variation. A similar situation would have actually been presented should we have included all participating students (§8 and §9).

Figure 6

The reader can easily notice that, except for items 18, 22, 23, and 35, there are no significant differences in the boxplots of the same CAD items in both forms. The median is the same for every item other than these four, and the significant range about that median (25th percentile through 75th percentile) does not vary significantly. Fluctuations in the four items in question are thus insignificant by comparison to the overall performance on the test.

Answer distributions were also similar on forms P12 and P99LB administered to comparable samples of U.S. and Lebanese students respectively. This can be seen in Figure 7 that is discussed in ample details in section 5 of this report.

Figure 7

By the time this report was sent to press, Form P20 data were still being collected in different countries around the world. Preliminary analysis shows that, like in the case of Form P12, answer distributions are virtually identical for similar groups of students taking VASS in different semesters. More ample details on this stability issue and other validity and reliability matters will be given elsewhere (Halloun et al., 2001).

Student Views and Profiles

VASS form P99LB was developed in the spring of 1999 and administered subsequently to over two thousand Lebanese students during that year. This form was revised in the manner described in §1 and evolved into Form P20. The latter form is, for all practical purposes, equivalent to form P99LB. All items are the same, except for some minor wording revision (§1.3), and given in the same order in both forms. Data obtained with the two forms are significantly similar (Figure 6, §3). Form P99LB has been administered to a larger number and to a bigger variety of students than form P20 so far. Therefore, and unless otherwise specified, our following discussion will concentrate on Form P99LB.

4. Samples

VASS Form P99LB was administered during the months of May and early June of 1999 to over one thousand Lebanese students enrolled in physics courses at Lebanese secondary schools (grades 10, 11, and 12) and universities (U.S. college level) spread throughout the country. In the following October and early November, it was administered to a sample of similar size and characteristics. Given the limited number of items in some VASS dimensions (Table I), students were dropped out of the analysis who failed to answer 5 questions or more in sections II and III of either test (cf. appendix), or who made 5 confusing marks or more on their answer sheets. Overall attrition rate was about 17%, which is not too bad given the unfamiliarity of Lebanese students with similar surveys and reluctance they sometimes express in taking tests that do not count in their course grades. In all the analysis that follows, we will thus account only for those participants with acceptable answer sheets on Form P99LB. Their number adds up to 813 students in the '99 spring term and 847 in the following fall term. These figures do not include students who participated in the Likert-CAD comparison (§2.3) and in the validation of Form P20 (§3).

Students were distributed in fifteen secondary schools and two universities in the spring term, and in twelve schools and three universities in the following fall term. Table III shows the characteristics of the two samples in question. No gender data was explicitly collected until Form P20 was devised in the spring of 2000. However, we have accounted for this variable indirectly in the manner described in §8, through the classification of participating schools according to the dominant gender in each school.

Table III

Fall data precede spring data in Table III because, in the following sections, we shall treat the former as pretest data and the latter as posttest data. This position will become evident during the course of our discussion. For now, we will point out two supporting facts. First, the performance of each of the sub-samples shown in Figure 6 is representative of the performance of the entire corresponding sample (§3 and §7). Hence, data collected in the spring of 1999 are not significantly different from data collected in the spring of 2000. Second, our prior experience does not show significant differences among similar samples across the years (Halloun, 1996, 1997; Halloun et al., 2001). Therefore, we can safely use the 1999 fall data as pretest indicators, and the 1999 spring data as posttest indicators.

Our discussion in §5 through §8 that follow in this report will concentrate on participants' answers to VASS questions corresponding to the taxonomy presented in Table I. Discussion on course achievement indicated by student responses on items 1, 2 and 3 of section I in Form P99LB is deferred to §9. An overview of student views on individual items of the taxonomy is first given in §5. Based on those views, students are classified into four groups, each characterized by a distinctive profile in the manner discussed in §6. Student views and profiles are compared, first, in §7, as they evolve from pretest to posttest in the same grade, then across different grade levels, and, in §8, across various demographic strata. All along our discussion, Lebanese students' performance is compared to that of their US peers.

5. Student views: An overview

Fall data assess student state before instruction, while spring data assess it after instruction. Table III shows that fall participants are apparently better distributed than spring participants across demographic variables distinguished in this table (we shall see later that this has virtually no effect on our results). For these reasons, and in order not to burden the reader with excessive and redundant data, we will analyze data obtained only in the fall semester of 1999 in this section and in the following one.

In prior VASS forms, we used to characterize student views on every single item in one of three categories, expert, mixed, or folk, based on whether a student answer on the item matches or not the corresponding answer given by university professors and experienced high school teachers (§1 and §3). A student answer would be considered as *expert* in the event of a match, as *folk* when it is on the opposite end, and as *mixed* when it is somewhere in between. The expert choice, and thus the other two, varied from one item to another. It corresponded, in Form P12, to choice 7 exclusively on the 8-point CAD scale (Figure 2), when alternative (b) is the expert pole (or 1, when alternative (a) is the expert pole), or to a combination of this choice with the following one or two choices (6 and 5, or 2 and 3). The variation was imposed by the nature of the 8-point scale used then (Halloun, 1996, 1997; Halloun & Hestenes, 1998).

With the new 5-point CAD scale, the same three view categories may be kept, at least for the sake of comparison with old data. The cutoffs in Forms P99LB and P20 are set though in more straightforward manner. An *expert* view now correspond to choice 5 or 4 on any CAD item where alternative (b) is the expert pole, and to choice 1 or 2 on any CAD item where alternative (a) is the expert pole. In contrast, a *folk* view corresponds to choice 1 or 2 in the first case, and to choice 4 or 5 in the second case. In any case, choice 3 denotes a mixed view.

The distribution of all participants' answers on the entire VASS Form P99LB is shown in Table IV and Figure 8 (Fall 1999). No distinction is made at this point between various strata identified in Table III. Stratum comparison is the object of later sections. In both displays, and to make the comparison easier across the entire instrument, student answers on a given item have been rearranged, if necessary, so that the expert pole is always scored as 1 in dichotomous items and 5 in CAD items (as well as in multiple choice items in section I). Section I items have been deferred to the end of both displays because they are not analyzed in this section. Items 1, 2 and 3 will be addressed in §9. Items 4 and 5 were not included in prior VASS forms; they are only analyzed in §6. Item 39 is slightly separated from the rest of the items in Table IV to highlight the fact that it is a control item that served its purpose in §3 and that it does not contribute to our analysis of student views targeted in Table I.

Table IV *then* Figure 8

Table IV and Figure 8 make it clear that a good majority of participants expressed expert views on all dichotomous items (Q6 through Q15), but not as much on CAD items (Q16 through Q38), which is expected given the nature of addressed issues and of the scale. The following are major characteristics of student answers on the 23 CAD items:

- The *median* corresponds to the mixed view (response 3) in 8 items. These being items 16, 17, 20, 26, 28, 33, 34, 35. It is located in the expert range for the rest of the items, but only at position 4 for all of these items except for item 30 where answers are polarized toward the expert pole (response 5).
- Among the eight items with a *mixed view median*, student answers in the significant range (between the 25th percentile and the 75th percentile) are clustered above the median in 5 items (16, 17, 33, 34 and 35 with a lesser extent on the latter), and below the median in 1 item (20); they are distributed without significant difference on both sides of the median in the remaining 2 items (26 and 28).
- Among the fifteen items with an *expert view median*, no answers in the significant range in question are clustered above the median. Answers are clustered within the expert range in only item 30, and below the median down to the mixed view in 4 items (18, 19, 25, and 29), and further down to a folk view in 6 items (21, 22, 24, 27, 31, and 36). Distribution on both sides of the median can be noticed in the remaining 4 items (23, 32, 37, and 38).

A closer analysis of response distribution across the three view types, expert, mixed and folk, in table IV and Figure 8 reveals the following:

- Five CAD items (22% of CAD items) have respondent answers *significantly polarized* toward the *expert* pole. These are items 23, 29, 30, 32, and 38. The first four of these items belong to the reflective thinking dimension, and the fifth (item 38), to the methodology dimension (Table I).
- Eight CAD items (35%) have respondent answers *polarized* toward the *expert* pole but with a heavy clustering on the *mixed* view. These are items 16, 17, 18, 19, 25, 33, 34, and 37. Except for item 37 that belongs to the methodology dimension, they all belong to cognitive dimensions (Table I). More specifically the first seven items are distributed

as follows: learnability (item 16), reflective thinking (items 19, 25, 33, and 34), and personal relevance (items 17 and 18).

- Seven CAD items (30%) have respondent answers *unpolarized* toward either pole, and distributed across all three view types (expert, mixed and folk) without significant differences. These are items 21, 22, 24, 27, 31, 35, and 36. Two items belong to a scientific dimension, validity (items 35 and 36), and the remaining five are spread across the three cognitive dimensions: learnability (item 22), reflective thinking (items 21, 27, 31), and personal relevance (item 24).
- Two CAD items (9%) *are slightly polarized* toward the *folk* pole but with a heavy clustering on the *mixed* view. These are items 26 and 28, which belong to the methodology dimension (Table I).
- Only one CAD item (4%) is characterized by an *overall folk view*. This is item 20 of the learnability dimension.

In sum, participants' views do not fall systematically in one type or another. Except only for item 30, there is no clear-cut polarization of student views toward one pole or another (expert/folk) on any single item. However, and except for items 20, 26 and 28, more respondents tend to express overall mixed or expert views than folk views in any particular VASS item. The mixed-expert trend is a little more pronounced in cognitive dimensions than in scientific dimensions. The trend in question reflects though the overall state of the entire sample and not the state of any single respondent. In fact, not a single respondent gave answers that fall entirely in one category or another. This can be easily inferred from Table V that shows the distribution of respondents giving one type of view or another, first on CAD items (items 16 through 38), then on dichotomous items (items 6 through 15), than on all 33 CAD and dichotomous items.

Table V

The situation is significantly reminiscent of the state of students of the same educational level elsewhere in the world, especially in USA. To show this point, let us compare results reported above to those obtained with VASS Form P12 in USA. This form consists of 31 CAD items (8-point scale). Twenty-three of these items are identical in both forms, P12 and P99LB. These are items 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38 and 39, in Form P99LB (or in equivalent Form P20 given in the appendix). The remaining eight items have parallel dichotomous items in Form P99LB. These are items 7, 8, 9, 10, 11, 13, 14 and 15 in the latter form.

By comparison to pretest results in the '96 fall semester in USA, we notice the following (Halloun, 1996, 1997; Halloun & Hestenes, 1998):

- Except for items 13 and 15, the medians of dichotomous items in Form P99LB and of parallel CAD items in Form P12 are all in the expert range. Furthermore, corresponding respondent answers in both forms are clustered close to the median. This was actually one major reason behind the transformation of these CAD items in Form P12 into dichotomous items in Forms P99LB and P20.
- Items 13 and 15 have actually been reworded in the new form in a manner that makes it unfair to include them in the comparison.
- The medians are exactly the same for 18 items out of the 22 common CAD items (82%

of these common items), as well as for control item 39 in the new form. On item 24 of the learnability dimension, the median is one point higher for U.S. students than for Lebanese students. In contrast, the median is one point higher for Lebanese students on the remaining three items. These are items 29 (reflective thinking), 35 (validity) and 37 (methodology) in the new forms.

- The distribution about the median is almost identical on thirteen items (59% of the 22 common CAD items with identical means). These are items 17, 18, 22, 23, 24, 25, 27, 28, 30, 31, 32, 33, 36 and 38 in the new forms. Less spread out about the median is detected on four items in U.S. data (items 26, 34, 35 and 37), and on five items in Lebanese data (items 16, 19, 20, 22, and 29). However, the spread out difference is not statistically significant on either one of the last nine items.

Virtually the same pattern of responses can be detected in posttest data shown in Figure 7 for comparable samples of U.S. and Lebanese students. Data now coming from many other countries around the world are also showing similar patterns. Whence the conclusion that, by and large, *student views about knowing and learning science assessed in VASS are universal*. They are similar on virtually every single scientific or cognitive dimension. The significant similarity shall become more evident when we discuss student profiles next.

6. Student profiles

A major objective of VASS, from the date of its conception, was to identify patterns in student views and classify them in general profiles. To this end, and as discussed above in §5, we originally classified the response to each VASS item in one of three categories: expert, folk or mixed. Subsequently, a profile was built for every student characterized by the number of answers provided in each category. Student profiles fell then into four types: expert, high transitional, low transitional, and folk.

6.1 Toward a new profiling scheme

The four profiles distinguished with Form P12 were then characterized by the number of items with expert and folk views as given in Table VI. Cutoffs between profiles are based on a detailed analysis of experienced teacher responses (Halloun, 1997; Halloun et al., 2001). The four profiles are further distinguished qualitatively along the six dimensions of VASS shown in Table I (Halloun & Hestenes, 1998). Distinguishing features are discussed later in this section.

Table VI

The profiling scheme shown in Table VI was the best scheme we could come up with, given the 8-point CAD scale used in all VASS forms up to Form P12. It has enabled us to classify students in distinctive categories, and to assess meaningfully the relationship between student views about knowing and learning science on the one hand, and achievement in various science courses, on the other (Halloun, 1997; Halloun & Hestenes, 1998). However, certain teachers, especially those with little statistical background or with limited resources, have complained that this scheme is a bit involved, and wished if we could find a more user-friendly approach. In fact, we had explored other schemes with the old forms, but they have all proven not to be valid and reliable alternatives (Halloun, 2001a).

One scheme that all teachers wished we could resort to was traditional sum of scores. The 8-point CAD scale previously used was more of an ordinal nature than of an interval nature, thus preventing us from taking the desired path. With the new 5-point scale, along with the revision of some items, it became possible to treat CAD items in forms P99LB and P20 more like *scaled* items (*interval* type) than *ranked* items (*ordinal* type). Hence, a sum of scores or perhaps a weighed sum of scores looked like viable alternatives, should they result in profiles in line with those distinguished in Table VI while obtained and employed in more efficient ways.

In order to explore such alternatives, we proceeded in two steps. First, a sum of scores was calculated and analyzed within each of the six VASS dimensions distinguished in Table I. Then an overall sum of scores – and a weighted sum of scores – was obtained for the entire instrument and analyzed, especially in terms of participants' achievement in their physics courses (§9).

Histograms of participant responses are displayed in Figure 9 across the six VASS dimensions distinguished in Table I, and corresponding descriptive data are shown in Table VII. The maximum score for each dimension is computed in function of the numerical value associated with each answer as displayed in Table IV. Hence, the maximum score is 1 for dichotomous items, and 5 for CAD items, and the minimum score is 0 and 1 respectively for the two types of questions. We had explored with different scoring schemes, including keeping the same maximum score for both types of questions, but things did not work out well, especially when we tried to correlate VASS score with achievement in physics courses.

Figure 9 *then* Table VII

An examination of data displayed in Figure 9 and Table VII reveals the following:

- Means of the sum of scores in the six dimensions are virtually all equal to about two-third the maximum score in any given dimension. This shows at least two important points:

  1. The overall trend of respondents' answers is oriented more in the mixed-expert direction than in the mixed-folk direction as discussed in §5 above. This is further supported by the fact that the sum of scores in every dimension is above the corresponding 3/5th of the maximum score for two thirds or more of the respondents (whose score is in the range of the mean ± 1 S.D.).

  2. Relative item difficulty as estimated by means is the same across all six dimensions, which supports the relative coherence of VASS dimensions and the instrument reliability. Standard deviation implications are discussed in the next bullet.

- The standard deviation is smallest for the structure dimension and largest for the reflective thinking dimension; it is significantly close for the other four dimensions. The difference between the two extremes should be interpreted very cautiously. All items in the structure dimension are dichotomous, whereas all items in the reflective thinking dimension are CAD items. They are a mix of both in the other dimensions. Furthermore, the structure dimension includes five items whereas the reflective thinking dimension includes eleven items (not counting new items 4 and 5 in Table I). As a consequence, the latter dimension has the highest maximum sum of score (55) among all dimensions, a maximum that is eleven times bigger than that of the structure dimension, and about equal to, or three times more than, that of every other dimension. The wider CAD scale, the bigger number of items, and the subsequently highest maximum score thus account to a large extent for the relatively large standard deviation of the reflective thinking dimension. Therefore, the conclusion we made in the second point of the preceding bullet still holds.

- The same holds for the sum of scores in all three scientific dimensions combined, on the one hand, and the sum of scores in all three cognitive dimensions on the other. The difference in the corresponding standard deviations can be largely accounted for by the bigger number of dichotomous items in the scientific sub-scale (8 out of 14) than in the cognitive sub-scale (2 out of 19 items), and by the overall bigger number of items in the latter, and thus by the significantly larger maximum sum of scores (87 to 38, cf. Table VII).

These results are highly compatible with those obtained in USA and introduced in the previous section. The reader is hereby spared unnecessary details that may be found elsewhere (Halloun, 1996).

In order to decide on a new profiling scheme, we followed two parallel paths. First, we carefully examined data displayed in Figure 9 and Table VII along with the distribution of the overall sum of scores shown in Figure 10, and this, by correspondence to data obtained in the 1999 spring term (§5 and §9) and in contrast to the scheme of Table VI (Halloun & Hestenes, 1998). Second, and in a stepwise regression analysis, we analyzed the correlation of the various VASS dimensions with achievement in physics courses participants were enrolled in.

Figure 10

We finally settled on a scheme based on a regular sum of scores for three reasons. First, simple sum of scores correlates significantly with course achievement, though not as significantly as a VASS indicator computed in a stepwise regression analysis (details in section 9 of this report). Second, it is a convenient approach that teachers of all levels can resort to, irrespective of their background and their resources in statistics. Third, and most importantly, the sets of profiles we came up with, following, first, simple sum of scores, then weighted sum of scores, did not differ significantly, neither from a quantitative perspective (§6.3) nor from a qualitative perspective (§6.4).

6.2 Profile cutoffs

Our statistical analysis led us to identify three cutoffs in the overall sum of scores that would attribute to a given student one of four profiles:

1. The expert profile (EP), characterized by a sum of scores greater than 92.
2. The high transitional profile (HTP), characterized by a sum of scores in the [83,92] range.
3. The low transitional profile (LTP), characterized by a sum of scores in the [73,82] range.
4. The folk profile (FP), characterized by a sum of scores smaller than 73.

In the '99 fall term, 14.5% of participants were characterized with EP, 44.8% with HTP, 31.4 with LTP and 8.3 with FP. Only 0.7% of all students (4.8% of EP students) scored above 102 in the expert range [93,125], i.e., more than 10 points above the upper limit of the next profile, and only 1.3% of all students (15.6% of FP students) scored below 63 in the folk range [0,72], i.e. more than 10 points below the lower limit of the next profile. The practical range of each outer profile is thus significantly close to that of the middle profiles. This gives a special significance for the 10-point range of the middle profiles (or the 9-point difference between the range upper and lower limits which is equal to one S.D.).

The new profiling scheme may appear to give a more optimistic picture of student views than the old scheme described in Table VI. That scheme placed about equal proportions of students in the folk profile and each of the two transitional profiles (Halloun, 1997). As we shall in §6.3 and §6.4, this is only a matter of profile interpretation. The actual distribution of student views expressed on individual VASS items is the same, under any profiling scheme, virtually anywhere in the world (§5). With the new scheme, we have redefined the demarcation line between various profiles, especially between the folk and the low transitional profile, so as to sharpen the contrast between folk and expert profiles.

Some caveats regarding the interpretation of VASS scores and profiles are worth noting before we go any further with our discussion:

- VASS items, and mainly CAD items, are *interval*, *not ratio* type variables. Hence a sum of score on VASS has to be interpreted and operated with accordingly. Thus, a VASS score of 100 may not be interpreted as twice as good as a score of 50. The interpretation can only be done in terms of the profiles identified above, i.e., that a score of 100 puts a student in the expert profile category while a score of 50 places her/him in the folk profile category. The two profiles have the distinctive features discussed below.
- VASS scores may be subject to virtually all kinds of statistical tests that ratio type of scores can be subject to (correlation, analysis of variance, etc.). However, the results of such tests provide us only with *rough indicators* of the relationships we are assessing between variables, and may not be interpreted with the same degree of confidence as results obtained with ratio type variables.
- Score variation within profiles is not significant, especially within a 10-point range. A student with a score of 82 is not significantly better than a student with a score of 73. Both students have a low transitional profile and share the features discussed below. A difference between scores of more than 10 points within the folk profile or the expert profile may be significant from an ontological and developmental point of view. The higher a student scores above 102, the closer s/he gets to becoming a real expert, and the lower s/he scores below 63, the more entrenched s/he would be in the folk profile and the harder her/his transition toward upper profiles. However, such extreme cases rarely occur as we noted above, and, when they do, they do not have a significant impact on course achievement within the same profile (§10).

6.3 Profile comparison: Quantitative differences

Like the four profiles associated with Form P12 (Table VI), the new four profiles are mainly distinguished by the number of expert views and folk views as expressed in the first part and second part respectively of each statement in Table I. More specifically, and as can be seen in Figures 11 and 12 and in Table VIII, the four profiles have the following quantitative characteristics:

Figure 11, *then* Table VIII *adjacent to* Figure 12
----

1. No one of the four profiles is characterized by exclusively one type of view or another (expert, mixed, folk). Virtually every student expressed an admixture of the three view types over the entire instrument, including those students at the two ends of the spectrum. For instance, in the 1999 fall semester, respondents with the highest VASS sum of scores expressed expert views on a maximum of 30 items (out of 33), and those with the lowest sum of scores expressed folk views on a maximum of 23 items. One of the former EP respondents expressed folk views on all three items remaining; the others expressed mixed views on one or two of these items, and folk views on the other(s). One of the latter FP respondents expressed expert views on all ten items remaining; the others expressed an admixture of expert and mixed views on these items.

2. The percentage of respondents expressing expert views on virtually any given VASS item increases gradually, and significantly, from the folk profile to the expert profile, across the two transitional profiles (Figure 11). The following are related results pertaining to the '99 fall sample. The other samples' data do not differ significantly:

- The proportion of respondents with a *folk profile* who expressed expert views did not reach 50% on any of the 23 CAD items (items 16 through 38), as well as on half the dichotomous items (specifically items 6, 9, 10, 13 and 15).
- The proportion of respondents with a *low transitional profile* who expressed expert views exceeded 50% on all but two dichotomous items (6 and 15), but remained below 50% on all but 4 CAD items. The latter are, in ascending order, items 23 (52%), 32 (56%), 38 (62%) and 30 (68%).
- The proportion of respondents with a *high transitional profile* who expressed expert views exceeded 50% on each of the ten dichotomous items, ranging from 55% (item 15) to 92% (item 8), and on 16 CAD items, ranging from 53% (items 16 and 35) to 89% (item 30).
- The proportion of respondents with an *expert profile* who expressed expert views exceeded 78% on all but two dichotomous, these being items 15 (56%) and 29 (70%), and ranged from 55% (item 16) to 97% (item 30) on all but 4 CAD items. The latter are, in descending order, items 28 (50%), 34 (44%), 26 (37%), and 20 (27%).

3. Only respondents with an expert profile express expert views on all ten dichotomous items (Q6 through Q15), and on 26 items (79%) or more in the entire instrument (Table VIII). In the 1999 fall term, respondents who expressed expert views on at least 79% of all items in VASS make up 26% of respondents with an expert profile, and 4% of the entire sample.

4. In parallel, the percentage of respondents expressing folk views decreases gradually, and significantly, from the folk profile to the expert profile (Figure 11). Only respondents with a folk profile express folk views on more than half the items in VASS (17 items or more). In the 1999 fall term, respondents who did so make up 36% of those with a folk profile, and 3% of the entire sample.

Figure 13 compares answers given on the entire instrument by the two extreme groups identified in the last two points: the EP group of students expressing expert views on at least 79% of VASS items and the FP group of students expressing folk views on more than half the items. The reader can easily notice the wide disparity between the two groups of respondents on every VASS item.

Figure 13
----

5. The disparity between profiles becomes more noticeable when we concentrate on the expert and folk poles in CAD items (answers with scores 5 and 1 respectively in Table IV). The proportion of respondents who picked the expert pole on any given item reached a maximum of: 24% within the folk profile, 35% within the low transitional profile, 64% within the high transitional profile, and 86% within the expert profile. These maxima are reached: in item 18 within the folk profile, and in item 30 within the other three profiles. In contrast, the highest proportion of respondents who picked the folk pole decreases gradually from 46%, within the folk profile, to 26% then 15% then 6% within the other three profiles respectively. These maxima were reached in item 31 within all four profiles.

6. The percentage of respondents expressing mixed views on any particular VASS item, as found in detailed frequency analysis, is relatively the lowest within the expert profile then within folk profile (except for seven items where the order is reversed). In contrast, the percentage in question is relatively the highest within the low transitional profile then within the high transitional profile (except for six items where the order is reversed).

7. The range of mixed views is the widest within the two transitional profiles (0 to 14 items with mixed views in each profile), and the narrowest within the expert profile (0 to 9 items, with the smallest median of 4). However, and as can be seen in Table VIII and Figure 12, the distribution of mixed views as expressed by the respective median and percentiles does not vary significantly across the four profiles. In these respects, the four profiles are more significantly distinguished by the total number of expert views and folk views expressed on the entire VASS instrument than by the number of mixed views:

- The expert profile is characterized by predominantly expert views. In this profile, the number of expert views ranges from 19 to 30 with a median of 24, whereas the number of folk views ranges from 1 to 9 with a median of 5 (Table VIII).
- The folk profile is characterized by predominantly folk views. In this profile, the number of folk views ranges from 5 to 23 with a median of 16, whereas the number of expert views ranges from 7 to 17 with a median of 12. Hence, the number of folk views within the folk profile can reach a maximum that is more than 2.5 times the respective maximum in the expert profile. Moreover, the maximum number of 17 expert views within the folk profile is smaller than the minimum number of 19 such views within the expert profile (Table VIII).
- The range and median of expert views increase gradually from the folk profile to the expert profile, across the two transitional profiles, with no overlap between the respective 25th percentile and 75th percentile (Figure 12). Moreover, and as we go down gradually from the expert profile to the folk profile, we notice that the median number of expert views in each profile is virtually equal to the maximum number of such views in the profile immediately below (Table VIII).
- The range and median of folk views decrease gradually from the folk profile to the expert profile, across the two transitional profiles, again with no overlap between the respective 25th percentile and 75th percentile (Figure 12). Moreover, and as we go up gradually from the folk profile to the expert profile, we notice that the median number of folk views in each profile is virtually equal to the maximum number of such views in

the profile immediately above (Table VIII).

All in all, data show that a person with a given profile is not warranted to hold a particular view type on any given VASS item. However, the points mentioned above, coupled with a chi-square analysis between respondents' actual answers on individual VASS items and their respective profiles, show that each of the following scenarios is significantly more likely to occur than any other scenario:

1. An expert view to be expressed by a person with an expert profile, especially when the view is polarized toward the expert end on the 5-point CAD scale.

2. A mixed view to be expressed by a person with a low transitional profile or, to a lesser extent, by one with a high transitional profile.

3. A folk view to be expressed by a person with a folk profile, especially when the view is polarized toward the folk end on the 5-point CAD scale.

Table IX shows the chi-square values and their significance for a cross-tabulation done between respondents' actual answers on the various VASS items and their respective profiles. The table and the underlying cross-tabulation show that, except for items 11, 15, and, to a lesser extent, for items 16 and 20, a person's answer on any VASS item is significantly determined by her/his VASS profile in the manner just described in the three points above.

Table IX

The disparity is most significant between folk and expert profiles with regard to virtually every aspect assessed in VASS, and especially those aspects assessed in items 6, 7, 17, 22, 23, 25, 27, 29, 30, 31, and 32 (Table IX). Thus, when it comes to learning physics, respondents with a folk profile are especially distinguished from all others by the fact that they are: (a) the least motivated and the least perseverant, (b) the most authority-dependent, (c) the most keen on rote learning and on avoiding self-evaluation, and (d) the most inclined on gathering piecemeal and situation-specific information. These four distinguishing features, as well as others, get gradually reversed as we move through the two transitional profiles up to the expert profile (§6.4).

6.4 Profile comparison: Qualitative differences

The four distinguishing features just identified and others are consistent with those obtained for the four profiles conceived with VASS Form P12 (Table VI). We had these profiles described qualitatively elsewhere with ample details (Halloun & Hestenes, 1998). In the following, we review some of the major features that we discussed then and that are being sustained with the current results of VASS Form P99LB and P20. The review is done within the context of the two VASS subscales: the set of scientific dimensions and the set of cognitive dimensions. But first, we call our readers' attention to two major points in our presentation.

First, our attempt to systematize student views within profiles does not necessarily reflect an actual coherence of student views within any given dimension. For instance, and as we mentioned in §6.3, not a single student expressed views of the same type (expert, mixed, or folk) on all VASS items, and rarely did a student with an expert profile (EP) express expert views on all items within a given dimension, or did a student with a folk profile (FP) express folk views on all these items. However, and as we showed in §6.3, EP students consistently expressed more expert views than mixed or folk, while FP students expressed more folk views than other students.

Second, our analysis of VASS results necessarily touches on deep philosophical issues about the nature of science and cognition, but it ignores many philosophical subtleties in order to capture *broad tendencies* in philosophical viewpoints. In particular, our overall classification of VASS results into profiles aims to distinguish a prototypical expert (scientific) viewpoint from a prototypical folk (unscientific) viewpoint. Experts sometimes choose folk responses to individual VASS items for good reasons, which VASS is not designed to detect. However, the expert profile is defined broadly enough to encompass such differences of opinion within the expert camp. Our interviews had shown that: (a) students with folk profiles seldom have well considered reasons for their choices on individual VASS items, whereas those with expert profiles often exhibit considerable insight in their justifications, and that (b) the folk views are generally more heterogeneous than the expert views.

In the following we use terms like ‘scientific realism’ loosely, to indicate broad philosophical perspectives. We are not concerned with the technical definitions needed to articulate a sharp philosophical position. Our meanings for such terms should be sufficiently clear from the context.

Scientific dimensions:

Research on student views about the nature of scientific knowledge has produced conflicting results. Edmondson and Novak (1993) reported that the “majority of college students hold essentially positivist views [that knowledge] is discovered through observation, unfettered by previous ideas or beliefs”, and that scientific knowledge consists of “separate, objective truths that are domain-specific and constant”. However, Aikenhead (1987) found that only 25% to 36% of Canadian high school graduates hold such positivist views, while 45% recognize scientific knowledge as a human construction and a partial representation of reality. Aikenhead also reported that while “almost all students would seem to agree that scientific knowledge is tentative, but ... in different and often conflicting ways”, 44% believed that this knowledge may be subject to change, and 31% believed that it may not. Songer and Linn (1991) classified middle school students’ views of science “into three groups: static, mixed, and dynamic. Those who view science as static [21%] assert that science consists of a group of facts that are best memorized. Those who view science as dynamic [15%] believe that scientific ideas develop and change... Students with mixed beliefs [63%] hold some static and some dynamic views”. Though not completely in agreement, VASS results are closer to the findings of Aikenhead and especially of Songer and Linn than to those of Edmondson and Novak.

VASS results have consistently shown that, with respect to the scientific dimensions shown in Table I, no more than 20% of college students and 25% of high school students (in Lebanon and USA) have views indicative of *positivism* or *naive realism*, while at least 20% of students of any level hold opposing views consistent with *scientific realism*. The remaining students hold admixtures of both types of views. Most naive realists have folk profiles. Virtually all students with an expert profile and some students with a transitional profile (mostly those with a high transitional profile) are essentially scientific realists.

Naive realists believe that the physical world is exposed directly to our senses and that scientific knowledge mirrors this reality. Consequently, they often believe that scientific knowledge is exact, absolute and final, as well as situation-specific, piecemeal, and developed from arbitrary rules of thumb. Naive realists often believe that physics is guided by mathematical rules for manipulating formulas.

In contrast, scientific realists believe the physical world cannot be known directly through sense perception, but only indirectly through theoretical constructions. Consequently, they believe that scientific knowledge is approximate, tentative and refutable, as well as generic, coherent and systematically structured and applied.

Cognitive dimensions:

There is, of course, an enormous literature pertaining to each of the cognitive dimensions probed by VASS (Table I). Here is a sampling of relevant work:

- Arons (1984), Reif (1987), and Reif and Larkin (1991) discuss the widespread belief among students that physics can be learned by memorizing factual knowledge and formulas piecemeal.
- Analyzing the discourse of college students about physical phenomena, Cobern (1993) found that "most students assigned science a minor role in their lives", noting that "what was most striking about the interview texts was the conspicuous absence of scientific talk, although these students had successfully completed several college science courses and were majors in a science-related field".
- In a national survey conducted by the National Science Board, only 52% of 17-year old high school students considered that "most of what [they] learn in science classes is useful in everyday life" (NSB, 1993).
- In independent studies with junior high school students, Simpson and Oliver (1985) and Baker and Piburn (1991) found that student attitudes toward science declined significantly following science instruction. Ebenezer and Zoller (1993) reported that, although 73% of students "feel the study of science in school is important", only 38% "would like to study more science".

VASS results in the cognitive dimensions (Table I) show that, when it comes to learning physics, no more than 28% of college students and 22% of high school students (in Lebanon and USA) can be characterized as *passive learners,* while at least 14% of students of any level are *critical learners.* The remaining students hold mixed cognitive views. Most passive learners have folk profiles. Virtually all students with an expert profile are critical learners, as is a good fraction of students with a high transitional profile.

Passive learners are little motivated to learn physics and not perseverant in their study. They are authority-dependent, believing that their understanding of physics depends more on instruction than personal effort. They tend to concentrate on isolated facts and formulas in physics, memorizing them by rote without relating them to prior knowledge. They see little relevance of physics to everyday life, and so their concern with physics is limited to satisfying course requirements.

In contrast, critical learners are authority-independent, believing that their understanding of physics depends more on personal effort than instruction. They tend to concentrate more on reasoning processes than factual information in physics. They are reflective thinkers, seeking a coherent understanding of physics, striving to detect and resolve discrepancies between accepted scientific knowledge and their own. Critical learners see physics as relevant to everyday life, so they pursue the study of physics more for personal benefit than for fulfilling curriculum requirements. Moreover, and as shown by new items 4 and 5 in

Forms P99LB and P20, critical learners are more curious than passive learners to learn about scientific developments covered in various media.

Relation between scientific and cognitive dimensions:

Many researchers have argued that student beliefs about the nature of scientific knowledge are coupled to their learning styles (Edmonson & Novak 1993; Hammer 1994; Reif & Larkin 1991; Songer & Linn 1991; Tobias 1990). VASS results support this conclusion.

Analysis of cross-tabulation between the scientific domain of the four profiles and the corresponding cognitive domain have consistently resulted in a Chi-Square value greater than 250 (p=.000). This shows that students' views about the nature of physics are significantly related to their views about learning physics. Thus, a naive realist is likely to be a passive learner, and a scientific realist is likely to be a critical learner. In fact, an Odds Ratio analysis between the extreme profiles have consistently revealed that the likelihood of a naive realist being a passive learner or a scientific realist being a critical learner is more than twenty times the likelihood of a naive realist being a critical learner or a scientific realist being a passive learner.

Universality and Intransigence of Student Views

At the end of section 5 of this report, we have concluded that student views expressed on individual VASS items are universal. The same can actually be said about profiles' distribution and properties, and, as we shall see in this last part of the report, about the impact of conventional instruction of lecture and demonstration on these views and profiles.

7. Tenacity of student views and profiles

Researchers have been constantly showing that student views about the nature of science are hardly affected by conventional science instruction, and that, if they do, they mostly get worse and not better after instruction (§6.4). Our results all along this project show that Lebanese students' views about knowing and learning science, like those of their international peers, are not improved by conventional physics instruction of lecture and demonstration.

VASS data was compared on individual items for the 1999 fall and spring samples, as well for the sample of students who, in the 99-00 academic year, took Form P99LB as pretest and Form P20 as posttest (Figure 6). Further comparison was made on the six VASS dimensions as shown in Table X in the 1999 spring term (used as posttest indicator) and the following fall term (used as pretest indicator). All results indicate a significant tenacity of most student views across all six dimensions, and a shift of about 5% in the negative direction, i.e., toward the folk profile, across the four profiles.

Table X

The non-significant shift detected in Table X is primarily due to a decline in student views regarding some issues addressed in the two cognitive dimensions pertaining to reflective thinking and personal relevance. More specifically, following conventional instruction, some students loose interest in physics. Consequently, they shy away from learning this science meaningfully, and they merely get satisfied instead with studying things by rote for the sole purpose of passing course exams.

A little more pronounced shift in the same direction, and for the same reasons, was detected when we compared results across the three secondary school grade levels. While the overall sum of scores remained about the same in the broad scientific domain, it decreased gradually by about four points in the broad cognitive domain across these grades. No significant difference was observed between the group of all secondary school students taken together and the group of all college students taken together (i.e. freshmen and sophomore university students).

We obtained similar results in the U.S. (Halloun & Hestenes, 1998). Actually, all VASS results seem to indicate that, by and large, under conventional instruction, *student views about knowing and learning science are determined by the time they get to secondary school,* and that these views are *hardly modified afterwards, at least up to the time they finish their sophomore studies in university*.

VASS was also administered to a sample of 46 fourth year students at the Faculty of Education whose specialty includes teaching science and mathematics in primary school. These students' overall average score on VASS was 92.15 (S.D.=9.43), i.e., about one standard deviation above the mean of 84.2 shown in Table X. This small sample of senior university students does not warrant any conclusion. However, it seems to indicate that student views about knowing and learning science might change significantly after advanced university studies. Further research is though needed to corroborate this assertion.

8. Uniformity of student views across demographic strata

VASS data were analyzed by education sector, county of residence and gender. Given the limited number of participating universities, the effect of these demographic variables will be analyzed only at the secondary school level in this section. Corresponding data are given in Table XI.

Table XI

This table shows that results appear to be slightly more affected by school sector than by school geographic location. Private schools appear to fair better than public schools on VASS as shown in this table, and schools in Beirut and the Mountain appear to fair best (in terms of both means and standard deviation, S.D.), followed by schools in North Lebanon then by those in the South.

One should interpret these results with caution, especially that the disparity between the two sectors (and among counties) is mainly caused by two schools, both located in North Lebanon, whose results are at the opposite ends of the spectrum. One is a public school whose students did the worst both in terms of the mean and the standard deviation (Mean=78.33; S.D.=10.97). The other is a private school whose students did the best in both respects (Mean=87.59; S.D.=6.09). Immediately following the former school is a public school in the South with a mean of 80.53 (S.D.=8.24), and immediately preceding the latter is a private school in the Mountain with a mean of 84.92 (S.D.=8.40).

Therefore, one can conclude that, despite what is shown in Table XI, VASS results may actually not be significantly affected by school sector or location.

Up until Form P20 was devised in the spring of 2000, respondents' gender was not accounted for in VASS answer sheet. Hence, the gender comparison shown in Table XI is made with respect to whether a participating school is one for boys, for girls or mixed. Boys' schools account for 10% of respondents, girls' schools for 35%, and mixed schools for the remaining 55%. Therefore, given the disparity in sample size, and notwithstanding the fact that mixed schools possibly provide a better learning environment than monosexual schools, we prefer to hold judgment, at the time being, regarding gender results shown in the table in question. Such a judgment will be better made when Form P20 results are tabulated and analyzed. We only note here that Form P12 results had shown no gender differences at the high school level in the U.S., and minor differences in favor of female students, at the university level (Halloun, 1996).

9. Views about science and course achievement

Interest in assessing student views about science has emerged within the educational community because of the suspicion that these views might affect students' achievement in various science courses. To this end, we assessed how respondents' views expressed in VASS relate to achievement in their respective physics courses.

To get an achievement index, we relied on actual final course grades whenever they were made available to us by participating schools, and on respondents' answers on questions 1, 2, and 3 in VASS, otherwise (Appendix). We followed this approach because we were able to get final course grades for only 333 participating students (41% of all students) in the 1999 spring term. In fact, we had the three VASS questions included in the new forms because we originally anticipated schools' reluctance to provide us with their students' grades.

Students' answers on the three VASS questions significantly correlated with actual course grades where provided (Pearson correlation coefficient = .441, .592, and .697 respectively, all p's=.000). The high correlation made us confident about respondents' honesty and reliability in making a self-evaluation of their actual performance in their physics courses as expressed by their answers on the three VASS items in question. Besides, these items may provide a more normalized indicator of course achievement than actual course grades, given the wide array of assessment schemes followed in various Lebanese schools.

In order to assess how student views correlate with course achievement, we explored four complementary VASS indices: (a) sum of scores on each of the six dimensions distinguished in Table I, (b) sum of scores in the two broad scientific and cognitive domains, (c) overall sum of scores, and (d) the profiles distinguished in §6. Statistical tests included Pearson correlation, linear stepwise regression analysis, and/or ANOVA between each of the four VASS indices and each of the course achievement indices (i.e., actual course grade and VASS items 1, 2, 3).

Pearson correlation analysis showed the following:

- ♦ Overall views expressed in the broad cognitive domain correlate better with course achievement, whether measured with actual course grades or with anyone of the first three VASS question, than overall views expressed in the broad scientific domain.
- ♦ Within scientific dimensions, views expressed in the structure dimension correlate the best with course achievement, and views expressed in validity correlate the worst. The validity dimension is in fact the only dimension that does not correlate significantly with

any course achievement index.

- Within cognitive dimensions, views expressed in the personal relevance dimension correlate the highest with course achievement when measured with any of the first three VASS questions. Learnability comes ahead when correlation is made with final course grades. These two dimensions are the only ones that relate to *attitudes* toward science education in VASS. It should thus be expected to find that course achievement depends most on how much students are motivated to learn science and interested in doing so, and on whether the locus of control is intrinsically or extrinsically oriented.

The same results held in a linear stepwise regression analysis done between the six VASS dimensions and the various achievement indices. Validity constantly showed a negative loading on achievement, and the remaining five dimensions loaded in the following order: personal relevance, then reflective thinking, then learnability, then structure, then methodology (F=14.353; p=.000).

Pearson correlation coefficient between a simple VASS sum of scores and achievement varied between .29 and .35 (p=.000), depending on the achievement index considered. However, and as one would expect, the coefficient gets higher when measured through stepwise regression. It reaches a highly significant R-value of .49 (p=.000). As far as this author knows, this is the *highest coefficient ever* reported in the literature for an instrument assessing views about science.

Pearson correlation is measured with a simple sum of scores, whereas regression R is measured with a weighted sum of scores (different weights are associated with different dimensions). A weighted sum of scores is more reliable than a simple sum of scores in assessing the relationship between student views about knowing and learning science expressed in VASS, on the one hand, and course achievement, on the other. However, and as we have noted in §6.2, because CAD items in VASS are interval type, correlation or regression analyses are thereby used only to provide us with a rough assessment of the relationship in question. Profile analysis is more reliable than simple or weighted sum of scores in this direction, not only from an ontological point of view, but also from a statistical perspective as can be seen in Figure 14 and Table XII.

Figure 14, *then* Table XII

Figure 14 gives the distribution with 95% confidence intervals of average final scores in physics courses across the four profiles distinguished in §6.2. It shows, along with Table XII, that student views about science and achievement in physics courses go hand in hand. The closer a student profile is to being an expert profile, the better her/his final course grade (ANOVA F = 12.63, p=.000), and vice versa. Course grades go down, on the average, as we move down from expert to folk profiles. This relationship between profiles and course achievement is even more distinctively pronounced when achievement is assessed through the first three VASS questions as indicated in Table XII. Students' answers to these questions provide achievement indices that are better normalized than actual course grades, given the disparity of grading schemes followed in different Lebanese schools. A more normalized index can yet be provided by standardized instruments like the Mechanics Diagnostic Test (Halloun, 2001b; Halloun & Hestenes, 1985a) or its equivalent, the Force Concept Inventory (http://modeling.asu.edu). We have just begun administering concurrently VASS and standardized physics instruments, including those just mentioned, to samples of students in Lebanon and abroad. Results will be published and analyzed elsewhere (Halloun et al., 2001).

10. Teacher practice and student views

A major objective in the current project was to assess the impact of teaching practice on student views about science. To this end, we devised a *Teacher Survey* presented in the appendix. This new complementary survey consists of 13 questions aimed at ascertaining factors in teaching practice that could affect student views about science.

The survey is by no means a comprehensive survey. It was meant to assess, in a snapshot, instructional aspects that bear directly on VASS dimensions, and this, in an efficient and friendly way that would require little effort and time, and that would not put off teachers, especially those who prefer not to disclose much information about their teaching practice. Participating teachers found our survey adequate enough for our purpose, and no one added anything to it when asked to provide us with additional information they could deem necessary to better reflect what they actually do in the classroom about student views on knowing and learning science.

In the 1999 spring term, 24 teachers whose students participated in our project filled out this survey. Our analysis of teaching practice was then done at two levels; first at the level of individual issues assessed in the teacher survey, then at the level of a broad perspective given by the entire survey.

At first, we computed means of scores obtained by the students of every participating teacher on each of the six VASS dimensions distinguished in Table I, and on the entire VASS instrument, in the manner shown in Table X. We then compared student VASS means to the respective teacher answers on the teacher survey in two respects: (a) student means on every VASS dimension versus teacher answers on the corresponding item in the teacher survey (e.g., student means on the reflective thinking dimension versus teacher answers on items 1 through 5 in the teacher survey, or student means on the personal relevance dimension versus teacher answers on item 8 in the teacher survey), and (b) student means on the entire VASS instrument versus teacher answers on every single item of the teacher survey.

Analysis of variance revealed that none of the computed student VASS means was significantly affected by whatever a teacher's answer was on any given item in the teacher survey. In other words, participating students' views about knowing and learning science were not significantly affected by whatever their teachers claimed they do in the classroom with respect to any of the thirteen issues addressed in the teacher survey.

We then computed a teacher practice index as a sum of scores on the thirteen items in the teacher survey. The answer to each item was given a maximum score of 5 corresponding to choice A, and a minimum score of 1 corresponding to choice E. Hence, the teacher practice index can have a maximum value of 65 indicating that a teacher is claiming to do everything we asked for in our survey, and a minimum value of 13 indicating just the opposite. Participating teachers' index ranged from 36 to 65, with a mean of 49, and a standard deviation of 9. This index was then correlated with students' mean score on VASS corresponding to every teacher, and we obtained a non-significant Pearson correlation coefficient of .21 ($p=.34$). Figure 15 shows the respective scatter plot that clearly shows the absence of any relationship between teaching practice and student views expressed in VASS.

Figure 15

The results of our *Teacher Survey* need to be interpreted very cautiously. In a first glance, one may get impressed by the practice of our participating teachers as reflected by their answers on the survey. A closer scrutiny reveals that these teachers all followed some form of *conventional* teaching practice of lecture and demonstration, and that their answers on the survey reflect what *they tell* their students to do and not what they actually engage their students in doing inside the classroom.

For example, these colleagues strive to *point out* student misconceptions in class and provide the good alternatives. They would have done better guiding students in a reflective process whereby individual learners would evaluate their own conceptions, on their own, in order to *detect and self-regulate* any incommensurability with science that could emerge in the process. Our colleagues did apparently their best in *telling* students explicitly what physics and science is all about, and how they should go about studying physics and doing traditional laboratory experiments. Students were then treated more like regimented soldiers expected to follow instructions to the letter (often blindly) than like actual stakeholders in the learning process who need to actively participate in contriving every step of the way (Halloun, 2001b).

Students' views about knowing and learning science are hardly affected by such or any other conventional practice of tell and show, no matter how explicitly teachers spell out their guts. This is actually what educational research has long been showing, and this is what we hope that teachers and administrators would eventually realize and account for when reflecting back on their educational practice!

11. Conclusion

Major findings reported in this document can be summed up as follows:

- Views About Science Survey (VASS) is a valid and reliable paper-and-pencil instrument for assessing student views about knowing and learning science.
- Contrasting Alternatives Design (CAD) is an efficient and reliable rating scale assessment format, and it is more so than other rating scales like Likert.
- Student views expressed in any VASS item can be classified into three categories, expert, mixed or folk. Expert views are typical of those expressed by scientists, philosophers of science and experienced science teachers. Folk views are commonly expressed by the lay community in disagreement with experts. Mixed views fall in between.
- No secondary school or college student expresses systematic views throughout VASS. Virtually every student expresses an admixture of the three view types over the entire instrument. More than two-third of respondents, though, tend to express mixed or expert views rather than folk views in virtually any item.
- Response pattern is such that students of any level can be grouped into four categories of distinctive profiles. These are the expert profile (EP), the high transitional profile (HTP), the low transitional profile (LTP), and the folk profile (FP).
- Each of the four profiles can be distinguished quantitatively by an overall sum of scores

falling within a specific range of values. The average sum of scores falls, overall, in the HTP range.

- No profile is characterized by exclusively one type of view or another (expert, mixed, folk). However, a gradual shift from folk views to expert views takes place as one evolves from FP to EP, and a larger proportion of students tend to express mixed views within HTP and LTP than within FP and EP.
- There is a bigger chance, on any VASS item, for an EP person than any other to express an expert view, and for an FP person to express a folk view. People with LTP, then those with HTP, are more likely than others to express a mixed view.
- People with folk profile are most likely to be naïve realists and passive learners, while people with expert profile are most likely to be scientific realists and critical learners. The transition from one extreme to the other takes place gradually across the two transitional profiles.
- Student profiles do not evolve following conventional instruction of lecture and demonstration. This holds irrespective of how explicitly teachers might explain the nature of science or spell out specific guidelines for studying physics in conventional settings. If anything, student profiles are more likely to shift slightly in the FP direction than in the EP direction.
- Student views about knowing and learning science appear to be determined by the time they get to secondary school, and to remain practically unchanged, under conventional instruction, at least until the time they finish their university sophomore years.
- Course achievement and VASS profiles are significantly related. High achievers are most likely to be EP students, while low achievers are most likely to be FP students. Students with HTP and LTP rank in the same descending order in the middle.
- Demographic factors seem not to have a significant impact on student views and profiles.
- Secondary school and college student responses on individual VASS items and corresponding profiles are universal, and so is the Intransigence of these views and profiles under conventional instruction.

Acknowledgment

Many thanks are extended to all those who have lent a hand during this project: school teachers and university professors who gracefully gave VASS to their students and who offered insightful comments on various stages of the project, school and university administrators who opened their classrooms for us and encouraged our efforts, students who participated in the survey, and graduate students and research assistants at the Faculty of Education at Lebanese University whose unsurpassed zeal and dedication made it all possible.

The work reported in this report is supported in part by a grant from Educational Research Center at Lebanese University.

References

Aikenhead, G. S. (1988). An analysis of four ways of assessing student beliefs about STS topics. *Journal of Research in Science Teaching, 25* (8), 607-629.

Aikenhead, G. S. (1987). High-school graduates' beliefs about science-technology-society. III. Characteristics and limitations of scientific knowledge. *Science Education*, *71* (4), 459-487.

Aikenhead, G. S., Fleming, R. W., and Ryan, A. G. (1987). High-school graduates' beliefs about science-technology-society. I. Methods and issues in monitoring student views. *Science Education, 71* (2), 145-161.

American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy. Project 2061*, New York: Oxford University Press.

American Association for the Advancement of Science. (1990). *Science for All Americans. Project 2061*, New York: Oxford University Press.

Arons, A.B. (1984). Student patterns of thinking and reasoning. Parts II & III. *The Physics Teacher, 22,* 21-26 & 88-93.

Baker, D.R, & Piburn, M. (1991). Process skills acquisition, cognitive growth, and attitude change of ninth grade students in a scientific literacy course, *Journal of Research in Science Teaching, 28* (5), 423-436.

Barrington, B. L. & Hendricks, B. (1988). Attitudes toward science and science knowledge of intellectually gifted and average students in third, seventh, and eleventh grades. *Journal of Research in Science Teaching, 25* (8), 679-687.

Bernard, C. (1865). *Introduction à l'Étude de la Médecine Expérimentale.* Paris: Flammarion.

Bunge, M. (1973). *The Methodological Unity of Science.* Dordrecht: D. Reidel.

Changeux, J. P. (1983). *L'homme Neuronal.* Paris: Fayard.

Cobern, W. W. (1993). College students' conceptualizations of nature: An interpretive world view analysis. *Journal of Research in Science Teaching, 30* (8), 935-951.

Cooley, W. W. & Klopfer, L. E. (1961). *Test on Understanding Science.* Educational Testing Service, Princeton.

Dewey, J. (1933). *How We Think. A Restatement of the Relation of Reflective Thinking to the Educative Process.* Boston, MA: D.C. Heath & Co.

Ebenezer, J. V. and Zoller, U. (1993). The No Change in Junior Secondary Students' Attitudes Toward Science in a Period of Curriculum Change: A Probe into the Case of British Columbia. *School Science and Mathematics, 93* (2), 96-102.

Edmondson, K. M. & Novak, J. D. (1993). The interplay of scientific epistemological views, learning strategies, and attitudes of college students. *Journal of Research in Science Teaching, 30* (6), 547-559.

Ericsson, K.A. & Charness, N. (1994). Expert Performance. Its Structure and Acquisition. *American Psychologist, 49* (8), 725-747.

Fleming, R.W. (1987). High-school graduates' beliefs about science-technology-society. II. The interaction among science, technology and society. *Science Education*, *71* (2), 163-187.

Gardner, H. (1985). *The Mind's New Science. A History of the Cognitive Revolution.* New York, NY: Basic Books.

Germann, P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching, 25* (8), 689-703.

Giere, R.N. (1988). *Explaining Science: A Cognitive Approach.* Chicago, IL: University of Chicago Press.

Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching, 28* (1), 73-79.

Gilovich, T. (1991). How We Know What Isn't So. New York: The Free Press, Macmillan.

Ginsburg, H. & Opper, S. (1979). *Piaget's Theory of Intellectual Development.* Englewood Cliffs, NJ: Prentice Hall.

Glass, A.L., Holyoak, K.J., & Santa, J.L. (1979). *Cognition.* MA: Addison Wesley.

Grossberg, S. (1982). *Studies of Mind and Brain.* BSPS. Dordrecht: D. Reidel.

Halloun, I. (2001a). *Contrasting Alternatives: An alternative assessment rating scale*. In preparation.

Halloun, I. (2001b). *Apprentissage par Modélisation : La Physique Intelligible.* Beirut: Phoenix series / Librairie du Liban.

Halloun, I. (2000). Model-laden inquiry: A prescription for effective physics instruction. *THEMES*, *1* (4), 339-355.

Halloun, I. (1997). Views about science and physics achievement. The VASS story. In E. F. Redish & J. S. Rigden (Eds), *The Changing Role of Physics Departments in Modern Universities. Proceedings of ICUPE.* pp. 605-614. College Park, Maryland: American Institute of Physics Press.

Halloun, I. (1996). *Views about Sciences Survey.* Paper presented at the annual meeting of the National Association for Research in Science Teaching. St. Louis, MI. *ERIC document No. ED394840.*

Halloun, I. (1994). *Assessing Student Views about Physics. How Adequate are Available Instruments?* Invited paper presented at the Annual Meeting of the American Association of Physics Teachers, Notre Dame, IN.

Halloun, I. (1993). *Lebanese public understanding of science.* Junieh: CREST.

Halloun, I. (1988). Compétence en mathématiques des étudiants de physique. *Lebanese Journal of Science & Mathematics Education, 1* (2*),* 21-23.

Halloun, I. (1986). Le réalisme naif et l'apprentissage de la physique. *Recherches Pédagogiques, 17,* 23-47.

Halloun, I, Hestenes, D., & Osborn-Popp, S. (2001). *Validation of the Views abou Science Survey.* In preparation.

Halloun, I, & Hestenes, D. (1998). Interpreting VASS dimensions and profiles, *Science & Education, 7* (6), 553-577.

Halloun, I. & Hestenes, D. (1985a). The initial knowledge state of college physics students. *American Journal of Physics, 53 (11),* 1043-1055.

Halloun, I. & Hestenes, D. (1985b). Common sense concepts about motion. *American Journal of Physics, 53 (11),* 1056-1065.

Hammer, D. (1994). Epistemological Beliefs in Introductory Physics. *Cognition and Instruction, 12* (2), 151-183.

Harré, R. (1959). *The Principles of Scientific Thinking.* Chicago, IL: U. of Chicago Press.

Hestenes, D. (1992). Modeling Games in the Newtonian World. *American Journal of Physics*, *60 (8)*, 732-748.

Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*, 141-158.

Johnson-Laird, P.N. (1983). *Mental Models.* Cambridge: Cambridge University Press.

Jones, E.E. (1986). Interpreting Interpersonal Behavior: The Effects of Expectancies. *Science, 234,* 41-46.

Kimball, M. E. (1968). Understanding the nature of science: A comparison of scientists and science teachers. *Journal of Research in Science Teaching, 5*, 110-120.

Klahr, D. (1976). *Cognition & Instruction.* New York, N.Y.: John Wiley & Sons.

Klopfer, L. E. (1969). Effectiveness and effects of ESSP astronomy materials. *Journal of Research in Science Teaching, 6*, 64-75.

Krynowsky, B.A. (1988). Problems in assessing student attitude in science education: A partial solution. *Science Education, 72* (4), 575-584.

Kuhn, T. (1970). *The Structure of Scientific Revolutions.* 2nd Ed. Chicago: University of Chicago Press.

Lakatos, I., Musgrave, A. (1974). *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press.

Lakoff, G. (1986). *Women, Fire and Dangerous Things: What Categories Reveal About the Mind.* Chicago: University of Chicago Press.

Lecourt, D. (1974). *Bachelard: Epistémologie.* Paris: Presses Universitaires de France.

Lederman, N. G. (1992). Students' and Teachers' Conceptions of the Nature of Science: A Review of the Research. *Journal of Research in Science Teaching, 29* (4), 331–359.

Lederman, N. G. and O'Malley, M. (1990). Students' perception of tentativeness in science: Development, use, and sources of change. *Science Education, 74* (2), 225-239.

Lochhead, J. & Clement, J., Eds. (1979). *Cognitive Process Instruction.* Philadelphi, PA: The Franklin Institute Press.

Mackay, L. D. (1971). Development of Understanding about the Nature of Science. *Journal of Research in Science Teaching, 8* (1), 57-66.

Margolis, H. (1987). *Patterns, Thinking, and Cognition. A Theory of Judgment.* Chicago, IL: U. of Chicago Press.

McComas, W.F., Ed. (1998). *The Nature of Science in Science Education.* Dordrecht: Kluwer.

Meichtry, Y. J. (1992). Influencing Student Understanding of the Nature of Science: Data from a Case of Curriculum Development. *Journal of Research in Science Teaching, 29* (4), 389–407.

Meichtry, Y. J. (1993). The impact of science curricula on student views about the nature of science. *Journal of Research in Science Teaching, 30* (5), 429-443.

Munby, H. (1983). Thirty studies involving the "Scientific Attitude Inventory": What confidence can we have in this instrument? *Journal of Research in Science Teaching, 20* (2), 141-162.

The National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform.* Washington, DC: U.S. Government Printing Office.

National Resesarch Council. (1996). *National Science Education Standards*. Washington, DC: National Academy Press.

National Science Board. (1996). *Science & Engineering Indicators-1996.* Washington, DC: U.S. Government Printing Office.

National Science Board. (1993). *Science & Engineering Indicators-1993.* Washington, DC: U.S. Government Printing Office.

National Science Teachers Association. (1995). *Scope, Sequence, and Coordination of Secondary School Science. Volume 3. A High School Framework for National Science Education Standards.* Washington, DC: NSTA.

National Science Teachers Association. (1993). *Scope, Sequence, and Coordination of Secondary School Science. Volume 1. The Content Core.* Washington, DC: NSTA.

Newell, A. & Simon, H.A. (1972). *Human Problem Solving.* Englewood Cliffs, N.J.: Prentice-Hall.

Perry, W. G. (1970). *Forms of Intellectual and Ethical Development in the College Years.* Holt, Rinehart and Winston. New York.

Piaget, J. (1972). *The Psychology of Intelligence.* Totowa, N.J.: Littlefield Adams.

Popper, K.R. (1983). *Realism and the Aim of Science.* Totowa: Rowman & Littlefield.

Redish, F. E., & Saul, J. (1995, January). *The Distribution of Student Expectations and Attitudes in Introductory University Physics.* Paper presented at the AAPT Winter Meeting, Orlando, Florida.

Reif, F. (1987). Instructional design, cognition, and technology: Applications to the teaching of scientific concepts. *Journal of Research in Science Teaching*, *24* (4), 309-324

Reif, F. & Larkin, J.H. (1991). Cognition in scientific and everyday domains: Comparison and learning implications, *Journal of Research in Science Teaching, 28* (9), 733-760.

Rennie, L. J., & Parker, L. H. (1987). Scale dimensionality and population heterogeneity: Potential problems in the interpretation of attitude data. *Journal of Research in Science Teaching, 24* (6), 567-577.

Resnick, L. B., Ed. (1989). *Knowing, Learning, and Instruction.* Hillsdale, N.J.: Lawrence Erlbaum.

Roth, W. M., & Roychoudhury, A. (1993). The nature of scientific language, knowing and learning: the perspectives of four physics students. *International Journal of Science Education, 15* (1), 27-44.

Rubba, P.A., & Andersen, H.O. (1978). Development of an Instrument to Assess Secondary School Students' Understanding of the Nature of Scientific Knowledge. *Science Education, 62* (4), 449-458.

Ryan, A.G. (1987). High-school graduates' beliefs about science-technology-society. IV. The characteristics of scientists. *Science Education*, *71* (4), 489-510.

Schibeci, R.A, & Riley, J.P. (1986). Influence of students' background and perceptions on science attitudes and achievement, *Journal of Research in Science Teaching, 23* (3), 177-187.

Schmidt, D.J. (1967). Test On Understanding Science: A comparison among several groups. *Journal of Research in Science Teaching, 5*, 365-366.

Simon, H.A. (1979). *Models of Thought.* New Haven, MA: Yale U. Press.

Simpson, R. D. & Oliver, J. S. (1985). Attitude toward science and achievement motivation profiles of male and female science students in grades six through ten. *Science Education, 69* (4), 511-526.

Songer, N. B., & Linn, M. C. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28* (9), 761-784.

Squire, L.R. (1986). Mechanisms of Memory. *Science, 232,* 1612-1619.

Stewart, J. (1985). Cognitive science and science education. *European Journal of Science Education, 7* (1), 1-17.

Symington, D. & Spurling, H. (1990). The 'Draw a Scientist Test': interpreting the data. *Research in Science & Technology Education, 8* (1), 75-77.

TIMSS. (1994). IEA *Third International Mathematics and Science Study. http://wwwcsteep.bc.edu/timss.*

Tobias, S. (1990). *They're Not Dumb, They're Different. Stalking the Second Tier*. Tucson, AZ: Research Corporation.

Tobin, K. Ed. (1993). *The Practice of Constructivism in Science Education.* Hillsdale, NJ: Lawrence Erlbaum.

Ullmo, J. (1969). *La Pensée Scientifique Moderne.* Paris: Flammarion.

UNESCO. (1993). *Projet 2000+. Forum international sur la culture scientifique et technologique pour tous.* Paris: UNESCO.

Welch, W. W. and Pella, M. O. (1967). The development of an instrument for inventorying knowledge of the processes of science. *Journal of Research in Science Teaching, 5*, 64-68.

تحديد مناهج التعليم العام ما قبل الجامعي وأهدافها.(1997). مرسوم رقم 10227 وتعميم رقم 24/م/97.

Appendix

VASS Form P20 and Answer Sheet

Teacher Survey

CAD vs. Likert Comparison Instrument

Views About Science Survey

Form P 20 – Student Survey

A copy of this instrument is provided separately

(6 pages)

Views About Science Survey

Form P 20 – Teacher Survey

Thank you for taking this survey which is intended to identify factors that affect people's understanding of physics, and to assist in the design of instructional material.

The survey is designed by Prof. Ibrahim A. Halloun in collaboration with Lebanese and U.S. researchers. For any information, please call Prof. Halloun at: 01-680382, or visit his web page at: http://www.inco.com.lb/halloun.

All data are ***confidential****. Your identity will not be disclosed to any party.*

Class(es): ____ Date: ______________

Please answer each of the following questions by choosing one of the following alternatives:

A: More than once a week; **B**: About once a week; **C**: About once a month; **D**: Seldom; **E**: Never

1. How often do you discuss with your students how they should go about using their physics textbook for study?	A	B	C	D	E
2. How often do you discuss with your students how they should go about solving homework problems on their own?	A	B	C	D	E
3. How often do you discuss in class misconceptions that students typically have about real world systems and phenomena?	A	B	C	D	E
4. How often do you discuss in class mistakes that students make in their homework?	A	B	C	D	E
5. How often do you discuss in class mistakes that students make in their exams?	A	B	C	D	E
6. How often do you get students engaged in experimental activities at school?	A	B	C	D	E
7. How often do you asssign experiments or other practical activities for students to do at home?	A	B	C	D	E
8. How often do you discuss with your students the applications of physics in everyday life?	A	B	C	D	E
9. How often do you discuss with your students the relation of physics to other scientific disciplines?	A	B	C	D	E
10. How often do you discuss with your students the relation of physics to technology?	A	B	C	D	E
11. How often do you discuss with your students the nature of scientific laws?	A	B	C	D	E
12. How often do you discuss with your students the nature of scientific thinking?	A	B	C	D	E
13. How often do you discuss with your students the role of mathematics in physics?	A	B	C	D	E

ID: ________________ Class: ________________ Gender: ☐ Male ☐ Female

Each of the following 12 questions consists of a statement about a given issue. Please tell us what is your opinion about the statement by choosing ***only one*** of the following five alternatives:

1	2	3	4	5
I strongly agree	I agree	I do not know	I disagree	I strongly disagree

1. Learning physics requires a serious effort.	1	2	3	4	5
2. Reasoning skills that are taught in physics courses can be helpful to me in my everyday life.	1	2	3	4	5
3. For me, doing well in physics courses depends on how much effort I put into studying.	1	2	3	4	5
4. The first thing I do when solving a physics problem is to represent the situation with sketches and drawings.	1	2	3	4	5
5. After the teacher solves a physics problem for which I got a wrong solution I discard my solution and learn the one presented by the teacher.	1	2	3	4	5
6. To me, physics is important as a source of factual information about the natural world.	1	2	3	4	5
7. Learning physics requires a special talent.	1	2	3	4	5
8. Reasoning skills that are taught in physics courses can be helpful to me if I were to become a scientist.	1	2	3	4	5
9. For me, doing well in physics courses depends on how well the teacher explains things in class.	1	2	3	4	5
10. The first thing I do when solving a physics problem is to search for formulas that relate givens to unknowns.	1	2	3	4	5
11. After the teacher solves a physics problem for which I got a wrong solution I try to figure out how the teacher's solution differs from mine.	1	2	3	4	5
12. To me, physics is important as a source of ways of thinking about the natural world.	1	2	3	4	5

ID: ________________ Class: ________________ Gender: ☐ Male ☐ Female

Each of the following 6 questions consists of two statements about a given issue, followed by five contrasting alternatives regarding the two statements. Please answer each question by choosing ***only one*** of the corresponding five alternatives. The example below describes the five choices for question 1.

Example

Learning physics requires:
(a) a serious effort.
(b) a special talent.

1. Mostly (a), rarely (b) **2**. More (a) than (b)
3. Equally (a) & (b)
4. More (b) than (a) **5**. Mostly (b), rarely (a)

What would each one of the five choices mean?

1. Mostly (a), rarely (b):	Learning physics requires ***mostly*** *a serious effort* and ***rarely*** a special talent (or *mainly* the former and *hardly ever* the latter).
2. More (a) than (b):	Learning physics requires ***more*** *a serious effort* than a special talent.
3. Equally (a) & (b):	Learning physics requires ***as much*** a serious effort *as* a special talent.
4. More (b) than (a):	Learning physics requires ***more*** a *special talent* than a serious effort.
5. Mostly (b), rarely (a):	Learning physics requires ***mostly*** a *special talent* and ***rarely*** a serious effort (or *mainly* the former and *hardly ever* the latter).

1. Learning physics requires:
   (a) a serious effort.
   (b) a special talent.
   ☐1 ☐2 ☐3 ☐4 ☐5

2. Reasoning skills that are taught in physics courses can be helpful to me:
   (a) in my everyday life.
   (b) if I were to become a scientist.
   ☐1 ☐2 ☐3 ☐4 ☐5

3. For me, doing well in physics courses depends on:
   (a) how much effort I put into studying.
   (b) how well the teacher explains things in class.
   ☐1 ☐2 ☐3 ☐4 ☐5

4. The first thing I do when solving a physics problem is:
   (a) represent the situation with sketches and drawings.
   (b) search for formulas that relate givens to unknowns.
   ☐1 ☐2 ☐3 ☐4 ☐5

5. After the teacher solves a physics problem for which I got a wrong solution:
   (a) I discard my solution and learn the one presented by the teacher.
   (b) I try to figure out how the teacher's solution differs from mine.
   ☐1 ☐2 ☐3 ☐4 ☐5

6. To me, physics is important as a source of:
   (a) factual information about the natural world.
   (b) ways of thinking about the natural world.
   ☐1 ☐2 ☐3 ☐4 ☐5

1	2	3	4	5
← Toward { "Mostly" or "Most often" } (a)		Equally (a) & (b) or (a) as often as (b)		Toward { "Mostly" or "Most often" } (b) →

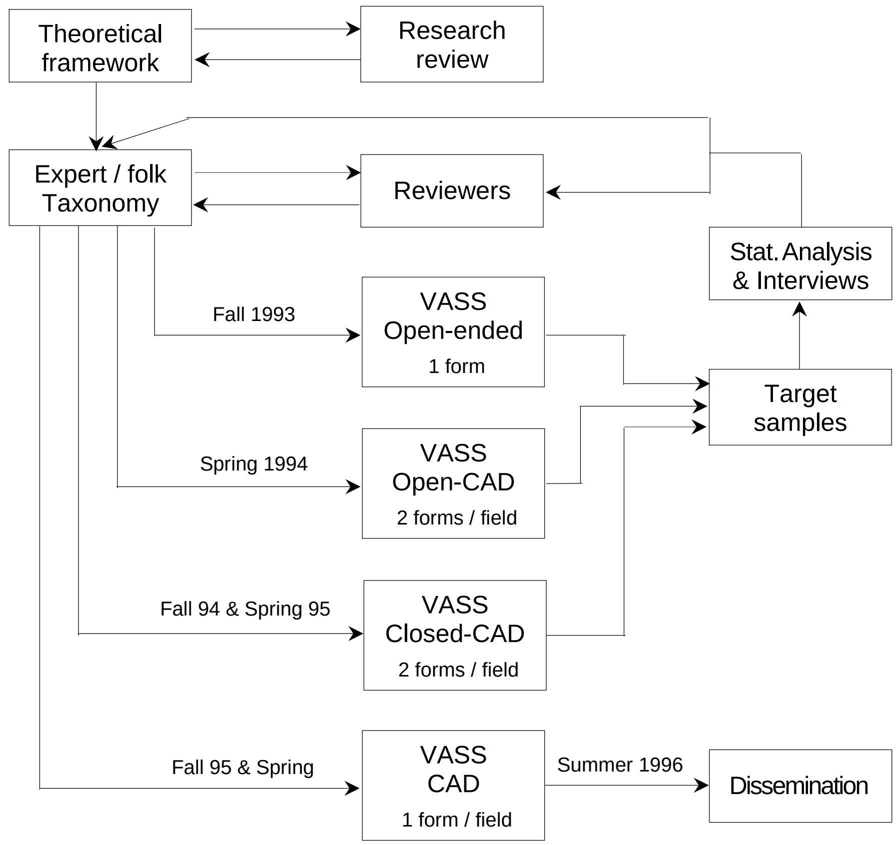


Figure 1: Evolution of VASS.

Learning physics requires:
(a) a serious effort.
(b) a special talent.

What would each one of the eight choices mean?

① Only (a), Never (b):	Learning physics requires ***only*** *a serious effort* and ***no*** special talent *at all*.
② Mostly (a), Rarely (b):	Learning physics requires ***far more*** *a serious effort* than a special talent.
③ More (a) Than (b):	Learning physics requires ***somewhat more*** *a serious effort* than a special talent.
④ Equally (a) & (b):	Learning physics ***equally*** requires ***both*** a serious effort and a special talent.
⑤ More (b) Than (a):	Learning physics requires ***somewhat more*** a *special talent* than a serious effort.
⑥ Mostly (b), Rarely (a):	Learning physics requires ***far more*** a *special talent* than a serious effort.
⑦ Only (b), Never (a):	Learning physics requires ***only*** a *special talent* and ***no*** serious effort *at all*.
⑧ Neither (a) Nor (b):	Learning physics requires ***neither*** a *special talent* ***nor*** *a serious effort*.

Figure 2: A typical item from VASS Form P12 with an 8-point CAD scale (§2).

Science Process Inventory - Form C (Welch & Pella, 1967):

Once a statement of science becomes a law of science, it will not be changed.
(*Agree / Disagree*).

Nature of Scientific Knowledge Scale (Rubba & Andersen, 1978):

Today's scientific laws, theories, and concepts may have to be changed in the face of new evidence. (*Likert Scale*)

VASS – Forms P99LB and P20:

14. Physicists' current ideas about particles that make up the atom:
    (a) will always be maintained as they are.
    (b) may eventually be modified in some respects.

15. Newton's laws of motion:
    (a) will always be used in their present form.
    (b) may eventually be modified in some respects.

Figure 3: A given issue like the falsifiability of science is assessed in specific contexts in VASS rather than in the abstract as commonly done in traditional instruments.

Test Of Understanding of Science (TOUS, Cooley and Klopfer, 1961):

A scientific theory should:

A. provide the final solution to scientific problems.

B. suggest directions for making useful things.

C*. tie together and explain many natural events.

D. suggest good rules for carrying out experiments.

* considered as “best response” by the authors of TOUS.

Germann (1988):

Science makes me feel uncomfortable, restless, irritable, and impatient.

(Likert scale)

Figure 4: Sample questions from traditional instruments addressing many factors in a single question.

I: Describe what you normally do when solving a physics problem. List all steps you often follow, in order.

S: First step in any problem would be to read the problem and list what's given and what you need, variables or what not. And the next step would be to find formulas that include these variables. And then, the third would be to solve for the unknowns. That's basically it.

I: So this would be an algorithm you would work through in any kind of problem?

S: Basically, I would agree. It's a basic general, general outline of how to solve a problem.

I: Do you ever consider drawing some kind of a diagram?

S: Uh-huh... I'd consider that helpful, yeah, I'd probably include that in step one. Draw, label, find out what you have and don't have.

I: So that becomes then, your first step.

S: Uh-huh.

I: Would that be true for any kind of problem?

S: Visualization helps a lot. I would say it would be a good step to try in any problem. If you can't visualize it, I wouldn't try to draw it. Yeah, I would agree that would have to be helpful for any kind of problem.

I: Do you usually do it?

S: Do I do it? Usually yes. It's almost asked of us in physics class: force diagrams, free body diagrams. I would say they're probably most helpful. I would say, yeah.

Figure 5: Excerpt from an interview with a university physics student whose written answer to an open-ended question does not reflect his actual position (Halloun, 2001b).

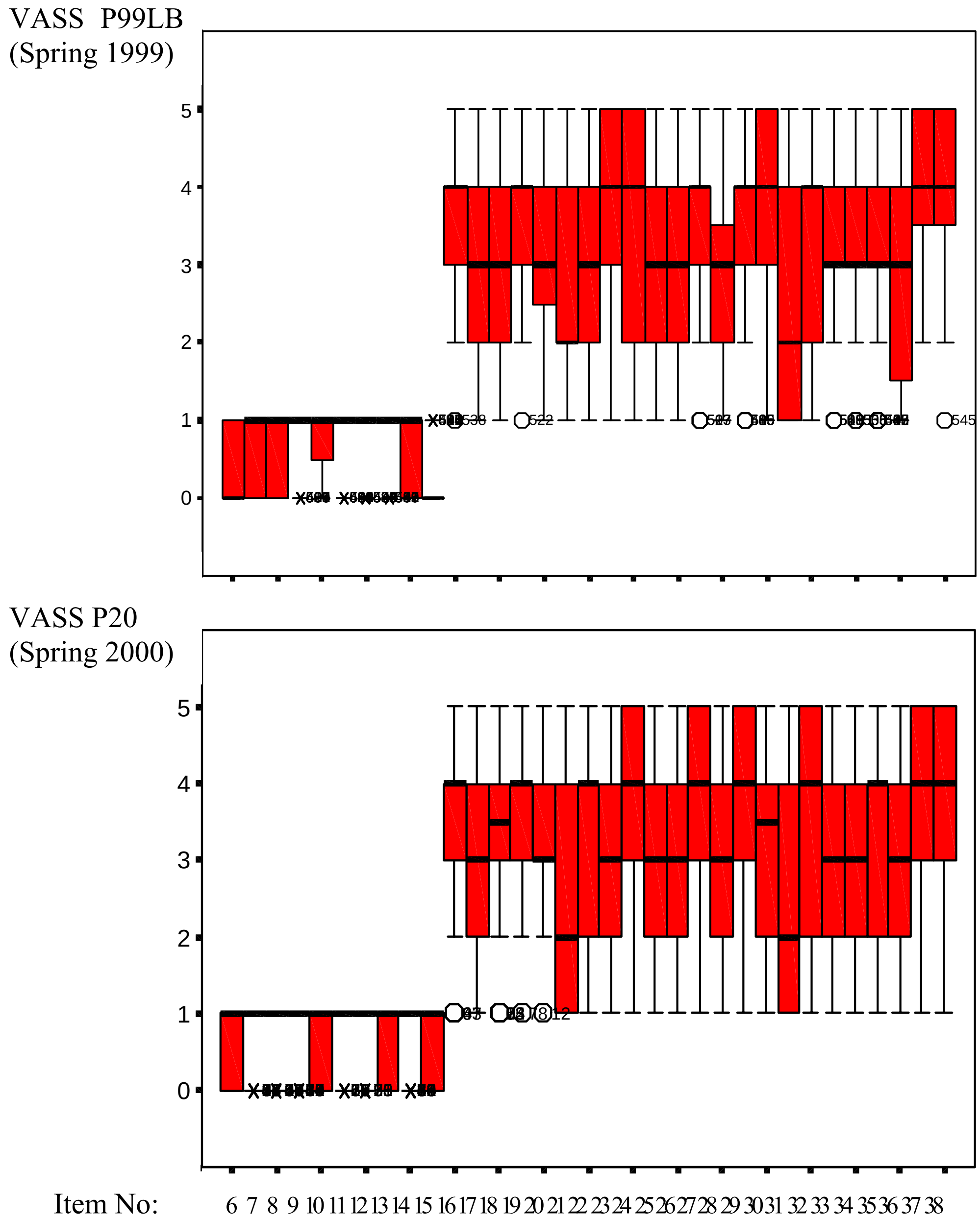


Figure 6: Boxplot comparison of answers provided, on VASS forms P99LB (top) and P20 (bottom), by students enrolled, in two consecutive years, in the first secondary grade at the same school.

Answers are recoded so that the expert choice is always 1 in dichotomous items (6 through 15) and 5 in CAD items (16 through 38).

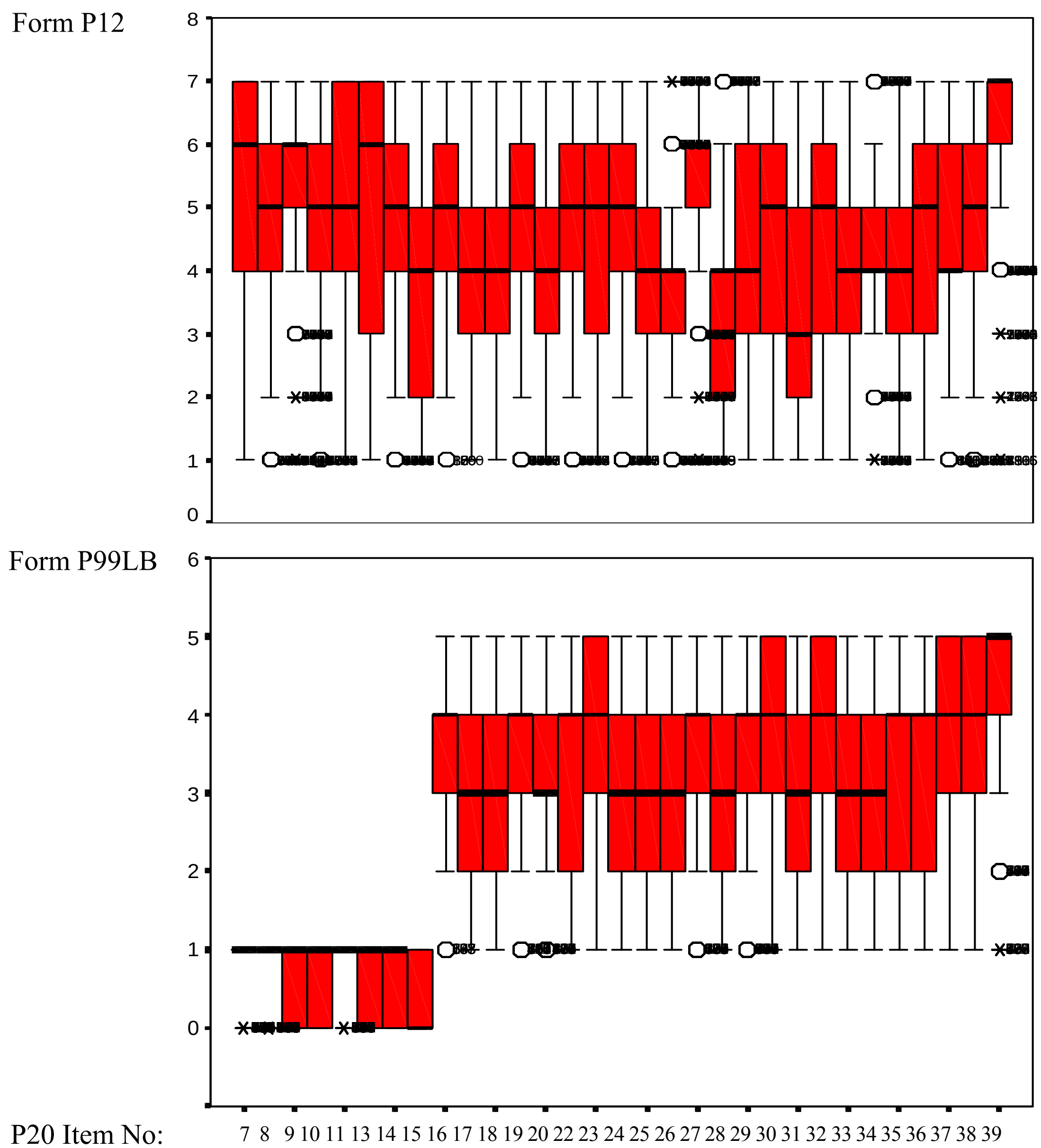


Figure 7: Boxplot comparison of responses provided by 1838 U.S. students and 813 Lebanese students on similar items in VASS Form P12 (Spring 97) and VASS Form P99LB (Spring 99).

Responses are recoded so that the expert end pole is always 7 in Form P12, and 5 (or 1) in Form P99LB.

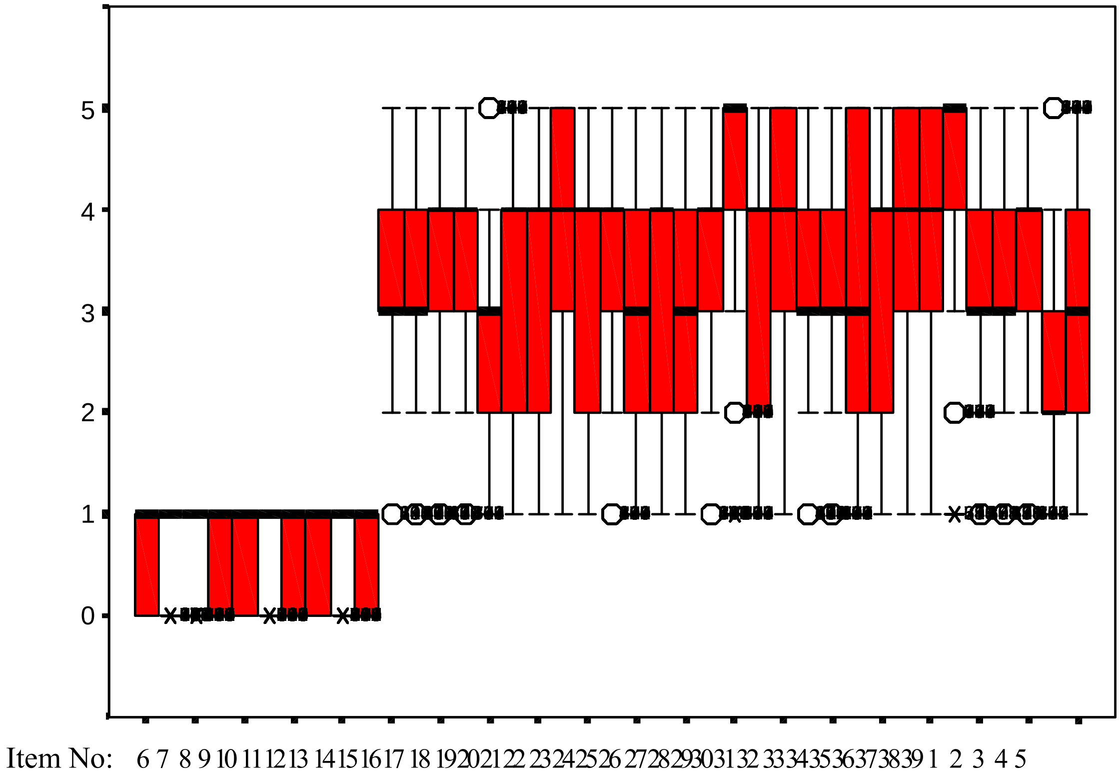


Figure 8: Boxplots of participants' answers on VASS Form P99LB corresponding to Table IV.

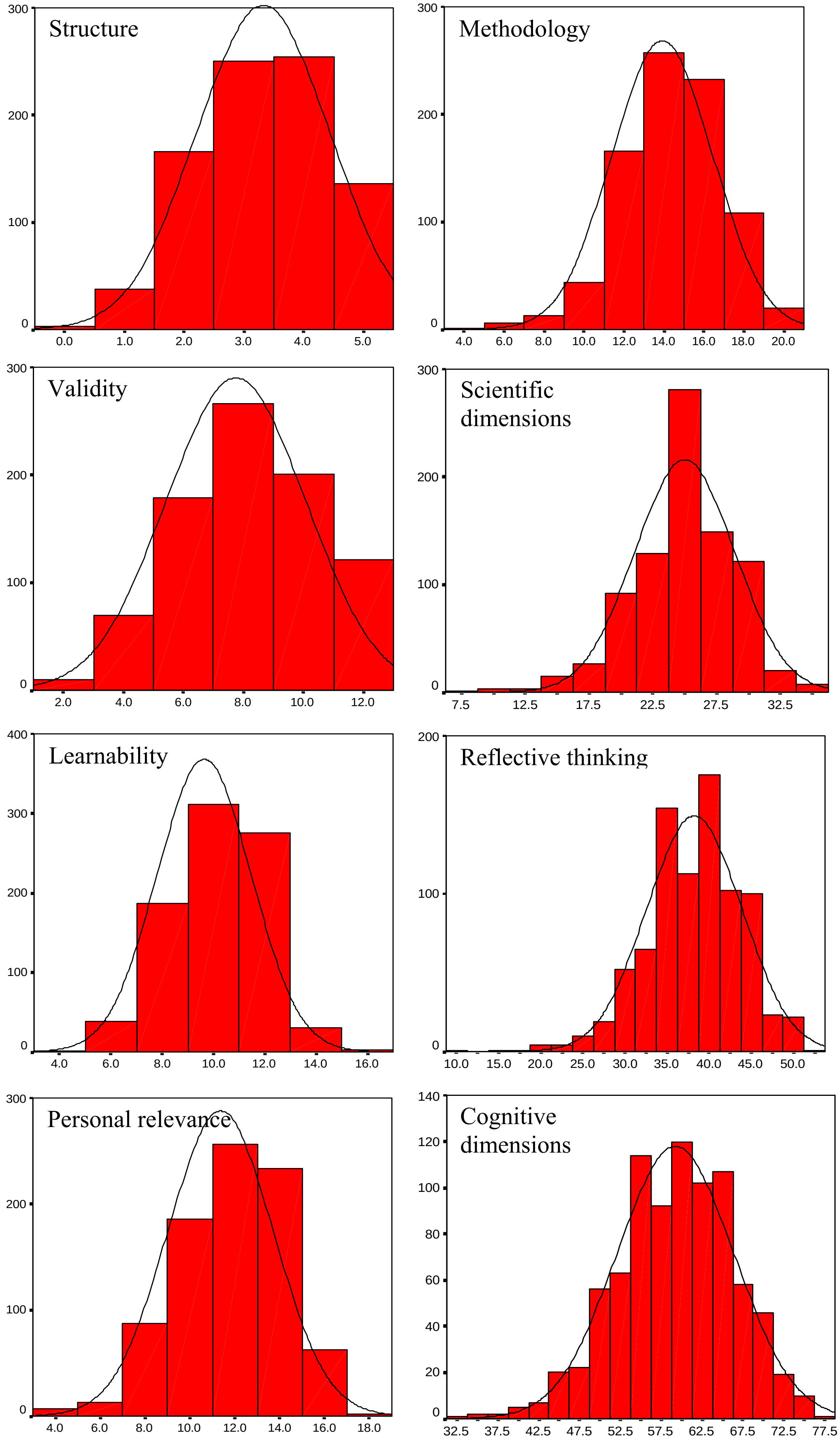


Figure 9: Histograms of respondents' sum of scores on VASS dimensions (Fall 1999).

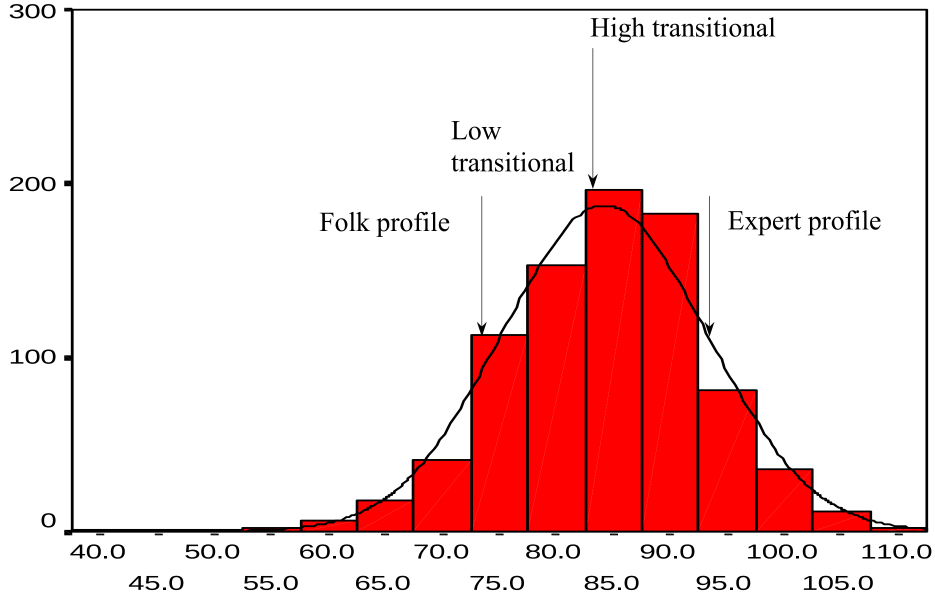


Figure 10: Histogram of respondents' sum of scores on all 33 VASS items along with profile cutoffs. Scores ranged, in the 1999 fall term, from 42 to 111 out of 125 points.

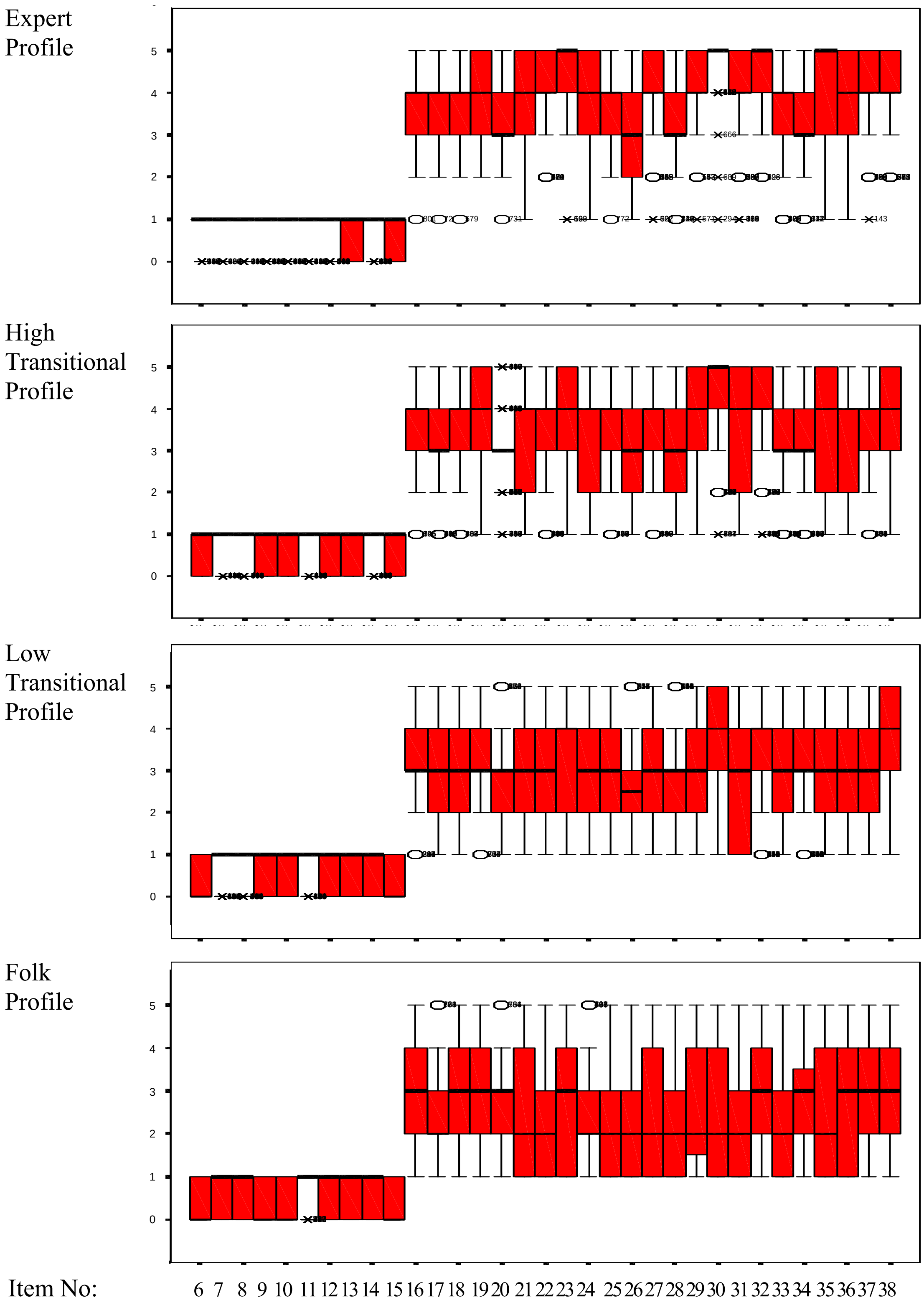


Figure 11: Boxplots comparing the distribution of respondents' answers on the various VASS items within each of the four profiles.

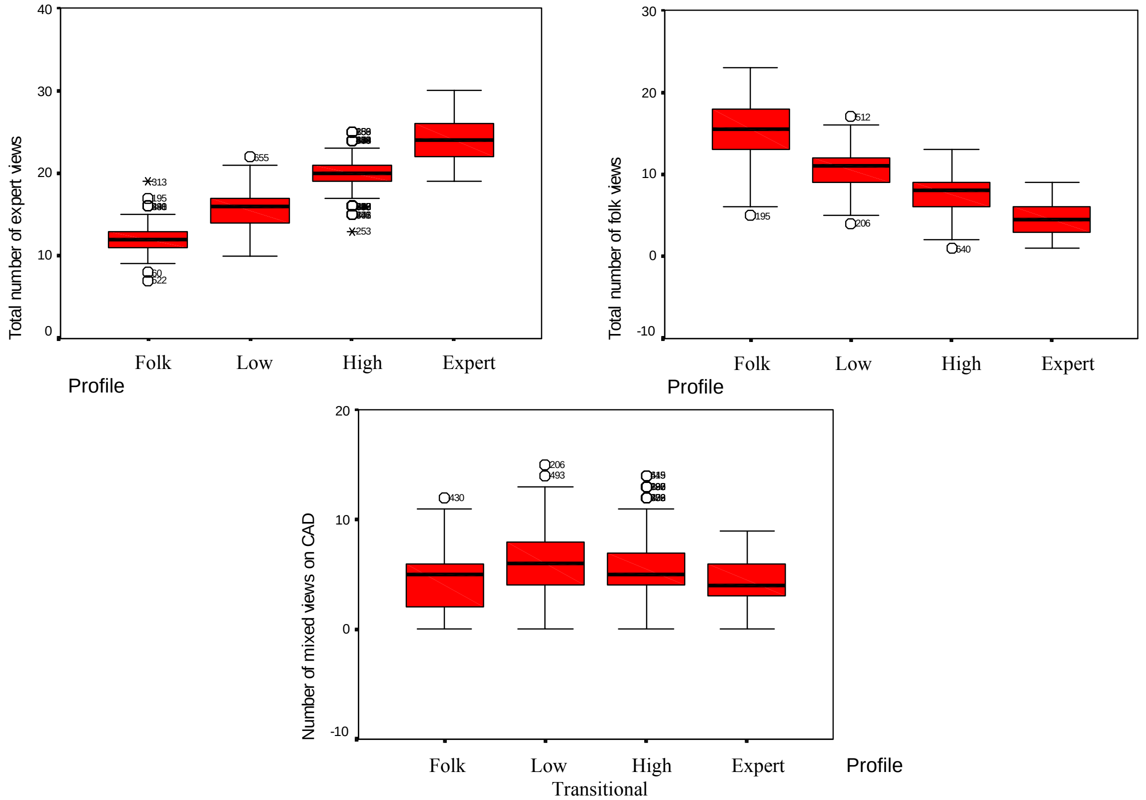
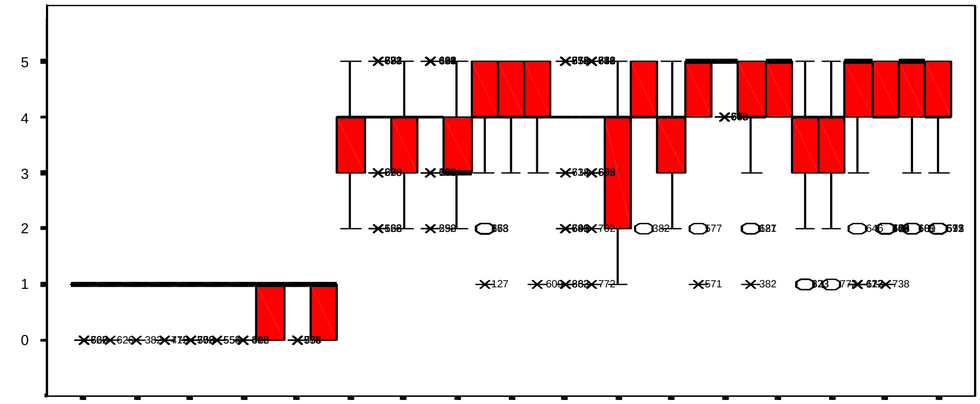


Figure 12: Boxplots comparing the four profiles with respect to the total number of views of a given type expressed by respondents in each profile.

Students with an *expert profile* and expressing expert views on 26 items or more



Students with a *folk profile* and expressing folk views on 17 items or more

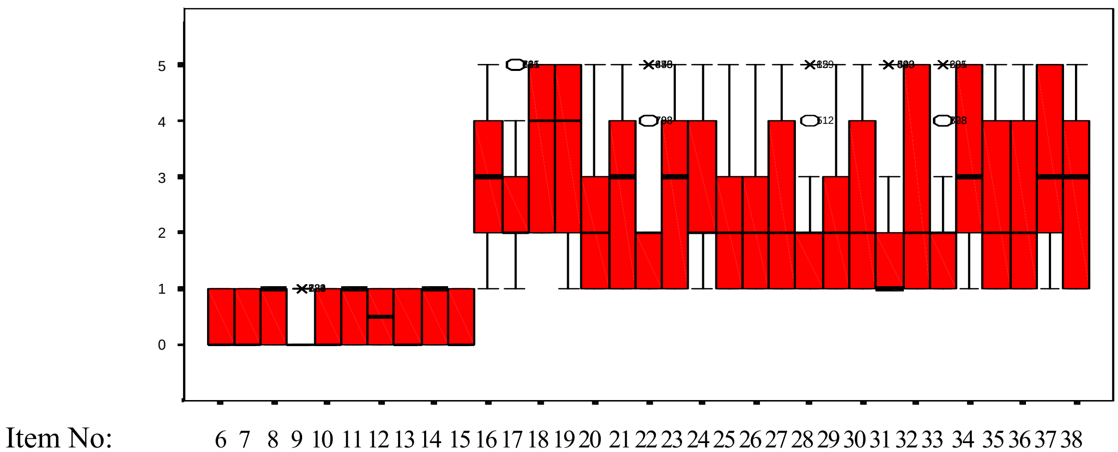


Figure 13: Comparison of VASS answers given by two extreme groups of respondents in the1999 fall term.

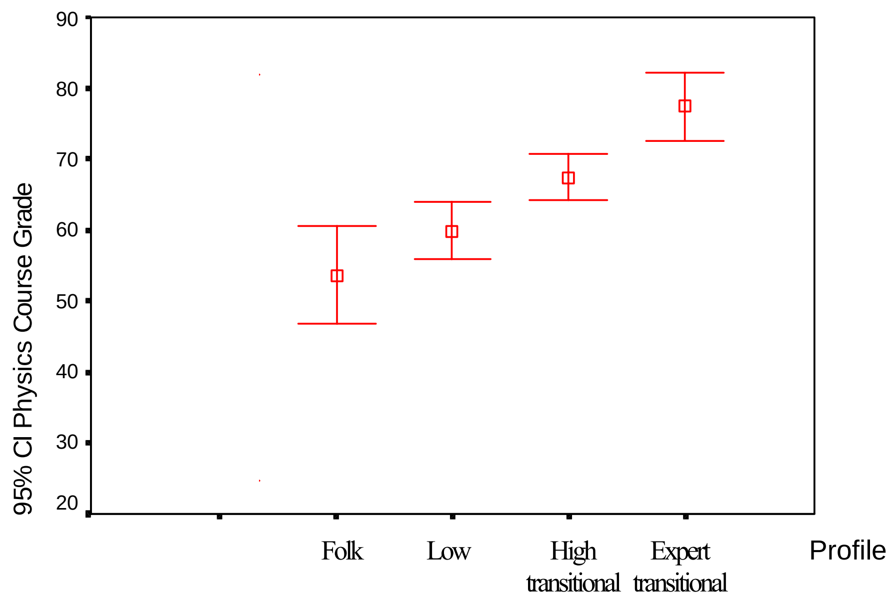


Figure 14: Error bars relating course final grades to profiles with 95% confidence intervals (CI).

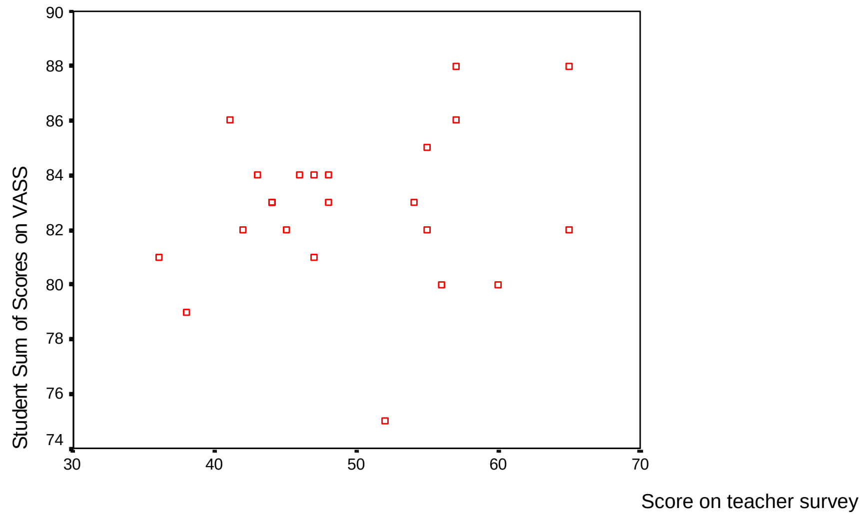


Figure 15: Student VASS scores vs. teacher practice index.

Table I

Taxonomy of VASS – Form P20

	Scientific dimensions	Item Number
1.	***Structure*****:** Science is a *coherent body* of knowledge about *patterns* in nature revealed by *careful investigation* — rather than a loose collection of directly perceived facts.	9, 10, 11, 12, 13
2.	***Methodology*****:** The methods of science are *theory-laden*, *systematic* and *generic* — rather than idiosyncratic and situation specific.	8, 38
	Mathematics is a *tool* used by scientists for describing, connecting and analyzing ideas — rather than a source of factual knowledge.	26, 37
	Mathematical modeling for problem solving involves *more* — than selecting mathematical formulas for number crunching.	26, 28
3.	***Validity*****:** Scientific knowledge is *approximate*, *tentative*, and *refutable* — rather than exact, absolute and final.	14, 15, 35, 36
	Cognitive dimensions	
4.	***Learnability:*** Science is *learnable by anyone* willing to make the effort — not just by a few talented people.	16
	Achievement depends more on *personal effort* and *perseverance* — than on the influence of teacher, peers or textbook.	20, 22
5.	***Reflective thinking:*** For meaningful understanding of science, one needs to: (a) concentrate more on the development of *generic methods* for *construction* and *application* of scientific ideas — than on memorizing facts and procedures;	25, 27, 29, 34, 19, 33
	(b) examine situations in *many ways* — instead of following a single approach;	30
	(c) seek knowledge from a *variety of sources* — instead of relying on a single authority;	4, 5
	(d) continuously *evaluate* one's own work for *consistency* and *effectiveness* — instead of just accumulating new information from presumed authorities;	31, 32 21
	(e) *reconstruct* new subject knowledge in one's own way — instead of memorizing it as given.	23
6.	***Personal relevance*****:** Science is relevant to everyone's life. — It is not of exclusive concern to scientists.	18, 24
	Science should be *enjoyed* and studied more for *personal benefit* — than for fulfilling curriculum requirements.	6, 7, 17

Table II

Comparison of student answers (in percentages) on six VASS items
first asked in Likert format then in CAD format

Answer / Item	1	2	3	4	5	Answer / Item	1	2	3	4	5
Likert 1a	31	52	6	11	0	Likert 4a	32	42	9	15	3
Likert 1b	14	32	18	25	11	Likert 4b	25	53	7	11	3
CAD 1	21	26	38	11	4	CAD 4	17	15	39	19	10
Likert 2a	33	55	7	4	0	Likert 5a	20	23	9	31	16
Likert 2b	35	35	12	14	4	Likert 5b	50	39	6	4	1
CAD 2	24	29	26	19	3	CAD 5	11	10	13	31	35
Likert 3a	30	42	7	19	2	Likert 6a	27	51	12	8	3
Likert 3b	37	42	7	11	3	Likert 6b	28	46	17	4	4
CAD 3	6	21	45	19	8	CAD 6	4	13	51	20	11

Table III

Sample Characteristics
(percentage of participants)

Term (1999)	**Grade Level**			
	1st sec	**2nd sec**	**3rd sec**	**College**
Fall	38	33	12	17
Spring	53	29	9	9

Education Level	
Secondary	**University**
83	17
91	9

Term (1999)	**County**		
	Brt-Mtn	North	South
Fall	36	29	36
Spring	43	44	13

Sector	
Public	**Private**
54	46
35	65

Table IV

Percentages of participants' answers on VASS Form P99LB in the 1999 fall term

Recoded response:

	0	1
Q6	37	63
Q7	18	82
Q8	14	86
Q9	37	63
Q10	35	65
Q11	16	84
Q12	33	67
Q13	42	58
Q14	24	76
Q15	48	52

Recoded response:

	1	2	3	4	5
Q16	2	15	33	40	10
Q17	6	17	36	31	10
Q18	4	19	27	30	20
Q19	2	12	27	33	25
Q20	4	22	52	17	5
Q21	16	15	16	31	21
Q22	12	18	14	39	16
Q23	7	11	15	40	27
Q24	9	23	18	28	21

Recoded response:

	1	2	3	4	5
Q25	6	15	28	37	14
Q26	15	29	31	18	8
Q27	9	18	15	35	24
Q28	10	26	34	19	11
Q29	9	13	15	39	23
Q30	6	8	7	25	54
Q31	20	17	12	27	25
Q32	7	8	15	36	34
Q33	11	14	32	33	10
Q34	9	14	46	21	10
Q35	15	15	19	25	26
Q36	14	20	14	32	19
Q37	5	17	22	30	26
Q38	3	9	19	36	33
Q39	2	4	6	22	66
Q1	2	12	43	39	5
Q2	1	10	44	38	8
Q3	2	7	38	45	9
Q4	16	41	18	17	8
Q5	6	29	21	26	19

Answers have been rearranged so that the expert pole is always scored 1 in dichotomous items and 5 in CAD or multiple-choice items.

Table V

Distribution of respondents' views on VASS items in the 1999 fall term

Number of views (out of 33 items, Q6 through Q38)	Percent of respondents expressing a given view type: on CAD items			on dichotomous items		on Entire test	
↓	Expert	Mixed	Folk	Expert	Folk	Expert	Folk
0		3	1		4		
1		4	3		13		1
2		7	8	0	21		2
3	0	10	11	1	24		4
4	1	15	13	5	20		6
5	1	14	15	12	12		7
6	3	14	13	20	5		9
7	5	12	11	24	1	0	11
8	8	7	9	21	0	0	11
9	8	4	5	13		0	11
10	10	4	4	4		1	10
11	12	2	2			3	8
12	13	1	2			2	7
13	10	1	1			5	5
14	9	0	1			4	2
15	9	0	0			5	3
16	4		0			9	2
17	3		0			8	1
18	3					10	1
19	1					11	1
20	1					10	0
21	0					8	0
22						8	
23						5	0
24						4	
25						2	
26						2	
27						1	
28						1	
29						0	
30						0	

Table VI

Profiling scheme associated with VASS Form P12 (Halloun, 1997)

Profile		Number of Items
Type	Code	out of 30
Expert	EP	19 items or more with *expert views*
High Transitional	HTP	15 to 18 items with *expert views*
Low Transitional	LTP	11 to 14 items with *expert views* and an equal or smaller number of items with *folk views*
Folk	FP	11 to 14 items with *expert views* but a larger number of items with *folk views,* or 10 items or less with *expert views*

Table VII

Respondents' sum of scores on VASS dimensions in the 1999 fall term

Dimension	Max. Score	Mean (%)	S.D.
Structure	5	3.3 (66)	1.12
Methodology	21	13.9 (66)	2.51
Validity	12	7.8 (65)	2.33
Scientific domain	38	25.0 (66)	3.91
Learnability	15	9.6 (64)	1.83
Reflective thinking	55	38.2 (69)	5.65
Personal relevance	17	11.4 (67)	2.34
Cognitive domain	87	59.2 (68)	7.15
Total sum of scores on entire VASS	125	84.2 (67)	8.99

Table VIII

Distribution of view types across VASS profiles

Profile	*Range*	*Median*	*Percentile 75*	*95*	*99*
Total number of Expert views					
Folk	7-17	12	13	16	17
Low transitional	11-20	16	17	19	20
High transitional	15-25	20	21	23	24
Expert	19-30	24	26	28	30
Total number of Folk views					
Folk	5-23	16	18	20	23
Low transitional	4-16	11	12	15	16
High transitional	2-12	8	9	11	12
Expert	1-9	5	6	8	9
Total number of Mixed views					
Folk	0-12	5	6	10	12
Low transitional	0-14	6	8	11	13
High transitional	0-14	5	7	10	13
Expert	0-9	4	6	8	9

Table IX

Chi-Square pertaining to the cross-tabulation of actual respondents' answers on VASS items and respective profiles ('99 Fall data)

Item	Chi-Square	Significance	Item	Chi-Square	Significance
6	69.65	.000	25	145.02	.000
7	68.13	.000	26	44.93	.000
8	35.28	.000	27	113.68	.000
9	35.04	.000	28	89.84	.000
10	23.74	.000	29	150.54	.000
11	2.05	.562	30	222.95	.000
12	14.95	.002	31	126.34	.000
13	11.57	.009	32	156.90	.000
14	12.08	.007	33	88.19	.000
15	4.59	.204	34	36.59	.000
			35	79.99	.000
16	19.70	.073	36	58.86	.000
17	107.21	.000	37	67.32	.000
18	52.16	.000	38	80.16	.000
19	63.07	.000	39	49.17	.000
20	19.48	.078	1	76.12	.000
21	65.27	.000	2	92.20	.000
22	155.86	.000	3	53.42	.000
23	123.86	.000	4	65.67	.000
24	89.45	.000	5	34.42	.001

Table X

Comparison of VASS results before instruction (fall term) and after instruction (spring term)

Dimension	Mean		S.D.	
	Before	After	Before	After
Structure	3.3	3.4	1.12	1.06
Methodology	13.9	13.8	2.51	2.70
Validity	7.8	7.8	2.33	2.17
Scientific domain	25.0	25.0	3.91	3.91
Learnability	9.6	9.9	1.83	1.89
Reflective thinking	38.2	36.9	5.65	5.95
Personal relevance	11.4	10.5	2.34	2.54
Cognitive domain	59.2	57.2	7.15	7.65
Total sum of scores	84.2	82.2	8.99	9.56

Profile	Percentage	
	Before	After
Folk	8	14
Low Transitional	31	35
High Transitional	45	39
Expert	16	12

Table XI

Comparison of VASS sum of scores across demographic strata

Sector	**Mean**	**S.D.**	**F**	**p**
Public	82.26	8.43		
Private	84.72	8.34		
Total / ANOVA	83.52	8.47	15.10	.000
County				
North	83.34	8.80		
South	82.46	8.81		
Beirut & Mountain	84.73	7.69		
Total / ANOVA	83.52	8.47	4.58	.011
Gender				
Girls schools	82.34	7.93		
Boys schools	83.59	9.80		
Mixed schools	84.26	8.46		
Total / ANOVA	83.52	8.47	3.88	.021

Table XII

Achievement indices across VASS profiles

Profile		Physics Course Grade	Q1	Q2	Q3
Folk	Mean	53.64	3.09	2.99	2.84
	S.D.	22.35	.78	.91	.99
Low transitional	Mean	59.93	3.32	3.13	3.21
	S.D.	22.33	.80	.81	.91
Low transitional	Mean	67.47	3.63	3.58	3.59
	S.D.	18.60	.75	.79	.87
Expert	Mean	77.39	3.88	3.87	3.95
	S.D.	15.06	.74	.76	.70
Total	Mean	64.29	3.48	3.38	3.40
	S.D.	21.15	.81	.86	.94
ANOVA	F	12.63	26.58	36.07	37.49
	[illegible]	000	000	000	000