Summative Report of the Project:

Evaluation of the Impact of the New Physics Curriculum on the Conceptual Profiles of Secondary School Students

Evaluation de l'Impact du Nouveau Curriculum de Physique sur les Profils Conceptuels des Élèves du Secondaire

Ibrahim A. Halloun

Faculty of Education – Lebanese University

Phoenix Series

Lebanese University

Beirut, Lebanon

2007

Project sponsored in part by a grant from Lebanese University

EVALUATION OF THE IMPACT OF THE NEW PHYSICS CURRICULUM ON THE CONCEPTUAL PROFILES OF SECONDARY SCHOOL STUDENTS

Abstract

Lebanese secondary school students' conceptual profiles in physics are evaluated by comparison to their U.S. peers, in an attempt to ascertain the effectiveness of the Lebanese physics curriculum currently in place. A conceptual profile consists, among others, of conceptions and dispositions that students develop about a particular discipline. A battery of three instruments were developed and validated to assess student profiles in particular areas of physics. Two instruments, the Inventory of Basic Conceptions in Mechanics and the Inventory of Basic Conceptions about DC circuits, targeted student conceptions in classical mechanics and electricity respectively. A third instrument, the physics form of the Views About Science Survey, targeted student dispositions about knowing and learning physics. The three instruments were administered to over three thousand Lebanese and U.S. students between 2004 and 2007. Results show that Lebanese students: (a) enter secondary school encumbered with naive conceptual profiles that are at odds with scientific paradigm, (b) fail, after physics instruction, to enhance their profiles and develop them to the level aspired for in the official curriculum, and (c) lag, in most conceptual respects, behind their U.S. peers.

Table of contents

Keywords / Mots clés

ENG: Attitudes, authentic assessment, conceptions, conceptual profile, curriculum evaluation, diagnostic instruments, dispositions, electricity, mechanics, nature of science.

FR: Attitudes, conceptions, diagnostic, dispositions, électricité, évaluation authentique, évaluation d'un curriculum, mécanique, nature de la science, profil conceptuel.

1. RATIONALE

Lebanon began implementing in 1998 new curricula at all pre-college levels, including secondary school physics. In physics, like in all other disciplines, the new curriculum is supposed to be aligned with novel, research-based educational theories that have driven educational reform in numerous countries around the world in the last two decades (Decree 10227, 1997 ; Halloun, 1998a). In particular, the curriculum in question is supposed to be geared with all its components, from program, to means and methods, to assessment, in the direction of meaningful learning of physics whereby students would be empowered to develop what we call scientific conceptual profiles (Ibid, pp. 395, 419, 420). This research is meant to ascertain, in certain respects, to what extent Lebanese secondary school students actually develop the target profiles in certain areas of physics.

Bachelard (1940, p. 42) introduced the notion of *epistemological profile* as the mix of "conceptualizations" that one might have about a given scientific conception (mainly a scientific concept like the concept of mass), a mix commonly grounded in different epistemologies, and extending from naïve realism and positivism through different degrees of rationalism. Mortimer (1995) developed Bachelard's profile into what he called *conceptual profile* in order to include methodological as well as epistemological aspects of a given conception. I further developed the ideas in question, in conjunction with Kuhn's (1970) idea of a *paradigm*, and came up with what I called *paradigmatic profile*, so as to cover epistemological, methodological and meta-cognitive aspects not of a single conception, but of a set of related conceptions that pertain to the make-up of a particular scientific theory or set of related scientific theories (Halloun, 2004/2006, 2007, 2008).

For the purposes of this research, a *conceptual profile* is defined as the part of a paradigmatic profile that students develop about a particular scientific theory, and consisting of personal conceptions and dispositions related to the theory in question. A *conception* is, for us, a concept, a law or any other theoretical statement that enter in the make-up of *content knowledge* of a given person, be it a scientist or a lay person. A *disposition* is a viewpoint, an attitude, a trend or a habit that one possesses about constructing, deploying, and ascertaining

Core-disciplinary dimensions

Content knowledge, drawn from the *episteme* of a given discipline (i.e., from the body of established knowledge shared and accepted by the members of a particular professional community). For example, the episteme of a given scientific community (physicists, chemists, biologists) consists of corroborated scientific theories. A science course is normally about certain *conceptions* (concepts, laws and other theoretical statements that make up certain conceptual models) in a particular scientific theory or set of theories.

Process knowledge, drawn from the *methodology* of knowledge construction and deployment in a given discipline. In the case of science, this knowledge pertains primarily to *model* construction and deployment, along with associated *tools* and *rules*. The ultimate target of a given curriculum is to stabilize process knowledge so as to turn it into permanent *skills* in student profiles (Halloun, 2004/2006).

Meta-cognitive dimensions

Learning styles, which primarily are about processes of reflective thinking that help students regulate their own profiles in insightful and meaningful ways.

Emergent trends or *dispositions*, which include habits and attitudes that are particular to a given professional community. In the case of science, dispositions include habits and attitudes that have been commonly promoted in recent calls for scientific literacy. They also include respective student views about the nature and relevance of science.

Figure 1. Dimensions of student paradigmatic profiles anticipated in science curricula (Halloun, 2008).

content knowledge. In this research, we are concerned with conceptions mainly from an epistemological perspective, and with dispositions primarily from a meta-cognitive perspective, and to a lesser extent from a methodological perspective. Our research targets, in particular, *basic conceptions and dispositions*, i.e., elementary and fundamental conceptions and dispositions (some learning styles included) that are *most critical* for *meaningful* learning of a given physics theory as targeted in the Lebanese secondary school curriculum.

A science course, especially at the secondary school and college (introductory university) levels, is usually about a particular scientific theory, and sometimes about a set of interconnected theories. The course content can be organized around a number of models that may be graded into categories of increasing structural and functional, and thus *epistemic*, complexity. Each category characterizes a cognitive evolution level that students need to attain at a certain point of instruction. Our research suggests that models of a given theory, and thus course content, can be pedagogically classified into three categories of increasing epistemic complexity (Fig. 2). The first category includes *primary models*. These are *simple basic models* relative to which students usually have the richest repertoire of subsidiary models, and thus in the context of which students can begin to develop the most fundamental conceptions of the theory (generic concepts, laws and other theoretical statements). The second category includes the rest of, and more complex, *basic models*. The third category includes *emergent models*. For example, secondary school Newtonian mechanics courses are typically about five basic particle models (Halloun, 2001a, 2004/6, 2007). Two models, the free particle model and the uniformly accelerated particle model, make up the category of primary models. Three other basic particle models, the bound particle under uniform circular motion, the harmonic oscillator, and the impulsion-driven particle, make up the middle category of Figure 2. Emergent models in classical mechanics courses usually include the model of a particle in uniformly accelerated circular motion, models of centrally-bound particles in elliptical motion and other types of motion with variable acceleration.

The three categories are organized and graded in such a way that students cannot meaningfully learn any model in a given category before learning all models in the lower category. The three categories are thus separated by *critical* demarcation lines. They are critical in the sense that at the level of each line is set a *threshold* of understanding that students need to meet before crossing into the upper category. Two *critical thresholds* can thus be set in any given course: the *basic threshold* between primary models and the rest of

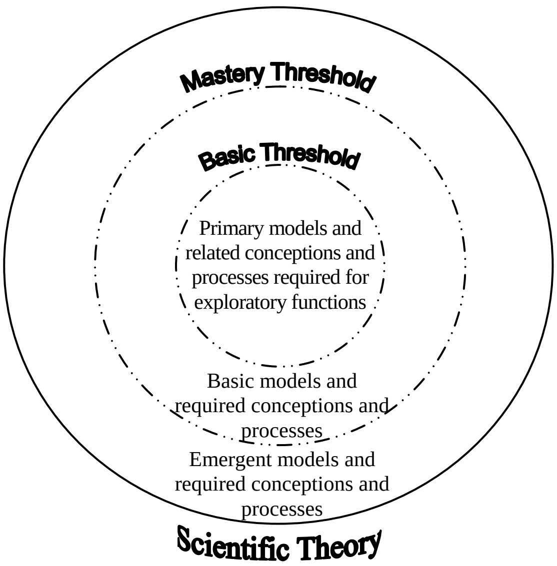

Figure 2. Thresholds of cognitive evolution in a science course. Evolution from one stage to another is not possible unless students meaningfully attain a particular critical threshold (Halloun, 2008).

basic models, and the *mastery threshold* between basic and emergent models (Halloun, 2008). Our research is concerned with conceptual profiles at the basic threshold.

The efficacy of any curriculum is function of the degree to which it allows individual students have their conceptual profiles evolve and become commensurate with the scientific paradigm which they correspond to. The object of this research is to ascertain the impact of the physics curriculum in place since 1998 on secondary school student conceptual profiles pertaining to classical mechanics and DC circuits. The two areas are chosen because: (a) they are at the entry level in secondary school physics, (b) they are the two most salient areas in learning physics at the secondary school and upper levels, and (c) because they are most closely related to everyday life, thus offering a natural and familiar context for students to develop their subsidiary models and their dispositions about knowing and learning physics and science in general.

Educational research in the last three decades has shown, worldwide, that: (a) students of all levels come to their mechanics and electricity courses with deeply seated conceptual profiles about the motion of physical objects and the structure and behavior of electric circuits, (b) that these profiles, which are developed in everyday life encounters, are grounded mostly in naïve realism and are at odds with scientific paradigm (theory and practice), and that (c) after the completion of respective physics courses, students fail to significantly resolve incompatibilities between their own profiles and scientific paradigms (Aikenhead, 1987; Bagno & Eylon, 1997; Cohen et al., 1983; Dupin & Joshua, 1987; Edmonson & Novak, 1993; Halloun, 1986, 1993, 1997, 2001a & b; Halloun & Assi, 1989; Halloun & Hestenes, 1985a & b, 1995, 1998; Abell & Lederman, 2007; McDermott, 1984; Meichtry, 1993 ; Rainson et al., 1994 ; Songer & Linn, 1991; Thornton & Sokoloff, 1998; Viennot, 1996). Research has also shown that curricular reforms that took place in many developed and developing countries in the last decade or so have often failed to significantly remedy the situation thus described.

Many movements have taken place worldwide in the last two decades to reform the state of science education at all educational levels. Some of these movements were led by international organizations such as UNESCO (UNESCO, 1993), others by professional associations (AAAS, 1990, 1993; NSTA, 1995), and some by public agencies (NRC, 1996) or local governments (like in the case of Lebanon). Virtually all these movements share foundations and aspirations that are similar to the ones behind our new curricula. The degrees of implementation and success vary from one country to another (Bibeau et al., 2002; IEA, 1995). However, no report has ever yet been published indicating that a particular reform movement has lived up entirely to what it was originally set to accomplish. Time may be a major impediment in this respect, since it takes any curriculum many years and cycles of reiteration and refinement until it may meet its ends, if ever. Some educators have even been skeptical about the efficacy of such reform movements, or at least about the reproducibility and sustainability of their impact when such impact turns out to be significantly positive in specific local settings (French, 1989; Hake, 2004).

The new Lebanese physics curriculum is claimed to be based on the same premises as the reformed curricula in question. However, analysis of our new curriculum has revealed that it actually falls behind what it claims (Halloun, 1998a). Subsequently, we were led to the following hypotheses that make the object of this research:

1. Lebanese secondary school students (Grades 10-12) come to their mechanics and electricity courses encumbered with naïve conceptual profiles that are at odds with scientific paradigm.
2. Student conceptual profiles do not significantly evolve in the direction of scientific paradigm after the completion of their mechanics and electricity courses.

3. Curriculum shortcomings are about both conceptions and dispositions, and perhaps more serious about the latter.
4. Conceptions and dispositions are closely related, and affect the evolution of one another.
5. Lebanese students lag behind their international peers, especially in USA, in resolving incompatibilities between their own conceptual profiles and respective scientific paradigms.

The last hypothesis was not part of the original proposal that was approved by the Research Board at Lebanese University (as per the research contract of March 26, 2004). It was made possible after I had the chance to spend about two academic years on leave in USA (October 2004 – April 2006). My work there has also made it possible to extend the scope of current research in the direction of the development of an authentic assessment framework for meaningful learning of science in the manner described in the last section of this report.

Measures taken to test these hypotheses and deploy results are described in the following three parts of this summative report. Research method is presented in the first part which includes details about participating samples of students in Lebanon and USA, the battery of instruments developed and validated to test the hypotheses, and respective procedures. Results are presented and analyzed in the following part. Their implications and prospects for their extrapolation are discussed in the last part. The report is supplemented with a two-part appendix. The first part of the appendix presents the personnel involved in the research along with budgetary expenditure. The second part presents the battery of instruments along with respective taxonomy.

2. METHOD

This research went into five phases:

Phase 1 extended from November 2003 through March 2004. This was a preparatory phase completed before the research project was approved and financed by LU administration. The object of this phase was to put together, in cooperation with Lebanese physics teachers: (a) a viable taxonomy of basic mechanics and electricity conceptions and general physics dispositions which would make up the *basic conceptual profiles* that the research should target, along with (b) paper-and-pencil drafts of the respective instruments.

Phase 2 followed immediately the official approval of the research project by LU administration in March 26, 2004, and extended through September 2004. The previously prepared instruments were piloted and validated in this phase with a sample of Lebanese secondary school students.

Phase 3 extended from October 2004 through September 2006. Arrangements were originally made to revise the instruments and administer them to a sample of Lebanese secondary school students, first as pretests during the fall of 2004 and then as posttests toward the end of the 2005 spring semester. The investigator being away on sabbatical leave, those arrangements did not materialize. Instead, revision and research took place in USA for a comparative study between Lebanese and US students, in order to test hypothesis 5. In addition, work began for the preparation of an authentic assessment framework that would apply, at any educational level, to any scientific discipline.

Phase 4 extended from October 2006 through February 2007. During this phase, the battery of instruments was administered to a Lebanese sample of students. The first draft was also written of the authentic assessment framework.

Phase 5 extended from March 2007 through September 2007. Data were analyzed in this phase, and this report prepared. Draft manuscripts were also prepared for prospective publications about this research. Work on the assessment framework continued.

Ample details about the five phases are presented in the following. Participating researchers and students are first presented. The battery of instruments is then described, followed with respective development, validation and implementation procedures.

2.1 PARTICIPANTS

Participants included collaborators and secondary school and college (university) students (hereby referred to as students). *Collaborators* were university professors and pre-service and in-service physics teachers, mostly Lebanese, and some Americans. Collaborators participated mainly in the preparation and dissemination of the instruments, and conducted interviews with participating students. Some contributed, and still are contributing, to the development of the authentic assessment framework. Participating students were distributed among a number of Lebanese secondary schools and one U.S. university. Students took various forms of the instruments, and some were interviewed to establish instrument validity.

Except for “assistant researchers” presented in the project expenditures and whose remuneration came from this research budget, collaborators either volunteered their work, or, in the case of American participants, were compensated by sources other than LU (mainly, from a National Science Foundation grant in USA). Except for U.S. students who were interviewed, no student was compensated for her/his participation.

Collaborators

Phase 1 collaborators included 18 Lebanese secondary school physics teachers with teaching experience ranging from two to twenty years of teaching physics at the secondary school level, as well as three university physics professors. These colleagues participated in the refinement of taxonomy and instruments in the manner described in § 2.3.

In Phase 2, nine graduate students at UL Faculty of Education participated in the research. Collaboration, as described in the project expenditures, ranged from instrument refinement and/or administration to data entry and/or analysis. Five of these collaborators also interviewed some participating students.

Phase 3 collaborators included six American university professors and two staff members at a state university in USA. Two faculty members participated in taxonomy and instrument refinement and conducted interviews with students. All participating faculty administered the instruments to their students. Staff members were mostly involved in putting some instruments on-line, monitoring on-line administration and gathering respective data. Two of the U.S. faculty and a Lebanese university professor took also part in breaking grounds for the Authentic Assessment Framework.

In Phase 4, a new group of nine graduate students at UL Faculty of Education participated in the research. These collaborators administered the instruments to Lebanese secondary school students and/or helped in data entry or analysis. No interviews were conducted in this phase.

I carried out Phase 5 mostly on my own. Some collaboration took place only in developing the Authentic Assessment Framework.

Students

Over 3,200 students participated in this research, i.e., more than three times the number of one thousand students originally anticipated in the proposal approved by LU administration. Students were administered three different instruments either in a paper-and-pencil format or electronically, on-line. Students were discarded if they missed about 20% or more of the items on any given test, or if they showed a pattern of answers that made it evident that they did not take the test seriously. As a consequence, 2,676 students were left for data analysis and inclusion in this research. 1,662 students (62%) were enrolled in Lebanese secondary schools, and 1,014 (38%) were freshman students enrolled in a USA university.

No secondary school students participated in Phase I of the research. Phase 2 involved a sample of over 700 students enrolled in eleven secondary schools situated in the greater Beirut area, and chosen at random. Six were private schools, five were public (official) schools. Following data scrutiny, about 20% of these students were discarded and 575 students were left for inclusion in data analysis. The sample is considered by all statistical criteria large enough for a pilot study. 49% of these students were female, 51% male. 75% of them were enrolled in the first secondary grade (G10), the rest in the second grade (G11) with science track. All participants studied physics in French.

About 1,200 students participated in Phase 3 which was carried out at a U.S. university. Participants were enrolled in numerous sections of two Newtonian mechanics courses of different levels in the 2005 Spring semester (thereafter referred to as USM1 and USM2), or of an electricity course the following Fall semester (USE). Following raw data clean-up and elimination of students who did not take the tests seriously, 608 students were left from the spring semester (241 students in USM1 and 367 students in USM2), and 406 USE students from the fall semester.

USM1 is an algebra-based mechanics course taken by students most of whom had never taken physics before, not even in high school. This course is somewhere between the first and the second mechanics course offered in Lebanese secondary school. USM2 is a calculus-based mechanics course taken by students who, on average, would have taken one mechanics course before, mostly at the high school level. This course is slightly above the last mechanics course typically offered in Lebanese secondary school. Most students enrolled in the electricity course (USE) would have never taken electricity before, not even in high school. Their physics background thus makes the U.S. sample, and as far as our research is concerned, comparable to participating Lebanese secondary school students. Research, including some conducted by this author, indicates that physics students at the U.S. university from which our sample was drawn are representative of the population of students enrolled in similar courses at various U.S. universities.

About 1,300 Lebanese secondary school students participated in Phase 4. 1,087 students were kept for data analysis after data clean-up. Students were enrolled in 14 schools, randomly chosen from different parts of Lebanon. 43% of students were female, 57% male. Two third took physics in English and the rest in French. 52% were Grade 10 students, 37% Grade 11 students (virtually all in the science track), and 11% Grade 12 students (all of whom in science tracks).

2.2 INSTRUMENTS

A battery of three instruments was developed and validated in this project to assess the impact of the new physics curriculum on student conceptual profiles, and more specifically on students’ content knowledge and dispositions – learning styles included – (Fig. 1). Profiles were targeted at the basic threshold (Fig. 2). Two instruments, the Inventories of Basic

Conceptions (IBC), ascertain basic conceptions in classical mechanics and DC circuits, and one instrument, the physics form of the Views About Science Survey (VASS), ascertain basic dispositions about knowing and learning physics that are common to all secondary school physics courses. The three instruments, as used in Phase 4 of the project, along with corresponding taxonomy, are presented in the second part of the Appendix. Earlier versions of these instruments used in Phase 2 were presented in the first progress report of November 2004. IBC-mechanics consisted of 33 multiple-choice items in all phases of the project. IBC-DC circuits consisted of 36 multiple-choice items in Phase 2, and of 33 items subsequently. VASS consisted of 48 Contrasting Alternatives rating scale (CArs) items with 3 control items in Phase 2, and of 50 CArs items with 5 control items subsequently (details below). All three instruments were written in English and French.

IBC-Mechanics emerged from the renowned Mechanics Diagnostic Test (Halloun & Hestenes, 1985a) and its successor, the Force Concept Inventory (Halloun & Hestenes, 1995; Hestenes et al., 1992). VASS emerged from an instrument that bears the same name (Halloun, 1997, 2001b; Halloun & Hestenes, 1998). IBC-DC Circuits is a new instrument developed specifically for this research. The taxonomy of IBC-Mechanics and VASS, as well as the instruments themselves, as presented in the Appendix, are major refinements of their respective predecessors.

The predecessors of IBC-Mechanics targeted a broad picture of classical mechanics and covered aspects from virtually all five basic models mentioned in the Rationale. The new instrument however, as can be seen in its taxonomy, focuses on basic Newtonian concepts and laws in the context of only the two primary models of the free particle and the uniformly accelerated particle models, models that define the *basic threshold* in Newtonian theory (Halloun, 2008); whence the term *basic* in the name of the instrument. Our research with the former two instruments has shown that content knowledge at the basic threshold is critical for upper level content knowledge. Students who do not reach this threshold are unable to meaningfully develop upper level knowledge, and the success of instruction is primarily determined by student ability to reach this threshold (Halloun, 1998b; 2008).

IBC-DC Circuits has been developed along the same philosophy as its mechanics counterpart. The taxonomy of this instrument was originally determined in cooperation with Phase 1 collaborators. As discussed in the following section, taxonomy and instrument were slightly refined throughout subsequent phases so as to reliably ascertain content knowledge that validly defines the basic threshold in DC circuit courses.

VASS was originally developed in the mid 90’s to ascertain student views about knowing and learning a variety of scientific disciplines (Halloun, 1997, 2001b). Its taxonomy and form gradually evolved to take the form presented in the Appendix. A major difference between the current form and its predecessors is that, unlike the latter, the current form distinguishes between what actually takes place in a physics classroom and what students would have wanted to happen there. Prior forms of VASS dealt with the nature of physics solely from physicists’ perspective, whereas the current form dealt with it from students’ perspective, as well, in the context of the physics classroom and not just in the context of physicists’ laboratories and fields of study. The new form of VASS thus provides a *direct* measurement of actual dispositions fostered, through practice, in the classroom. By contrast, prior forms, like all other instruments in the literature dealing with the nature of science, provide *indirect* measurement of how classroom practice *might* affect general dispositions about the discipline of physics. By the end of this research, VASS was additionally refined in this direction so as to become a more valid and reliable instrument than its predecessors (or any other instrument to that matter) for ascertaining what directly takes place in the classroom and bring about student dispositions in one direction or another. Parallel versions of the latest form are also being made available for mathematics and other scientific disciplines at www.halloun.net.

2.3 PROCEDURES

First versions of the three instruments were written in Phase 1. Collaborators were then provided with the taxonomy of the three tests for analysis and feedback, and were asked to actually take the tests (at home) in January and February 2004. Following analysis of their feedback and written answers they provided on individual items, taxonomy and some items were revised for enhancing the validity of the instruments.

Refined instruments were administered in Phase 2, specifically in May 2004, to the respective sample of Lebanese secondary school students. Student data were analyzed, and the instruments were subsequently revised to take the form presented in the first progress report submitted to LU administration in November 2004.

The three tests were given during the month of May 2004 when participating schools had already covered materials assessed in the tests, especially IBCs. Each group (class) of students took two tests during two different class periods, and students were afforded enough time (up to one class period of 50 minutes) to complete each test. 36% of students took IBC-Mechanics along with VASS, 43% took IBC-DC circuits along with VASS, and 21% took the two IBC tests but not VASS. Tests were given to each group by the physics teacher in charge of the course, often a graduate (or a graduate student) of this author, or in presence of such a graduate/student (administrator). Administrators (nine of them) were trained ahead of time on the procedures to follow. They were to: (a) explain what a given test is about, (b) encourage students to take the test seriously, (c) provide clear directions on how to take the test while refraining from helping students choose answers, and (d) ensure that questionnaires and answer forms are duly passed out and collected.

Versions of the tests written early in spring 2004 were then piloted with a sample of over 700 Lebanese secondary school students. For each test, each student received one questionnaire and a separate answer sheet. Filled answer sheets were scrutinized to tease out those with significant flaws. Two major flaws were considered: sheets with answers of a particular pattern (e.g., all “A” answers, or a repeated pattern of the same consecutive answers), and sheets with less than 80% of questions answered. Those sheets were discarded, and about 18% of participating students were subsequently dropped out of data analysis. Data reported thereafter pertain to the 575 secondary school students mentioned in the previous section.

In this Phase 2 pilot, students were asked to provide their own answer if they felt that no provided alternative reflects what they think about a particular item in a given test. A few students did so. When they did, it often turned out that the alternative they wrote was similar to one already provided. The same thing happened during interviews conducted with a number of students to assess the face validity of individual items. Whenever necessary, items and/or alternatives were modified to clarify things and/or express them in a more colloquial form for Phase 3.

Following administration of the written tests, and preliminary analysis of the data, each administrator chose at least three students for interview from each group. Interviews were conducted with a total of 45 students and were meant to assess the face validity of the instruments and reliability of students in choosing particular answers. Students were classified in three categories based on their performance on a particular IBC test: high, average and low performers. One or two students from each category were chosen for interview with a particular administrator. Particular items from IBC and VASS were discussed during the interview with each student in order to assess whether the student: (a) actually understood what each item is about and the scope of the respective answer, (b) answered an item seriously and can reproduce the same answer originally provided in the written test, (c) provided an answer that actually reflects what s/he thinks about the particular item, and

(d) can duly justify the provided answer. Administrators refrained in the process from telling interviewed students what their original answers were and whether or not they chose a correct answer, or from leading students to correct choices. Except in rare cases, interviews revealed that students took tests seriously, that their answers were reliable and provided support for the face validity of the various items on any given test. Where interviews showed problems with a particular item that are consistent with what statistical item difficulty analysis was showing, the item was duly revised for Phase 3.

Quantitative data and interview analysis revealed that a few IBC-Mechanics items required minor refinement, mostly in rephrasing a particular alternative for clarity purposes. The situation was different with the other two tests. In the 2004 May pilot version, issues discussed in items 20 through 27 in IBC-DC circuits were presented in 12 items (rather than 8), and written in a form that differs from what is shown in the Appendix. Data and interviews revealed that students had difficulty interpreting these items in the form then presented, which led to their revision. There was also an item (No. 18, then) that data analysis showed that it needs to be entirely dropped out from the taxonomy and test. Dropping the latter item out and revising the former ones made room to add items 15 and 16 which did not figure originally in the pilot version (1st progress report). IBC-mechanics consisted of 33 items in all three implementation phases (2, 3 and 4), while IBC-DC circuits consisted of 36 items in Phase 2, and of 33 items in Phases 3 and 4.

Phase 2 pilot form of VASS consisted of 48 items distributed in three subsets: 3 control items (originally items 1, 2 and 48), 23 items about the scientific dimensions of the taxonomy (originally items 8 through 30), and 22 items about the pedagogical dimensions of the taxonomy (originally items 3-7 and 31-47). Following the administration of this pilot form in May 2004, and analysis of respective data and interviews, it became evident that two more control items were needed (items 3 and 4 in the appendix of the 1st progress report and the current report), and that the other items needed to be clustered differently in the body of the test. More specifically, pedagogical items subsequently separated in two clusters, one about the way physics courses are *actually* being taught, and the other about the way students would *like* these courses to be taught. These were respectively clusters 28-39 and 40-49 in the VASS form used in Phase 3 and presented in the 1st progress report. Scientific dimensions were then covered in cluster 5-27, and control items in cluster 1-4 and item 50. Some items required minor revisions for clarity purposes.

I wrote the three tests and their respective taxonomy towards the end of the 2003 fall semester and beginning of the 2004 spring semester. I then revised taxonomy and tests based on statistical analysis of the quantitative data (using SPSS) and qualitative analysis of the interviews, as well as on feedback from three university physics professors who reviewed the refined instruments for content and face validity. In addition to handling the logistics of administering tests in various schools and assisting in data entry and analysis, Phase 2 collaborators contributed their insights for the refinement of the three instruments.

Plans were originally made to administer the instruments revised by the end of Phase 2 to a new sample of Lebanese secondary school students in October 2004 and May 2005 as pretests and posttests respectively, and to subsequently bring the research to closure. However, I moved to USA in that October and spent about two academic years there. As a consequence, the research took a new direction. The planned pilot took place in USA instead of Lebanon, and spanned through two consecutive academic years (2004-2006).

In the fall of 2004, results of Phase 2 were analyzed. With the assistance of U.S. personnel, the instruments were further refined that fall, and prepared for dissemination in the subsequent spring and fall semesters of 2005. All instruments were then put in two formats: paper-and-pencil, and electronic for online administration. IBCs consisted then, as they still do, of multiple choice items, while VASS consisted, and still does, of Contrasting

Alternatives rating scale (CArs) items. I developed CArs about thirteen years ago specifically for VASS (Halloun, 2001b). Many researchers around the world have since adopted this scale in developing new instruments for a variety of purposes.

In Phase 3, IBC-Mechanics consisted of 33 multiple-choice items, and so did IBC-DC Circuits, while VASS consisted of 50 CArs items. In their current form (used in Phase 4), the three inventories are slight refinements of their predecessors used in Phase 3 and reported in the 1st progress report. Refinements are about the wording of some questions and took place mostly for language clarity purpose. The taxonomy of the three instruments was about the same as the one presented in the current appendix.

As indicated in the previous section, three groups of U.S. university students enrolled in two mechanics courses (USM1 and USM2) and one electricity course (USE) took part of our research in Phase 3. USM1 and USM2 students took IBC-Mechanics as pretest the second week of the spring semester, and as posttest, three weeks before the end of that semester. The part of mechanics covered in that IBC was completed in the courses in question long before the posttest was administered. USE students took IBC-DC circuits only as a pretest at the beginning of their electricity course. USM2 students also took VASS along with their IBC pretest. All IBC pretests were administered in a paper-and-pencil format. Corresponding posttests and VASS were administered electronically, online. Students were assigned specific computers at their university to take the tests in question in due time. In all IBC pretests and posttests, students were given 30 minutes to complete each test. They were given 50 minutes to complete VASS. Test administration followed the same guidelines described in Phase 2.

A representative sample of 24 students was chosen for interview following the same criteria of Phase 2. Two U.S. professors conducted the interviews and corroborated face validity of the instruments and student answer reliability along the same lines of Phase 2. Following test administration and interviews, some items were slightly reworded to improve clarity, and the three tests took the form presented in the appendix for deployment in Phase 4.

The three tests were administered in Phase 4, late March or early April 2007, to a sample of Lebanese secondary school students. 493 students took IBC-Mechanics, 721 students took IBC-DC circuits (215 of whom had also taken the mechanics test as well), and 687 students took VASS in addition to either IBC test. By then, students had completed the part of their mechanics or electricity course covered in the respective IBC. Therefore, Phase 4 data can be considered as pertaining to an immediate posttest. Collaborators were all graduate students at the Faculty of Education (cf. appendix). They administered the tests in schools where they were teaching. Given their late entry to the said Faculty in November 2006, the battery of instruments could not be administered as pretests. As we shall discuss in the next section, pretest data were not relevant to this phase of the project. No interviews were conducted in this phase since the validity and reliability of the instruments were already established as we shall discuss in the following section.

3. Results

Results of Phase 2 and Phase 3 have already been partially presented in the previous two progress reports submitted to LU's Board of Research in November 2004 and February 2007 respectively. In this section, we present major results of Phases 2, 3, and 4 mainly in the following respects: (a) validity and reliability of the instruments, (b) the impact of the Lebanese curriculum on student conceptual profiles, and (c) profile comparison between Lebanese and U.S. students.

3.1 VALIDITY AND RELIABILITY OF THE INSTRUMENTS

Phase 2 was primarily concerned with establishing the validity and reliability of the three instruments. We begin this subsection with detailed analysis of those issues as ascertained in that phase, and we follow up with respective outcomes obtained in subsequent phases.

Taxonomy of IBCs is confined to the basic threshold of meaningful understanding of a given course as discussed in the Rationale (Fig. 2). Taxonomy of VASS is confined to those issues that 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 (Halloun, 2001b). Validity of each instrument was assessed in four respects. First, university professors and experienced high school teachers who are versed in educational research pertaining to our work verified the *content*, and more specifically the *sampling* validity of the taxonomies of our three instruments, as well as *item* validity, i.e., the validity of any particular item to assess what it corresponds to in the taxonomy of each instrument. Second, the same collaborators virtually all agreed on what we consider as correct or expert answers to all questions, thus corroborating the *face validity* of the instrument. Third, interviews with participating students revealed that students have understood most questions and the nature of the anticipated answers. When flaws were detected, refinements were made in the manner described in the previous section. Fourth, *predictive validity* was estimated indirectly through Pearson correlation coefficients between students' scores on the pairs of test they took.

Table 1

Pearson correlation coefficients between various test pairs in Phase 2

		IBC-Mechanics	IBC-DC Circuits	VASS
IBC-Mechanics	Pearson Correlation	1	.780(**)	.697(**)
	Sig. (2-tailed)	.	.000	.000
IBC-DC Circuits	Pearson Correlation	.780(**)	1	.430(**)
	Sig. (2-tailed)	.000	.	.000
VASS	Pearson Correlation	.697(**)	.430(**)	1
	Sig. (2-tailed)	.000	.000	.

** Correlation is significant at the 0.01 level (2-tailed).

Table 1 above shows results in Phase 2. Similar results were obtained in subsequent phases. As can be seen in this table, all coefficients are highly significant, which indicates a strong relationship in student minds between various issues assessed in the three tests. Furthermore, and as one would expect, given the similar nature of the two tests, the relation is strongest between the two IBC tests. As for VASS, it is stronger with the mechanics test than with the electricity test. At least two reasons may be behind this. First, the pilot version of the mechanics test was better written than its electricity counterpart. Second, the nature of items is perhaps within a closer context for VASS in the mechanics test than in the electricity test. In fact, some VASS items are phrased within the context of mechanics, and none within the context of electric circuits.

Reliability of the three instruments was essentially established in three ways, with Cronbach's alpha, through interviews with students, and by comparing outcomes obtained in various phases with each other on the three tests. Cronbach alpha was measured to establish the *internal consistency reliability* of the three instruments. This coefficient though is more suitable for IBCs than VASS, given the nature of the latter instrument (Halloun, 2001b). Still, and in the absence of other indicators in classical statistical theory, it may be used as a viable, rough indicator of the internal reliability of rated scale instruments like VASS. In Phase 2,

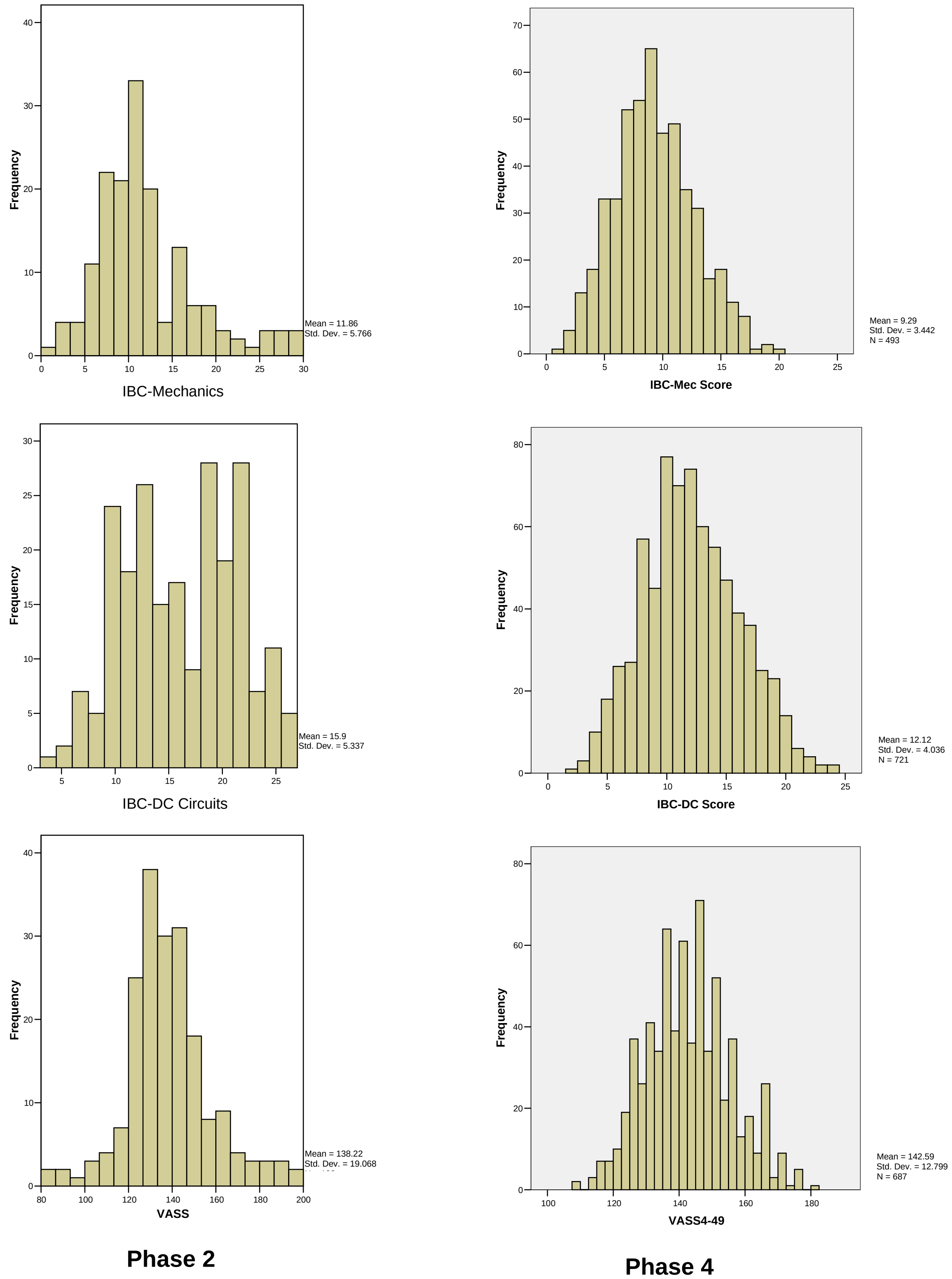

Figure 3. Comparative histograms of student scores on the three tests administered to Lebanese students in Phases 2 and 4.

this coefficient had a value of .84 for IBC-Mechanics, .79 for IBC-DC circuits, and .83 for VASS. Cronbach alpha came within 3% in phases 3 and 4 for all three instruments, higher in Phase 3 than in Phase 4. The drop in Phase 4 was primarily due to the heterogeneous nature of participating students who were distributed among the three grade levels in Lebanese secondary schools. Nevertheless, the obtained values of Cronbach alpha for all three tests are indicators of significantly high reliability. In fact, they are among the highest ever reported in the literature, especially with regard to instruments like VASS.

For *test-retest reliability* assessment, interviewed students in all three phases were asked to orally answer specific questions of each test, a few days after they had filled the written surveys, without reminding them of their written answers. Virtually all these students reiterated the same answers they had indicated previously. In a related aspect, the last item in VASS asked students how seriously they took the test. 96% of respondents expressed a positive position in this respect.

For *stability* (and *equivalence*) assessment, we compared results obtained in Phases 2 and 4 to each other and to those obtained with previous forms of the tests with Lebanese students. Figure 3 shows histograms of Lebanese student scores on the three tests in the two phases in question. The average score on IBC-Mechanics was 11.86 (S.D. = 5.77) for the entire sample of students who took this test in Phase 2, and 9.29 (S.D. = 3.44) in Phase 4 (i.e., within 8% of Phase 2 outcome). A result falling between the two phases' means was obtained with the Mechanics Diagnostic Test when administered to Lebanese students of similar background (Halloun, 1986). The average score on IBC-DC circuits was 15.90 (S.D. = 5.33) in Phase 2, and 12.12 (S.D. = 4.04) in Phase 4 (i.e., within 7% of Phase 2 outcome). The two mean scores constitute 44% and 37% respectively of the maximum possible score of 36 points in Phase 2 and 33 points in Phase 1; between them fell the average of 42% obtained on the open-ended form of this test which was piloted in Lebanon before this project took off. The average score on VASS was 138.22 (S.D. = 19.07) in Phase 2, and 142.29 (S.D. = 12.80). The two respective VASS mean scores of 61% and 63% of the maximum score of 225 points (45 items x 5 point each, in both phases) are significantly comparable to the 67% mean score obtained on a prior VASS form in Lebanon (Halloun, 2001b). As explained in the following two subsections, differences between Phase 2 and Phase 4 IBC scores are attributed to demographic variables that affect performance on IBCs but not VASS. It will thus later become evident that the two IBCs provide comparable results under the same demographic conditions, and thus are as stable as VASS. In fact, and as can be noticed in Figure 3, Phase 4 outcomes of both IBCs approach significantly better the normal distribution than Phase 2 outcomes, and Phase 4 standard deviations are significantly smaller than those of Phase 2. All in all, the overall reliability of the battery of instruments is well established in our research.

3.2 The Lebanese Curriculum and Student Conceptual Profiles

In Phase 2, data on the three tests were analyzed in relation to the documented demographic variables. These were, gender, school and class. As for gender, the two groups of participants, males and females, scored within 0.9% of each other on the mechanics test, within 1.4% on the electricity test, and within 0.43% on VASS. Females mildly outperformed males on the first and last tests, males outperformed females on the electricity test. ANOVA analysis though revealed that gender differences were insignificant in all three cases. Similar analysis revealed significant differences (p=.005) among participating schools on the mechanics test but not on the other two tests. Average scores on the mechanics test ranged from 8.86 (27%) to 12.21 (37%) with similar variation on standard deviations. Finally, ANOVA revealed significant differences (p=.000) between first secondary grade students (G10) and second secondary grade students (G11) on the electricity test but not on the other two tests. The average score on IBC-DC circuits was 15.43 (43%) for G10 students as opposed to 19.60 (54%) for G11 students.

In phase 4, language of instruction was an additional variable. Table 2 shows distribution of mean scores across values of each demographic variable and the corresponding ANOVA *F*-coefficient value and significance level *p*. As can be seen from this table, significant differences are detected between different values of any given variable for both IBC tests (p=.000), but not for VASS (all p's > .10). Mixed results were obtained with the language of instruction whereby French educated students outperformed their English educated peers on IBC-Mechanics, but the latter outperformed the former on IBC-DC circuits. Consistent results were obtained with both IBCs whereby, and systematically: (a) males outperformed females, (b) performance improved from grade 10 to grade 11, but not from grade 11 to grade 12 on IBC-mechanics (it did for IBC-DC circuits), and (c) performance varied in tandem with economic background for different schools.

Table 2

Means (S.D. between parentheses) distribution across demographic variables

Variable		IBC Mechanics		IBC DC-Circuits		VASS	
Type	Value						
Gender	Female	8.71	(2.84)	11.38	(3.80)	143.27	(12.14)
	Male	9.90	(3.51)	12.32	(4.11)	142.17	(13.49)
ANOVA	F	18.27		14.25		.43	
	p	.000		.000		.653	
Language	English	8.48	(2.56)	13.14	(4.15)	143.64	(13.44)
	French	10.17	(3.60)	11.40	(3.17)	141.19	(12.06)
ANOVA	F	19.08		66.06		2.29	
	p	.000		.000		.11	
Grade	10	7.91	(2.61)	11.74	(3.80)	142.45	(12.86)
	11	9.58	(2.85)	12.33	(4.67)	143.39	(13.83)
	12	9.53	(3.63)	14.01	(3.69)	142.73	(12.11)
ANOVA	F	12.01		26.22		.19	
	p	.000		.000		.90	
School	Lowest	7.36	(2.57)	9.21	(2.96)	142.03	(10.44)
	Highest	10.19	(3.48)	13.88	(3.69)	143.60	(13.56)
ANOVA	F	10.54		21.72		2.11	
	p	.000		.000		.13	

At first sight, Phase 4 results appear to be worse than Phase 2 results on IBCs but better on VASS. Closer scrutiny reveals that the economic background is a major factor that affects student performance on IBCs but not VASS. In fact, the apparent better VASS results in

Phase 4 are mostly due to the wording enhancement of some items. Still, the VASS difference in question (2%) is not significant by any statistical measure. On average, Phase 2 schools serve students from a better economic background than Phase 4, and some are even elitist schools that serve only high-performers (commonly referred to as gifted or talented students who can be self-educated). When the latter schools are left out from Phase 2 results, the average score drops down from 11.86 to 10.30 (31% of the maximum 33-point score) on IBC-mechanics, and from 15.90 to 14.79 (41% of the maximum 36-point score in Phase 2) on IBC-DC circuits. This would bring Phase 4 results to within 3% and 4% of those of phase 2 on the two tests respectively. Differences on the two IBC tests thus become statistically not significant between the two phases.

All-in-all, the average score, in any school, never exceeded 50% of the maximum score, at any grade level, on either IBC test, and remained idle in the low 60's% on VASS at all grade levels. As discussed in the Rationale, and as can be seen in the Appendix, the taxonomy of all three tests, and especially the two IBC tests, are chosen to meet the *basic threshold* (Fig. 2) in the evolution of student profile targeted in Lebanese secondary school physics. Taxonomy and items are chosen so that any student who reaches such threshold should ace all three tests, or at least the two IBCs. However, results show that this is far from being the case at any Lebanese school participating in our research, including those elitist schools mentioned above. Thus, in accordance with our criteria, *Lebanese schools fail to meet the expectations of the physics curriculum in place*. The situation is even worse when we look at the evolution of students across secondary school levels, and when we compare the current state of things to the state that prevailed before the current curriculum was implemented.

Given the uncontrollable constraints in this project, especially in Phase 4, none of the three tests could be administered as pretest and posttest to the same Lebanese students (the case was different in Phase 3 with U.S. students), and not especially in a longitudinal study across all three secondary school grade levels. However, our previous research suggests that pretests (or posttests) given, at the same time, to different groups of students in different grade levels at the same school are significant indicators of the evolution, through the years, of any given group of students across different grades (Halloun, 1986, 2001b; Halloun et al., 1985b, 1989). As Table 2 suggests, gain on either IBC test, from one grade level to the next, or even across all three secondary grade levels, was always less than 2.5 points. This gain is less than 8% of the maximum score of 33 points, and less than 6% of the average maximum possible gain on IBC-mechanics, and less than 10% of that gain on IBC-DC circuits. The average maximum possible gain is equal to the difference between the maximum score (33 points) and the average pretest score, in this case taken as Grade 10 average score on either IBC. Physics education research suggests that a curriculum, in any country around the world, cannot be considered to allow meaningful evolution of a student profile like the one expected in our curriculum unless it results in an average gain on inventories like ours that exceeds 60% of the maximum possible gain, from pretest to posttest within the same school year (Hake, 2004, and references therein). This further shows how drastically our schools fail to meet our curriculum ends.

The physics curriculum currently being implemented in Lebanon is claimed to be significantly better than its predecessor that was in place until 1998, and for many decades before then. In the 1980s, we conducted research on Lebanese secondary school students' understanding of mechanics and electricity (Halloun, 1986; Halloun & Assy, 1989). Results obtained then on instruments similar to IBC, especially in mechanics, show that the new curriculum does not fulfill what it claims. In fact, research using the Mechanics Diagnostic Test (MDT in Halloun, 1986) showed that Lebanese students complete their secondary school years with an average of 35% of the maximum possible score of 36 points on MDT. MDT is very similar to IBC-mechanics. It had a close taxonomy, and many MDT and IBC items are virtually the same. The average posttest score of our grade 12 students was well below the

35% level in Phase 4 (Table 2), and was at about this level in Phase 2. The situation looks even worse when we compare Lebanese students to their international peers, while keeping in mind that the new curriculum claims to be aligned with modern educational trends implemented in the rest of the world (§3.3).

Research suggests that student dispositions, like the ones assessed in VASS, are hard to change under any form of instruction that does not directly target such dispositions (Halloun, 2001b; Halloun & Hestenes, 1998). Furthermore, and as Table 1 shows, there is a significant correlation between students' core disciplinary knowledge, like the one assessed in IBCs, and their dispositions about the particular discipline of concern. VASS results shown in Table 2 indicate how drastically our physics instruction fails to meet the aspiration of our new curriculum regarding student dispositions, and provide perhaps an explanation for how deep-seated are student conceptions about the target areas of physics. Curriculum developers around the world have long argued, and research suggested, that unless students are guided to develop the kind ascertained in VASS of learning styles and epistemological framework about the discipline they are studying, secondary school students' profiles will fail to evolve into the realm of science. In fact, our current research, like previous ones, show that our *students get out of secondary school with the same kind of dispositions they bring in to their first secondary school year, and that their conceptions about the motion of physical objects and electric circuits are barely affected by three years of physics instruction at the secondary school level.* One may then argue, like many researchers do, that it is because of student inadequate learning styles and their erroneous beliefs about the nature of science, that they fail to change the sort of conceptions about the world that are ascertained in IBCs.

The quantitative data so far presented show that *our physics curriculum virtually has no impact on conceptual profiles which Lebanese students bring in to their secondary school*. These profiles are completely incommensurable with the target scientific profiles governed by Newtonian theory in mechanics and classical electro-dynamic theory in DC circuits. Student profiles are actually governed by the sort of naïve paradigms that dominated the pre-Galilean era. Unlike modern scientific paradigms, student naïve paradigms are somewhat positivist and driven by implicit tenets that: (a) heavily rely on direct sense perception of secondary features of the world rather than on the Galilean quest for primary features that are beyond direct sense perception, and that (b) erroneously consider that scientists' quest for objective reality is governed by the same socio-cultural paradigms that govern everyday life of ordinary people. Qualitative aspects of all paradigms in question are discussed in ample details in my publications cited throughout this report, and are beyond the scope of this summative report. What this research shows in their respect will be the object of subsequent publications as discussed in the last section of this report.

3.3 WHERE DO LEBANESE STUDENTS STAND RELATIVE TO THEIR U.S. PEERS

Phase 3 of this research made it possible to compare Lebanese students' profiles to their U.S. peers. The same battery of instruments used in Phase 4 of this project was used to ascertain student profiles attending particular freshman physics courses at a U.S. university. As argued in section 2.1: (a) USM1 students' background in physics is very close to the entry level of Lebanese Grade 10, secondary school students, (b) USM2 and USE students' background is very close to the entry level of Lebanese Grade 11 students. Age-wise, participating U.S. students are about the same age as our Lebanese Grade 12 students or one year older. Their background in other science disciplines is typically similar to their physics background (or lack of such background).

All three U.S. groups of students learned physics following a traditional approach similar to the one still followed in Lebanese schools. Students in each group attended large hall lectures (100-150 students hall capacity) for three 50-mn periods a week, and some attended small groups "recitations" (about 24 students each) of one weekly period dedicated to problem solving. Some have also attended a 2-period weekly laboratory session that engages small groups of students (about 24 students) in traditional physics experiments. Research had shown that the extra traditional recitation and laboratory sessions have little impact on enhancing student understanding of materials covered in lectures (Halloun & Hestenes, 1985a; Hestenes et al., 1992). Thus, all-in-all, participating U.S. students are not at a pedagogical advantage relative to their Lebanese peers. In fact, by the time they finish their secondary school years, Lebanese secondary school students would have taken many more physics and other science courses than their U.S. college peers. In this respect, and in terms of covered content in physics and other science courses, Lebanese Grade 12 students would, in principle, subsequently be at an advantage relative to all their U.S. peers. U.S. students may be at a slight advantage relative to their Lebanese peers regarding the language of education when it comes to English.

Table 3

Performance of U.S. college students

Course	Test	Pretest Mean (Max score %)	Posttest Mean (Max score %)	Gain (Possible max %)
USM1	IBC - Mechanics	8.48 (26)	12.93 (39)	4.46 (18)
USM2		12.21 (37)	16.94 (51)	4.73 (23)
USE	IBC-DC	12.58 (38)	---	---
USM2	VASS	---	150.65 (67)	---

Table 3 shows Phase 3 results for the three groups of U.S. students. As indicated in §2.3, USM1 and USM2 students took IBC-mechanics as pretest and posttest. USE students took IBC-DC circuits as a pretest only. Students in USM2 but not the other two courses took VASS, and only as a posttest.

IBC-Mechanics pretest mean of USM1 students indicates that these students begin their mechanics course slightly better than where Lebanese Grade 10 students end their course. Notwithstanding the age difference, this is a troublesome situation given the fact that the two groups of students begin their mechanics courses with little physics background. Pretest score for such students is a measure of basic conceptions students normally develop about moving objects in their everyday life and not in formal schooling. The IBC-Mechanics posttest score of Lebanese Grade 10 students being slightly smaller than the pretest score of USM1 students, this leads one to believe that Grade 10 physics course in Lebanon seems to have no impact whatsoever on Lebanese students enrolled in this course. The situation becomes gloomier when we compare IBC-Mechanics posttest mean scores of Lebanese Grades 11 and 12 students to the respective USM1 posttest mean score or USM2 pretest mean score. Not only Lebanese students fall significantly behind their U.S. peers, but they seem not to retain what they learn by the end of one grade to the beginning of a subsequent grade.

Pretest mean of USM2 students on IBC-Mechanics is close to USM1 posttest mean on the same test, which is to be expected since USM2 students would have taken, either in high school or in college, a course similar to USM1 before enrolling in USM2. By comparison to their Lebanese peers (Grades 11 and 12 outcomes in Table 2), U.S. students seem then to better retain what they have learned in previous physics courses and build on their background to have a better posttest performance on IBC-Mechanics as they move from one grade/course level to the next. Though not to the satisfactory level expected in efficient courses (Hake, 2004), U.S. students' pretest-posttest gain relative to the average maximum possible gain is significantly better (more than three times better) than the gain of their Lebanese peers on IBC-Mechanics. This is another indication of how bad the situation is at Lebanese schools. Participating U.S. students learn physics following a traditional approach of lecture and demonstration, just like their Lebanese peers. They are not engaged in active learning inside the classroom, and their curriculum by no means follow modern educational trends, which the Lebanese curriculum claims to follow, whether in content or in instructional approach.

The situation is about the same with IBC-DC circuits. The pretest score on this test of American students is about the same as the posttest score of their Lebanese peers in Grade 11. Like their Lebanese peers in Grade 10, U.S. students come to courses like USE with no physics background, especially not in electricity. The latter students would have picked up what they know about DC circuits mostly from everyday life experience and not from formal schooling. Yet, these students' knowledge is slightly better than, or about the same as, knowledge Grades 10 and 11 that Lebanese students develop after schooling and reflected in posttest scores in Table 2. Added to the insignificant difference between Grade 10 and Grade 11 results in Lebanon, this further shows how deficient Lebanese schools are in meeting the expectations originally set in the Lebanese physics curriculum.

Research cited above has long shown that everyday life confrontation with any of the issues addressed in either IBC leads ordinary people to develop alternative conceptions about mechanics and electricity that are often governed by a naïve realism that is at odds with scientific realism. Research also shows that, unless explicitly directed to reconsider their naïve paradigms: (a) the more naïve thinkers think about any particular issue, the more naïve conceptions they would develop about that particular issue, and that (b) the more their naïve paradigm is at odds with the respective scientific paradigm, the more deeply seated the naïve paradigm gets in time, and the bigger the gap gets between the two paradigms. Participating students' answers on VASS, in both Lebanon and the U.S., indicate that rarely these students reconsider their views about the physical world and science, even when confronted in the classroom with situations that contradict these views from a scientific perspective. This explains why students' naïve conceptions revealed in both IBCs resist to change in formal schooling.

VASS results are consistent with IBC results when comparing Lebanese and U.S. students. Their VASS mean score was respectively about 63% and 67% of the total maximum score of 225 points on this instrument (Tables 2 and 3). Such results have been consistent throughout the years in both countries (Halloun, 2001b; Halloun & Hestenes, 1998). On average, VASS results appear to be better than IBC results in both countries. However, given the different nature of the two instruments, and especially the absence of any change in students' views about knowing and learning physics as they move from Grade 10 to Grade 11, and then Grade 12, as shown in Table 2 (and through university in Table 3), is another indication of the failure of traditional physics instruction to help students evolve into the realm of science.

Major deficiency shown in VASS that directly relates to the unsatisfactory results in IBCs is about student learning styles. VASS shows that, in both countries, students' often learn physics passively. They try to memorize scientific definitions and statements, as well as problem solving routines offered in their physics courses without due comparison to their own naïve ideas about the physical world and problem solving. They subsequently fail to realize that a conflict exists between their own naïve paradigms and scientific paradigms, and that they need to regulate such a conflict. Given the significant cognitive disequilibrium between the two states of mind, students complete their physics courses with compartmentalized and dissociated knowledge, one compartment they deem useful only for passing course exams and another for dealing with everyday life situations like the ones ascertained in IBCs. As such, students consider their physics courses to be remotely related to everyday life, and see no need to reconsider their views and conceptions about the physical world in light of what they learn in their physics courses. The call for physics curricula, in Lebanon and the U.S., to help students develop *scientific literacy* thus fall short from getting materialized.

Naïve paradigms of Lebanese students may apparently be more deeply seated than those of their American peers, and instruction at Lebanese schools so far appeared to be more deficient than at U.S. schools when it comes to changing naïve paradigms about motion and electricity. One though should note here that the situation at U.S. schools and universities is far from being what one would had hoped for after all the efforts and money that have been spent on improving the state of science education in the U.S. in the last two decades. Weighing things in terms of such efforts and expenditures, the situation would look worse in the U.S. than in Lebanon. Still, this does not justify or alleviate the deficient Lebanese situation. Our physics curriculum have been changed under the claim to lead to meaningful learning of physics, i.e., to help students significantly evolve from naïve realism to scientific realism. Our research suggests that this is still a far fetched claim!

3.4 WHAT DO RESULTS TELL US ABOUT OUR HYPOTHESES?

In sum, this research shows that our five hypotheses virtually stand as originally formulated, i.e. (changes in italics):

1. Lebanese secondary school students (Grades 10-12) come to their mechanics and electricity courses encumbered with naïve conceptual profiles that are at odds with scientific paradigm.
2. Student conceptual profiles do not significantly evolve in the direction of scientific paradigm after the completion of their mechanics and electricity courses.
3. Curriculum shortcomings are about both conceptions and dispositions, *and the latter stand somewhat more deeply seated than the former*.
4. Conceptions and dispositions are closely related, and affect the evolution of one another.
5. Lebanese students lag behind their international peers, especially in USA, in resolving incompatibilities between their own conceptual profiles and respective scientific paradigms. *Nevertheless, conceptual profiles of American students do not significantly evolve in the scientific direction after the completion of traditional physics courses.*

Perhaps the most significant result was the corroboration of our position about the *basic threshold.* Our research suggests that no student can ever meaningfully learn course materials without first understanding *all materials falling under the basic threshold.* Such materials appear to be the ones covered by our current Inventories of Basic Conceptions. Ample details about the critical thresholds, and other issues, are provided elsewhere (Halloun, 2008).

4. Implications and Dissemination

Above discussion is only meant to provide the kind of sketch typically required in a summative report like ours. Ample details about our research, its outcomes and implications, will be provided in subsequent publications. These include peer-reviewed journal articles and monographs. A detailed picture of student profiles will then be drawn in order to allow concerned educators and policymakers make appropriate judgment about the status of our physics curriculum. Publications will go beyond the scope of the research herein reported to cover broad aspects of assessment and curriculum evaluation and development. A short document about the Authentic Assessment Framework mentioned in this report is already out (Halloun, 2008). More publications will follow.

In addition to these publications, papers about the research will be presented at local and international conferences. One paper has already been presented, in August 2005, at the Summer Meeting of the American Association of Physics Teachers, and another is scheduled, in December 2007, at the Antonine University conference on the role of ICT in higher education (cf. www.halloun.net).

Our battery of instruments is already being used by numerous researchers and educators around the world. Arrangements will be made to allow broader dissemination, and continuous refinement, of the three instruments. VASS has already been revised following this research, and a new 33-item version has subsequently emerged (available at www.halloun.net). Arrangements would include the institution of a dedicated portal/website that allows online administration of the instruments, and safe and confidential data collection and storage. Data would be continuously analyzed and results put, through the website, at the disposal of interested educators and policymakers around the world.

Acknowledgment

Many thanks are extended to Lebanese University in Lebanon, and to the National Science Foundation in USA, for sponsoring in part this research. Special thanks are extended to all those who contributed to this work in both countries, whether colleagues, staff or students. Along with my gratitude is expressed the commitment to pursue this work beyond its current scope, so that all interested stakeholders make the best of it.

Ibrahim A. Halloun

October 14, 2007

REFERENCES

Abell, S. & Lederman, N. (2007). *Handbook of Research on Science Education.*

Aikenhead, G. S. (1987). High-school graduates' beliefs about science-technology-society. III. Characteristics and limitations of scientific knowledge, *Science Education*, *71*(2), 145-161.

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

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

Bachelard, G. (1940). *La Philosophie du Non*. Paris: Quadrige (4th edition, 1994) / Presses Universitaires de France.

Bagno, E. & Eylon, B.S. (1997). From problem solving to a knowledge structure: An example from the domain of electromagnetism. *American Journal of Physics, 65(8),* 726-736.

Bibeau, J. R., Halloun, I., May, J. & Reddy, V. (2002). *Report on the Evaluation of UNESCO's Science and Technology Education (STE) Programme*. Internal report presented to the STE section at UNESCO, Paris.

Cohen, R., Eylon, B.S., & Ganiel, U. (1983). Potential Difference and Current in Simple Electric Circuits: A Study of Students' Concepts. *American Journal of Physics, 51,* 407-412.

Dupin, J.J., & Joshua, S. (1987). Concepts of French Pupils concerning Electric Circuits: Structure and Evolution. *Journal of Research in Science Teaching, 24,* 791-806.

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.

French, A. P. (1989). Learning from the past; looking to the future. *American Journal of Physics, 57* (7), 587-592.

Hake, R. R. (2004). Lessons from the physics-education reform effort (and related Hake's publications cited therein).

Halloun, (2008). *Authentic Assessment Framework. A Quick Reference.* Beirut, Lebanon: Phoenix series / Educational Research Centre.

Halloun, (2007). Mediated modeling for meaningful learning of science. *Science & Education, 16 (6-7),*

Halloun, I. (2004/6). *Modeling Theory in Science Education*. Dordrecht: Kluwer/Springer.

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

Halloun, I. (2001b). *Student Views about Science: A comparative Survey.* (2001). Beirut, Lebanon: Phoenix series.

Halloun, I. (1998a). Lebanon's new precollege science curricula and modern educational theory. *An-Nahar* (14, 15, 17 August 1998).

Halloun, I. (1998b). Schematic concepts for schematic models of the real world: The Newtonian concept of force. *Science Education, 82(2),* 239-263.

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). Schematic modeling for meaningful learning of physics. *Journal of Research in Science Teaching, 33(9),* 1019-1041.

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

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

Halloun, I. A. & Assi, M.M. (1989). Connaissances d'élèves du secondaire sur la charge électrique. *Revue Libanaise de Didactique des Sciences et des Mathématiques, 2(1),* 8-11.

Halloun, I., Hake, R., Mosca, G., and Hestenes, D. (1995). Force Concept Inventory – Revised. Available at: *www.modeling.asu.edu*.

Halloun, I, & Hestenes, D. (1998). Interpreting VASS dimensions and profiles, *Science & Education, 7*(5).

Halloun, I, & Hestenes, D. (1995). Interpreting the Force Concept Inventory. *The Physics Teacher, 33(8),* 502-506.

Halloun, I, & Hestenes, D. (1985a). Common sense concepts about motion. *American Journal of Physics*, *53,* 1043-1055.

Halloun, I, & Hestenes, D. (1985b). The initial knowledge state of college physics students. *American Journal of Physics, 53,* 1056-1065.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30* (3), 141-158.

International Evaluation Association, IEA. (1995, and subsequent years). Third/Trends in International Mathematics and Science Study (TIMSS).

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

Lebanese Government. (1997). *Decree No. 10227.* Beirut: Sader press.

McDermott, L. (1984). Research on conceptual understanding in mechanics. *Physics Today, 37,* 24-32.

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.

Mortimer, E. F. (1995). Conceptual change or conceptual profile change? *Science & Education, 4* (3), 267-285.

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

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.

Rainson, S., Tranströmer, G., & Viennot, L. (1994). Students' understanding of superposition of electric fields. *American Journal of Physics, 62(11),* 1026-1032.

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.

Thornton, R.K. & Sokoloff, D.R. (1998). Assessing student learning of Newton's laws. *American Journal of Physics, 66(4),* 338-352.

UNESCO. (1993). *Project 2000+. Forum International sur la Culture Scientifique et Technologique pour Tous*. Paris: UNESCO.

Viennot, L. (1996). Raisonner en Physique. La part du sens commun. Paris: De Boeck.

APPENDIX

Instruments and their Taxonomy

A. Inventory of Basic Conceptions in Mechanics
B. Inventory of Basic Conceptions about DC Circuits
C. Views About Science Survey