2. ÉVALUATION DES INSTRUMENTS D'ÉVALUATION DE DEUX COURS DU PROGRAMME

Dans le critère d'évaluation 5.2, on nous a demandé de vérifier, pour les cours «Initiation pratique à la méthodologie des sciences humaines» et «Économie globale», la capacité des moyens d'évaluation des apprentissages à mesurer adéquatement et équitablement l'atteinte des objectifs visés.

Pour réaliser cette tâche, nous avons une fois de plus utilisé la taxonomie de Bloom. La précision et l'efficacité de cet outil d'analyse nous a permis d'avancer de façon plus sécuritaire sur le terrain glissant des évaluations de nos collègues de travail.

C'est à partir de ces deux questions que nous avons évalué la concordance entre les outils d'évaluation et les objectifs de ces deux cours:

1. Est-ce que les moyens d'évaluation permettent de mesurer tous les objectifs spécifiques et les contenus planifiés dans les plans de cours?

2. Est-ce que les moyens d'évaluation utilisés permettent de bien mesurer les objectifs des cours? Autrement dit, sont-ils bien adaptés aux niveaux taxonomiques des objectifs?

Pour répondre à la première question, nous avons procédé à l'analyse des moyens d'évaluation de ces deux cours afin de vérifier si les contenus et tous les objectifs de cours sont évalués. Dans cette première étape de l'analyse, nous avons dressé un tableau dont la colonne de droite comprend les questions d'examens et les exigences pour les travaux de recherche. Dans la colonne de gauche, on retrouve le ou les objectifs se rapportant aux questions d'examens et aux exigences des travaux. Un tel tableau nous a permis de déterminer précisément les objectifs des cours qui étaient couverts par les évaluations et ceux qui ne l'étaient pas. Une fois cette identification réalisée, le professeur peut facilement s'ajuster pour s'assurer d'évaluer tous les objectifs de son cours.

En ce qui concerne la deuxième question, nous avons fait l'analyse de la concordance entre les moyens d'évaluation utilisés dans ces cours et le niveau taxonomique des objectifs poursuivis. Pour vérifier cette concordance, nous avons réalisé l'analyse taxonomique des questions des examens et des exigences des travaux de recherche pour ces deux cours. Nous avons alors été en mesure de constater si le niveau de difficulté des évaluations de ces deux cours est équivalent à ce qui est prévu par les objectifs ministériels.

Une telle analyse nous a donc permis de vérifier de façon précise et rigoureuse si les instruments d'évaluation mesurent l'ensemble des objectifs spécifiques des cours et s'ils respectent les niveaux taxonomiques des objectifs de ces mêmes cours.

CONCLUSION

Lors de l'évaluation du programme de sciences humaines au Centre collégial de Mont-Laurier, la question de la méthode s'est posée à plusieurs reprises. Soucieux de produire un rapport tendant le plus possible vers l'objectivité, nous avons utilisé la taxonomie de Bloom pour une bonne partie de l'analyse. Cet outil de travail nous a permis de recueillir des informations de qualité qu'un traitement intuitif ne peut fournir. Mais surtout, cette méthode d'analyse nous a permis de corriger certaines faiblesses dans la mise en œuvre de notre programme d'étude.

ASSESSMENT AND LEARNING IN PHYSICS

LESIE O. DICKIE, JOHN ABBOTT COLLEGE

ABSTRACT

This exploratory study determined the intellectual demands of quizzes, tests and final exams in physics using a scheme derived from Bloom's taxonomy.

It was found that the majority (70%) of assessment items required routine problem solving, while 28% required comprehension. The grade assigned to items requiring comprehension increased from Mechanics 101 (19%) to Electricity and Magnetism 201 (28%) to Waves and Optics 301 (32%).

The study also explored the relationships between the intellectual demands of assessment and the performance of the students. The students in the study wrote the Study Process Questionnaire and the Force Concept Inventory in Mechanics classes at the start of their first semester. Students who proceeded to Electricity and Magnetism rewrote the measures at the end of their second semester. The findings show that most incoming students approach physics with the intention of memorizing formulae rather than understanding concepts. They adopt surface or surface-achieving approaches. The approach to learning adopted by students was found to be related to the intellectual demands of the examinations, to the students' performances on the final examinations, and to their prior knowledge of the concept of force.

ASSESSMENT AND LEARNING IN PHYSICS.

The introduction of the new science program with its emphasis on competencies rather than content has meant that a number of issues relevent to assessment must be re-examined. There are close connections between the cognitive demands of learning tasks and two of the competencies; critical thinking and problem solving. This study looked at the cognitive demands of assessments in physics at three English language Cegeps and at the relationships between the cognitive demands of the assessments and performance in physics at one of these cegeps.

One of the most important and controversial issues in contemporary education is that of assessment: the assessment of student learning and the impact of assessment on student learning. When students enter their classrooms, they look to the teacher for guidance about what to learn and how to learn and, rightly or wrongly, they see the tests and other assessments as indicators of what the teachers consider to be important. Indeed the primary concern voiced by most students facing a learning task is. "Is this going to be on the test?" After reviewing over 200 studies of the impact of classroom evaluation, Crooks (1988) concluded that assessment guided the student's judgment of what it was important to learn, and affected their motivation and approach to studying; that is the how of their approach to the learning task. If the test focuses on factual knowledge, the student will learn to memorize; if the test requires analytical thinking the student will learn to reason analytically. The intellectual skills the students rehearse will depend on the cognitive demands of the tasks they are asked to undertake.

Once the teachers' expectations have been communicated the students can decide if they want to study and what learning strategies they want to use. The combination of strategy and motivation is called the approach to learning of the student. Three approaches have been identified; surface, deep, and achieving (Ramsden, 1991). In a surface approach the student focuses on memorizing to obtain a passing grade, a deep approach involves an intention to understand the material, while in an achieving approach a student adopts deep or surface level strategies according to what he or she judges to be most efficient for obtaining grades (Biggs, 1987). While students can control the approach to learning, they are just a part of a larger system. The boundaries of the system are set in part by the institution, in part by the participants' perceptions of one another, and in part by the habits and practices of both teacher and student (Bhushan, 1991; Brekelmans, Wubbels, and Créton, 1990; Roth 1994). Within these boundaries are many complex interactions that influence the quality of learning: one of these is the interaction between the student's motives and strategies and the assessment practices of the teacher.

The purpose of the present study was threefold: first to determine the approach to learning of students in physics classes, as measured by the Study Process Questionnaire (Biggs, 1987); second to determine the intellectual demands of final examinations, tests, and quizzes in physics, using a scheme based on Bloom's Taxonomy (Bloom, 1956); and third to search for relationships between these two variables, as well as the impact on the performance and persistence of the student. The performance of the students was measured by their grades in the physics final exams, and by their understanding of the concept of force as measured by their score on the Force Concept Inventory (Hestenes, Wells, and Swackhamer, 1992).

METHOD

This study was conducted in a two year cegep in Québec where students graduate from high school after grade 11 and those seeking further education enter a Cégep (Collège D'Énseignement Général et Professionnel) for two years before going on to university at what would be in the rest of North America their second year. The language of instruction was English. The first sample for the study consisted of 107 first semester physics students who completed the measures detailed below at the start of the year, followed a one-semester calculus-based mechanics class and wrote the final exam. The students were selected by having all students enrolled in six of the fourteen sections of the mechanics course complete the Force Concept Inventory in the first week of classes and the Study Process Questionnaire in the second week of classes. Out of a total of 670 students who were registered in the course, 267 wrote the first of the measures, the FCI. The different sections were taught by different instructors, but the assignments and final exam were common. The second sample consisted of the 35 academically on-track students who, after passing the mechanics course, followed the subsequent electricity and magnetism course, wrote the final exam and repeated the measures at the end of the academic year. This second group was asked to write the measures outside of class time: a small honorarium was paid to these students. This second sample comprised 78% of the on-track cohort.

Measures

The Study Process Questionnaire (SPQ) is a 42-item group-administered instrument (Biggs, 1987). Each item consists of an affirmative self-report statement that describes a student's strategy or motive. An example of a motive statement is " I find that at times studying gives me a sense of deep personal satisfaction." An example of a strategy statement is "I summarize suggested readings and include these as part of my notes on a topic." After consultations with students and teachers, the wording of the questionnaire was changed to better conform to common usage in the cegeps (e.g. tertiary to post secondary, lecturer to teacher, rote to learn by heart, ...). For each item of the questionnaire, the student responds on a five-point Likert scale. Seven items of the questionnaire are constructed to reflect each of the sub-scales: surface, deep, and achieving motivation, and seven items to reflect each of surface, deep, and achieving strategy. The sub scale scores are combined to give three-scale scores: Deep, Surface and Achieving Approach. The scale scores were used as dependent variables to identify a student's approach to learning, and as independent variables to assess changes in the approach to learning over the one-year period of the project. Students in this study were in the first year of the pre-university science stream at cegep, therefore results were analyzed using the norms for the instrument given by Biggs for science university students.

The Force Concept Inventory (FCI) is a twenty-nine-item multiple choice questionnaire that was used to determine a students prior knowledge of the concept of force and the agreement between the student's understanding and the Newtonian understanding (Hestenes, Wells, and Swackhamer, 1992 see also Heller and Huffman, 1995; Hestenes and Halloun, 1995; Huffman and Heller, 1995;). The score on the inventory was used in this study as a dependent variable. The questionnaire was administered at the start of the study as a measure of the students' prior knowledge, and again at the end of the academic year to those students who had continued in physics as a measure of the change in their understanding. Because the score on the inventory is a measure of the students' understanding of the concept of force rather than of their ability to apply formulas, relationships between the student's score and the student's approach to learning were also sought.

The Cote Finale (science) was used as a measure of the high school performance of the student. It is a weighted average of a student's high school grades for grades ten and eleven and is calculated for students who took high school in Québec: it is used in determining admission to the cegep.

Coding Scheme For Assessment Items

The intellectual demands of a learning task in physics are defined by the answers students are required to produce and the routes that can be used to attain these answers. In undertaking an analysis of the intellectual demands of the learning tasks, one is not looking at the physics content that is being asked for, but at the behaviors and processes that are being required. There have been a number of descriptive frameworks offered to classify and identify the intellectual demands of objectives and/or assessment items (Biggs, 1991; Donald, 1985; Doyle, 1983; Gagne, 1977; Lawrence et al 1994; Merrill and Tennyson, 1977). In physics, a number of authors have suggested different approaches to classifying the demands of learning tasks; for example Klatt (1991) used visualization and the implicit geometrical content while Niaz and others have used the limits on information processing imposed by working memory or M-space (Niaz, 1993; Roth, 1991). However, the most widely used classification scheme for intellectual tasks is that developed some forty years ago at the University of Chicago by a team under the editorship of Benjamin Bloom (Bloom, 1956). Blooms Taxonomy, as it is widely called, was developed in part to assess the demands of course objectives, and in part to examine the demands of examinations. It was developed in a pragmatic fashion from the ideas of several working groups and has come to be widely accepted, particularly as it has been modified to better suit the demands of different disciplines (Krathwohl, 1994). Its constructs have been applied to classifying learning tasks in physics by a number of authors (Aubrecht, 1990; Ferris, 1960; Crooks and Collins, 1986).

Bloom's taxonomy has six hierarchical levels; knowledge, comprehension, application, analysis, synthesis, and evaluation. Preliminary work showed that the examinations and other assessment items of this study involved the three lowest levels of Bloom's taxonomy, but that the definitions of these levels did not adequately represent the intellectual demands of assessment items in physics. Given the predominance of problem solving as both a teaching and an assessment methodology in physics, it was necessary to consider the place of problem solving in the taxonomy and in particular two linked issues; novelty versus rehearsal, and rote application of an algorithm versus understanding. The issue of whether a problem was novel was addressed by examining quizzes and tests which confirmed the initial hypothesis that almost all problem types were rehearsed. The second issue was rote application versus understanding. When a student was applying a problem solving algorithm to the solution of a typical or "textbook type" problem, was the student reflecting on each step and understanding why choices were being made or were they just following a well worn path? A path that the student had seen demonstrated in class or had rehearsed as assignments were completed. The consensus of discussions with teachers was that it was generally the latter. Accordingly, it was decided to adopt the point of view of Doyle (1983), and of Lawrence et al (1994) and place Comprehension after routine problem solving. This is somewhat at odds with Bloom's statement, "If a student really comprehends something he can apply it" (1956, p. 120), because it accepts that the converse of this statement is not necessarily true. Deciding whether a student has understood a procedure could be resolved by interviewing the student, but examination of their correct written answer to a typical exam question is unlikely to reveal whether they were following an algorithm, or had an understanding of why the procedure was appropriate and successful.

In the taxonomy that was developed for this study, the first category, Memory, demands the recognition or reproduction of information previously encountered. Bloom considers that such tasks do not require thinking and distinguishes this level from other intellectual tasks; tasks that require some content to act on. For example, one does not just think, one thinks about projectile motion.

In physics there are many short, routine problems requiring little thinking or understanding. Bloom places such operations as part of knowledge but they are so common in physics that a separate category, Procedural/Algorithmic, is warranted. Similar distinctions in different disciplines have been made by Doyle (1983), Lawrence et al (1994), and McGuire (1963).

In the third category, Procedural/Comprehension, the student has to make choices, has to make decisions and judgments about what procedure to follow, and may even have to carry out some analysis or form an opinion - but at a straightforward level. The problem is not novel. In discussions with cegep physics teachers, it was agreed that placing this level before Comprehension was appropriate and proper, a decision also made by Lawrence et al, (1994) in classifying mathematics items.

The label comprehension is used by Bloom, by Doyle, and by Lawrence to describe similar but slightly different levels of understanding. Bloom talks about transforming information to demonstrate that it has been *understood and* about applying a formula when its use is specified. Doyle defines it in terms of recognizing transformed information, and also in terms of choosing between several procedures in solving a problem. Lawrence talks about understanding the "gist" of a problem, the how and why procedures are used. In the present work, the four categories of abilities defined by Reif, Larkin and Brackett (1976) as constituting understanding of a relation are used as the basis for the working definition of comprehension. These include the transformation abilities of Bloom or Doyle, for example, translating from a table of values to a graph, or interpreting and using the information given in a graph. The understanding of why a problem-solving procedure works places this after procedural/comprehension.

In adopting these levels, the cognitive demands of problems have been split into four levels: memory, rote application, those requiring limited comprehension, and those requiring understanding of principles rather than just the demonstration that an algorithm can be applied. The working definitions of the levels are given in table I.

Table I
Taxonomy for Coding the Intellectual Demands of Physics Assessment Items

Memory:	Recalling information much as it was learned.
Procedural / Algorithmic:	Following a routine series of steps in solving a problem. The problem is familiar and the rule or formula is either given in the problem or very familiar from previous rehearsal.
Procedural / Comprehension:	Solving a problem that requires that choices be made about which rule or formula to apply based on the information given in the problem.
Comprehension:	Recognize transformed or paraphrased information. Draw inferences from previously encountered information. In applying a rule or formula, demonstrate understanding of when, why, and how the relation can be applied.

The primary data for this part of the study were the final examinations, term tests and in-class quizzes for the three one-semester physics courses, Mechanics , Electricity and Magnetism, and Waves and Optics, given at the cegep. The researcher and the two coders (who were experienced physics teachers from cegeps other than that of the researcher) reviewed all the exams and other assessment items to determine the novelty or otherwise of the questions of the assessments. In all, thirty quizzes, fifteen term tests, and ten final examinations, totaling 710 items were coded.

In addition, the final examinations and term tests of the three physics courses from two other anglophone cegeps were coded.

Coding Of Items Of The Quizzes, Tests, And Final Exams
The process of assigning a final code to each question or part question followed three steps. First, solutions to each question were prepared by the researcher. The majority of these solutions were annotated with the thinking processes that were followed. In solving and annotating the questions, the researcher was an experienced physics teacher trying to act like a novice. Second, after a training session, the coders were provided with a detailed rational for the coding scheme together with examples, and two final exams were coded. The team then met to discuss the scheme and to compare codes. The coding scheme was adjusted and finalized. Then the coders and the researcher independently coded all items.

In assigning a final code to an item, if the two coders a greed on the code, this became the final code. The rate of agreement was 72%. If there was no initial consensus, the code assigned by the researcher was considered. If it agreed with one of the codes assigned by the coders, this became the final code. This was the case for all but 16 cases. The remaining 16 items were discussed and final codes assigned. The chief cause of disagreement was between Memory and Comprehension. For example; was an electric field diagram, or the derivation of a simple harmonic motion formula, remembered or understood?

RESULTS

Approach to Learning
The procedure suggested by Biggs (1987) to determine the approach to learning of the students from the SPQ scores showed that most students entered the cegep adopting a Surface (12%) or Surface Achieving (29%) approach to learning physics, compared with just 5.6% who adopted a Deep approach. When Biggs' procedure was used to examine the changes in approach of the cohort of 35 who were academically on track after two semesters, it showed that for these students-the percentages adopting both . Deep and Surface approaches increased and the percentage adopting the more opportunistic Achieving Approach declined. The complete results are given in Table 2.

Table 2
Approach to learning
Percentage Of Students Classified As Adopting A Particular Learning Approach:

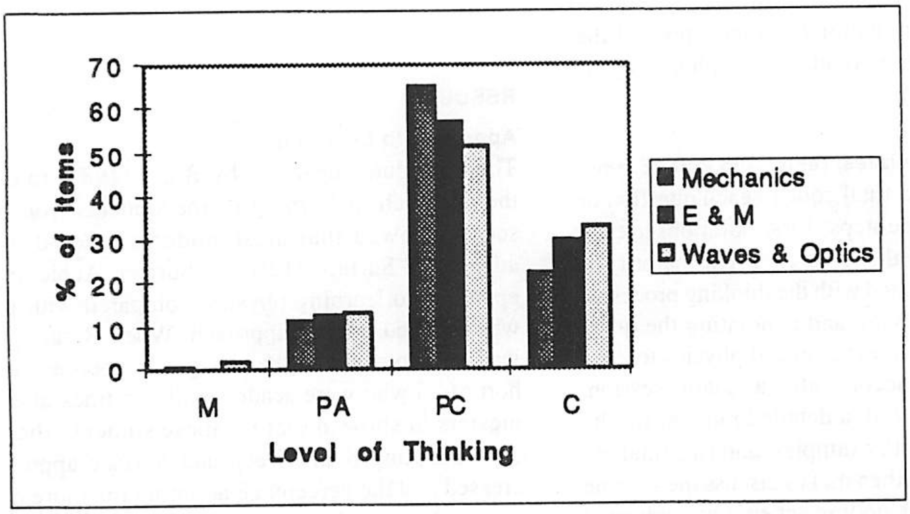
	Surface	Deep	Achieving	Surface Achieving	Deep Achieving	Un- determined
All Incoming Students at start of first semester (n = 107)						
	12	5.6	14	29	12	27
On-track cohort at start of first semester and end of second semester (n = 35)						
Pre	5.7	2.8	31.4	28.6	17.1	14.3
Post	14.3	20	2.8	11.4	14.3	37

Intellectual Demands of Assessment

The coding scheme developed classifies the level of thinking required by the assessment tasks into four hierarchical levels; Memory, Procedural/Algorithmic, Procedural/Comprehension and Comprehension. There was a steady increase of the demands made of students as they went from the first to the last course as is shown in Figure 1. In all, 253 items were coded for the Mechanics course, 236 items for the Electricity and Magnetism course, and 221 items for the third course, Waves and Optics; the percentage of items coded at the highest level, comprehension, increased from 22.9%, to 30% to 33% from the first to the third course. When the results were analyzed in terms of the grade assigned to an item, the overall pattern was maintained and it was found that teachers assigned more credit to items of greater difficulty.

Figure 1

Percent Of Items By Course And Level Of Thinking



Percent Of Items By Course And Level Of Thinking
It was found that the intellectual demands of the assessment items asked at the other two cegeps were similar to those asked at the cegep where the study took place.

Relationships Between Approach to Learning and Intellectual Demands of Assessment
The Final examination in a course accounts for 40% of the final grade, the term tests for 30%, the quizzes for just 10% and the laboratory reports for the remaining 20%. A step-wise linear regression of the Mechanics final exam grade (the dependent variable) with the scale scores of the SPQ, the Cote Final, and the FCI (the independent variables), showed that high school performance, as measured by the Cote Final, was more strongly correlated to performance than the other measures and that both the FCI and Achieving Approach scores were positively correlated while the Deep Approach score had a negative correlation. High school performance was also more strongly correlated to performance on the Electricity and Magnetism exam than the other measures. A step-wise linear regression between the Cote Final, the FCI, and the scale scores of the SPQ (both measures re-written at the end of the second semester) and the Electricity and Magnetism final exam grade showed that the Cote Final, the score on the FCI and the Achieving Approach score were positively correlated to performance, and Surface Approach score was negatively correlated with performance.

There was a clear relationship between intellectual demands of the exam, performance on the exam, and approach to learning. The final mechanics exam in the fall of 1993 had a much lower percentage of its grade determined by more intellectually demanding questions because the teachers who set the exam responded to a change in imposed admission criteria by setting a more routine exam than normal. Analysis of this exam showed that it had just 6.6% (cf. a more usual 20%) of its grade coming from items coded at the Comprehension level and there was a negative correlation between performance and Deep Approach. In contrast, the winter 1994 Electricity and Magnetism exam had a much higher (and usual) percentage, (28.4%) of its grade coming from items coded at the Comprehension level and there was a negative correlation between performance and Surface Approach. Results are shown in Table 3.

Table 3

Step-Wise Linear Regression Exam grade and Measures

Mechanics Exam Grade.
Measures Written at the Start of the First Semester
Squared Multiple $R = 0.948$ $(n = 107)$

Variable	Std Error	Std Coef	T	p
Cote Final	0.175	0.821	5.65	.000 **
Force Concept Inventory	0.608	0.275	3.67	.000 **
Deep Approach	0.392	-0.358	-2.37	.019 *
Achieving Approach	0.349	0.244	1.65	.102

Electricity And Magnetism Exam Grade
Measures Re-Written At The End Of The Second Semester
Squared Multiple $R = 0.957$ $(n = 35)$

Variable	Std Error	Std Coef	T	p
Cote Final	0.074	0.985	3.78	.001 **
Force Concept Inventory	0.177	0.313	2.93	.006 **
Surface Approach	0.142	-0.697	-2.84	.008 **
Achieving Approach	0.134	0.382	1.76	.089

RELATIONSHIP BETWEEN APPROACH TO LEARNING AND SCORE ON THE FORCE CONCEPT INVENTORY

For the incoming students who wrote the FCI and the SPQ at the start of their first semester, a Tukey HSD multiple comparison showed that the group of students identified as adopting a Deep Approach scored significantly higher on the FCI than those identified as adopting a Surface ($p < .01$), Achieving ($p < .001$), Surface Achieving ($p < .005$), or Deep Achieving ($p < .005$) approach. These relationships are consistent with the aims of the two measures. A Deep Approach indicates an intention to apply principles to real world situations as opposed to applying formulas in a rote fashion: the FCI is known to test conceptual understanding of Newton's laws and the ability to apply the concepts to real world situations. The mean score of the group of students identified as adopting a deep approach was very high, 70%. The significance of this score can be gauged by the reported average pre-instruction scores reported by Hestenes, Wells and Swackhamer (1992), of between 34% and 52% for students entering university, and post-instruction averages of between 63 and 68%, and in the case of a post-test conducted in a class consisting mostly of physics majors at Harvard, 77%. The SPQ is identifying those incoming students who have very good understanding of Newton's laws. The results are shown in table 4.

Table 4

Means And Standard Deviations For Score On Force Concept Inventory By Approach To Learning. Measures Written at the Start of the First Semester (n = 107)

	Approach				
	Surface	Deep	Achieving	Surface Achieving	Deep Achieving
	n=13	n=6	n=15	n=31	n=13
FCI (%)	46.4	70.1	46.9	47.1	48.7
S.D.	10.9	10.8	13.0	14.2	13.8

Analysis Of Variance For Approach To Learning And FCI Score

	df	SS	MS	F	p
Between groups	5	349.5	69.90	4.08	.002
Within-groups	101	1729.1	17.12		

DISCUSSION

This study found that incoming physics students approach physics with the intention of memorizing formulae rather than understanding concepts, they adopt Surface or Surface-Achieving Approaches. If instruction is to be effective, it must be aware of, and understand, the preconceptions students hold about both the content of physics and the learning task in physics. How should teachers use this information to counsel students and to guide the form of instruction to better match instruction to the beliefs and practices of incoming students, and to the goals they, as teachers, consider desirable?

At this cegep, as at many other colleges and universities, the course, which was taught by a number of teachers, had a common final examination, mid-term tests, and a number of quizzes and assignments. Final examinations, tests and quizzes from previous semesters were available in the library. While the course syllabus provides an overview of the course, most students look to the exam and other assessments to determine the aims and methodologies of the teachers.

However, teachers must respond to many influences. The students in this study entered the cegep as a new science program, and new admission criteria were being imposed. The instructors were very concerned with the demands the course made on the students both in terms of the amount of material covered and in terms of the level of difficulty of the course. Their solution was to develop a final examination that was, in their view, straightforward. The cognitive demands of the final examination were low. Over 90% of the questions involved routine problem solving and just 6% required comprehension. Seventy-seven percent of those who persisted and wrote the mechanics final examination passed and there was a negative correlation between Deep Approach and grade in the final exam. When the on-track students entered the subsequent electricity and magnetism course, they were faced with what many consider to be the most challenging of the three physics courses. The final examination that they wrote in this course had a much higher fraction (34.8%) of items that required comprehension. There was a negative correlation between grade in this exam and Surface Approach.

The goal of most students is to pass, and students adapt and adopt practices that they hope will ensure their success, however, success in a course does not ensure understanding of the material covered. It is accepted, albeit with some controversy, that the Force Concept Inventory measures understanding of the Newtonian concept of force (Heller and Huffman, 1995; Huffman and Heller, 1995; Hestenes and Halloun, 1995). Over the two semesters, the students who persisted showed a gain of 12% in their score on the FCI. This increase is in agreement with that reported in the literature for conventional physics courses, but is much less than has been achieved by more interactive courses (Hake, 1994). Such courses engage the student in tasks that require active participation and the use of higher level thinking skills.

Over twenty-five years ago, Rogers (1969), in talking about physics, pointed out that learning will be sabotaged if the final exam asks for numbers to be put into memorized formulae even if the classes are dynamic, the demonstrations intriguing, and students are forcefully exhorted to "understand the physics." The findings of this work suggest that the intellectual demands of assessment tasks influence the approach to learning adopted by students. Classroom assessment guides learning. A majority of the questions in the quizzes, tests, and final examinations required problem solving with some limited understanding of the principles and concepts. It would be possible to pass the courses without understanding the concepts. One can infer that the students are acquiring content knowledge but must ask if they are able to apply this knowledge to complex, unfamiliar situations.

A limitation of this study was that the cognitive demands of laboratory work, were not addressed. In carrying out laboratory work were students required to design an experiment or follow a cookbook? In writing a report, were students required to complete a table or synthesize data, theory and results? If we are to fully determine the thinking skills developed by current instructional practices, then the intellectual demands of laboratory work must be determined. Finally, while this work has shown the relationships that exist between assessment and approach to learning for students who persist and succeed, it does not answer directly the questions that arise about the relationships between approach to learning and assessment for those who drop out and fail. Such a study could help teachers give appropriate guidance, and design appropriate instruction to help those who currently fail.

What is known is that involving students with the tasks rather than encouraging silent listening or repetitive calculations does achieve increased understanding. However if time is to be devoted to allowing students to grapple with the ideas, then the content covered must be reduced. (However, we must ask if the content was covered by the student or by the teacher.)

A constant debate among cegep physics teachers is what topics to include and what to omit as they see themselves squeezed between the high schools and the universities and buffeted by the changes in curriculum and course structures dictated by others. Many traditional practices of teaching and assessment are no longer appropriate for the diverse population that fills present day physics classrooms. The background, outlook, and needs of students have changed. Society no longer accepts without question the value of physics as an intellectual discipline and as a subject that can provide solutions to societal problems. Faced with these challenges, physics teachers must re-examine their teaching and assessment methodologies and adopt strategies that will encourage meaningful learning of the mix of content and process they, the teachers, consider appropriate. What changes, if any, to current methods of teaching and assessment will ensure that students combine knowledge of current content and concepts with the ability to apply these in meaningful ways, and the ability to adapt to as yet unknown challenges and ideas?

REFERENCES

Aubrecht, G. J. (1990). Is there a connection between testing and teaching? Journal of College Science Teaching, 20(3), 152-157.

Bhushan, V. (1991). Learning environments and teacher attitudes in French-speaking Canada. In B. J. Fraser & H. J. Walberg (Eds.), Educational Environments: Evaluation, Antecedents and Consequences (pp. 245-254). Oxford: Pergamon Press.

Biggs, J. B. (1987). Student Approaches to Studying and Learning. Hawthorn, Victoria: Australian Council for Educational Research.

Biggs, J. B. (1991). Enhancing learning in the context of school. In R. F. Mulcahy, R. H. Short, & J. A. C. Andrews (Eds.), Enhancing Learning and Thinking (pp. 35-52). Westport, CT: Praeger.

Bloom, B. S. (Ed.). (1956). Taxonomy of Educational Objectives. The Classification of Educational goals. Handbook I: Cognitive Domain. New York: David McKay.

Brekelmans, M., Wubbels, T., & Créton, H. (1990). A study of student perceptions of physics teacher behavior. Journal of Research in Science Teaching, 27(4), 335-350.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. Review of Educational Research, 58(4), 438-481.

Crooks, T., & Collins, E. (1986). What do first year university examinations assess? New Zealand Journal of Educational Studies, 21(2), 123-132.

Donald, J. G. (1985). Intellectual skills in higher education. Canadian Journal of Higher Education, 25(1), 53-68.

Doyle, W. (1983). Academic work. Review of Educational Research, 53(2), 159-199.

Gagne, R. M. (1977). The Conditions of Learning. New York: McGraw Hill.

Heller, P., & Huffman, D. (1995). Interpreting the force concept inventory. The Physics Teacher, 33, 503, 507-511.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. The Physics Teacher, 30, 141-153.

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory: a response to march 1995 critique by Huffman and Heller. The Physics Teacher, 33, 502, 504-506.

Huffman, D., & Heller, P. (1995). What does the force concept inventory actually measure? The Physics Teacher, 33(3), 138-143.

Klatt, E. (1991) An Analysis of the Implicit Geometry Content of Physics, Ontario Academic Course (OAC). Master of Education, McGill University.

Krathwohl, D. R. (1994). Reflections on the taxonomy: Its past, present, and future. In L. W. Anderson & L. A. Sosniak (Eds.), Bloom's Taxonomy: a forty year retrospective (pp. 181-202). Chicago: National Society for the Study of Education.

Lawrence, J. H., Hart, K., Kingan, M., & Campbell, J. (1994 April). Estimating the intellectual demands of college courses through task analysis. In Annual Meeting of the American Educational Research Association, New Orleans.

McGuire, C. (1963). A process approach to the construction and analysis of medical examinations. Journal of Medical Education, 38, 556-563.

Niaz, M. (1993). Working memory, mental capacity and science education: towards an understanding of the 'working memory overload hypothesis'. Oxford Review of Education, 19(4), 511-525.

Ramsden, P. (1991). Study processes in grade 12 environments. In B. J. Fraser & H. J. Walberg (Eds.), Educational Environments: evaluation, antecedents and consequences (pp. 215-229). Oxford: Pergamon.

Reif, F., Larkin, J. H., & Brackett, G. C. (1976). Teaching general learning and problem solving skills. American Journal of Physics, 44(3), 212-217.

Roth, W.-M. (1991). Factors in the development of proportional reasoning strategies by concrete operational college students. Journal of Research in Science Teaching, 28(6), 553-566.

Rogers, E. M. (1969). Examinations: Powerful agents for good or ill in teaching. American Journal of Physics, 37(10), 954-962.

Roth, W.-M., & Roychoudhury, A. (1994). Physics students' epistemologies and views about knowing and learning. Journal of Research in Science Teaching, 31(1), 5-30.