

EVALUATING COMPETENCIES: THEORY AND PRACTICE



NICOLAS FAUCHER
Biology Teacher
Cégep Limoilou

TESTING THE TOOLS IN A BIOLOGY COURSE OF THE NURSING PROGRAM

Even when I was a student, the question of evaluation appeared to me to be a fundamental issue in teaching. In fact I had conducted a workshop on it at the 14th AQPC Symposium in Québec City in June 1994¹. Since then I have refined my thoughts on the subject but my preoccupation has remained the same.

FROM TRADITIONAL EVALUATION...

When I started teaching at the college level in 1998, I had to adopt a position as to how I should construct my professional practice. I arrived with the competency-based approach, and my first concern was to identify the impact that the competency-based approach would have on my teaching practice.

From the outset I felt comfortable with the concept of a “competency” which I will define simply as a complex ability to perform in a given context. Certainly this concept is comprehensive and has subtle nuances that require a more detailed explanation; but, since this is not the specific object of this article, I will not elaborate any further here. Having said that, the idea that one ought to assume responsibility for the performance of one’s students seemed obvious to me. However, as easy as it may be to talk about such grand ideals, concrete practice reminds us of the need for an effective pragmatic approach. If, as I understand it, the purpose of evaluation is to provide information regarding the level of understanding achieved by the learner, then it is also important that evaluation be viable as much for the students as for me, the teacher. Therefore I attempted to

construct my evaluation practice within a world of possibilities as defined by the many practices I had been able to observe². Such a task proved to be impossible and, in spite of my convictions regarding evaluation, I quickly resigned myself to what seemed to be an acceptable compromise between principles and practice.

This is how, in practice, my evaluations became, for the most part, realistic situations of varied complexity that required students to provide answers with more or less elaborate explanations. The overall judgement was based more on the quality of the reasoning and explanations than on the answers themselves.

The strengths of such an evaluation strategy resided in the possibility it afforded of rendering a relatively reliable judgement regarding the understanding and mastery of content elements and regarding the ability of the learner to use this learning in context: in short, a judgement on the level of understanding and on the transfer of knowledge.

A compromise? This evaluation strategy certainly had a few weaknesses. By reducing the judgement rendered to the sum of grades accumulated during the session, the final grade inevitably reflects at once what has been learned as well as when it was learned. In fact, even if mastery is total and complete at the end of the learning process, the grading process keeps track of early session difficulties. By the same token, if the judgement rendered makes it possible to confirm mastery of one given component of a competency, there remains a high risk that we will end up with a sum of graded components the total of which sheds little information on how the competency developed during the course as a whole. In addition, the attribution of a point value to a set of questions does not really inform learners on what they have achieved relative to the objectives of the course. So, let us call it an acceptable compromise until proven otherwise. Such a practice is part of a process that I will characterize by comparing it to a chain. The first link would represent the course objectives, which are themselves contingent on the requirements of the Ministry of Education, the program and the department. These objectives would then be translated by professors into teaching, the second link. The third link of this metaphoric chain would be the learning itself. Finally, there would be evaluation.

When we look at this model, we realize that evaluation is only taken into account at the end of the process. This increases the risk that it may become little more than a rubber stamp at the service of managing courses of study rather than a tool providing students with valid information on their learning. In addition to this, such a vision of evaluation often leads students to see evaluation as a necessary evil.

¹ Text appearing in *Pédagogie collégiale*, vol. 8, n° 1, October 1994, Nicolas FAUCHER and Clément LABERGE, « Être évalué, oui, mais bien! ».

² Among other things, these observations were based on the types of questions asked in exams (multiple choice, true or false, answers to be developed, etc.), the types of evaluation tools (exams, case studies, problems, oral presentations, etc.), the types of evaluation methods (number of evaluations, weighting, etc.) and the types of goals targeted by the evaluations (classification, diagnosis, regulation, attestation, certification, etc.).



TO A NEW EVALUATION STRATEGY

In time, I came to consider evaluation to be an integral part of the learning process. I am among those who see the learning process as the result of an iterative process involving interaction between a performance, a judgement rendered on this performance and an appropriate regulation mechanism. Moreover, judging one's own performance and making appropriate adjustments, are these not essential components of the concept of competency? If so, then the idea of only evaluating at the end of a process seemed to be less and less workable. After a number of meetings and talks, particularly with François Vasseur, an educational advisor at my college, I finally decided to go outside of my comfort zone and revisit my positions.

The main purpose of this exercise was always the same: to attest to the mastery of a given competency, but to do so in a better way. I wanted to avoid the weaknesses and paradoxes of my former evaluation practice without introducing any new inconsistencies. It was important to do better and not just to do something different.

Based on the principle that learners are the primary actors in their own learning process, I also wanted to give them the tools they needed for self-mastery. So it was also necessary to provide them with improved feedback on their performances.

THE FOUNDATIONS FOR EVALUATING COMPETENCIES

I will leave to more qualified experts the reasoning behind the theoretical foundations of the competency-based approach and of the approach of the program. That being said, I felt compelled to answer the following questions in order to achieve my objectives:

- What is expected of the learner?
- What indicates to us that a competency has been developed?
- In what context will the learning and performances be implemented?

Answers to the above questions should be provided in the final objective of the course and they should be in line with the broader course framework. As a case in point, I have provided below the terminal objective³ for the *Immunity* course (101-003-Li), that is, the course in which the tools referred to in this article are being tested:

At the end of the course the student is capable of linking immunological disorders and infections to physiological and metabolic mechanisms by:

- recognizing the major groups of microbial aggression based on their general characteristics;
- adequately using the homeostatic model to illustrate the impact and the extent of an infectious disorder on the human body;
- adequately using the specific and non-specific physiological immunity models in order to analyze and illustrate the extent and efficiency of the immune responses against infectious disorders;

³ At the time of writing these lines, the latest version of the course framework plan was still not formally adopted by the Nursing Program Committee of Cégep Limoilou. That should take place during the Fall of 2008.

- describing an infectious process using precise and appropriate vocabulary and identifying the elements relating to the pathogenic power of the aggression, its original flora and the main prophylactic measures;
- respecting basic principles that ensure aseptic conditions during handling; and
- carrying out correctly the different laboratory tests for diagnosing infections.

The *Immunity* course is a 2nd year course in the 3-year Nursing Care program. It provides students with the learning and knowledge they need to develop the competencies required in their program. In fact, the definition of the competency for this course makes reference to a superior intellectual ability. This is often the case for biology courses, regardless of the program of which they are a part.

Indeed, are not self-evaluation and self-regulation essential components of the concept of competency?

Having established the foundations, next I had to develop a plan that was coherent with the course objective. I proceeded to define the Key Learning and Evaluation objects (KLO). There are a total of five for the course, some comprising two or three elements. Each KLO consists of a group of learning components that define the essential capabilities which are linked to the specific competency components developed in the course and to the competency itself. Each KLO forms a whole, the study, the implementation and the confirmation of which can be carried out in a coherent manner.



Of course, KLOs are not isolated entities but rather closely interrelated. Collectively, they should be able to confirm the development of the course's competency; and a complex and genuine performance should make it possible to render a judgement on all the KLOs of a given course. On the other hand, each KLO can be treated individually and can be the object of one or several performances enabling the teacher to render a valid judgement. This is necessary because the targeted final integration can only be demonstrated once enough learning elements have been acquired. So, it is important to give the learner the opportunity to construct this learning and to provide feedback during the learning process.

After defining these KLOs, I needed to make them operational. This meant that I had to plan the course:

- taking into account the KLOs;
- with a view to a progressive mastery of the learning leading to integration;
- by integrating a formative process that makes ample use of feedback;
- by developing an evaluation plan that gives students more than one occasion to demonstrate what they have learned.

KLOs must also be developed keeping in mind how the course fits into the program⁴, in order to identify the essential capabilities students will have to demonstrate.

THE TOOLS

Figure 1 summarizes the evaluation plan for the Fall 2007 course. It provides the relative value of each KLO and its components, when applicable. Also included are the evaluation moments for each KLO. We can see that 55% of the final grade is linked to KLOs for which students are given more than one occasion to demonstrate their acquired learning. Indeed, my pedagogical approach has led me to introduce evaluation methods that not only cover newly acquired knowledge but also include a re-evaluation of previous knowledge in order to consolidate a student's overall learning. To this end, most components which can only be demonstrated once relate to the end-of-session evaluation which assesses the overall competency.

In this evaluation plan, theoretical exams 2 and 3 (X2 and X3) review elements previously assessed in one or more earlier evaluations. Generally speaking, acquired learning is evaluated more than once within a context that favours a gradual integration of this learning. Theoretical Exam 3 is the final exam for the competency of the course and it includes skills developed in the laboratory.

FIGURE 1 : TABLE OF KEY LEARNING OBJECTS FOR THE IMMUNITY COURSE

KEY LEARNING OBJETS (KLO)	X1	XL 1	X2	XL2	X3	Total / KLO
KLO 1: Micro-organisms		15		R		15 PT
KLO 2: Immunity						45 PT
a) Non specific immunity, sampling	10		10		R	20 PT
b) Specific immunity			10		10	20 PT
c) Homeostasy	5		R		R	5 PT
KLO 3: Infection						15 PT
a) Infectious process					10	10 PT
b) Anaphylaxis					5	5 PT
KLO 4: The artificial fight against infection				10		10 PT
KLO5: Handling						15 PT
a) The culture environment and Gram's colouration				5		5 PT
b) Handling and sampling				10+R		10 PT
Totaux by evaluation	15	15	20	25	25	100 PT

Legend: X = exam, XL = Lab exam, R = may require re-taking.

We wanted to make the process transparent and give learners the possibility to assume responsibility for their own learning, so not only did we need to explain to them the ins and outs of this evaluation plan, but we also needed to tell them what was expected of them and the criteria by which we could assess whether or not they demonstrated the relevant acquired learning. For this purpose, each KLO was developed in the form of a grid. Figure 2 provides an example of this.

⁴ There can be no coherent competency-based approach without an enlightened program approach.



FIGURE 2 : GRILLE DESCRIPTIVE DE L'OCA 2A DU COURS IMMUNITÉ

KLO 2 Immunity: Define the reactions of the organism under attack (competency element 2; 45%)

KLO 2a (20%) Non specific immunity

Performance criteria: 2.1 Accurate description of resistance offered by the organism under attack.

Demonstrating mastery of the KLO..

NO DEMONSTRATION (0%)	- No contextualization worthy of the name.
INSUFFICIENT DEMONSTRATION (30%)	- Did not adequately demonstrate the ability to explain the reason why two or all of the following reactions were triggered: inflammation, phagocytosis, fever. OR - Presence of first two deficits in the <i>Minimum Demonstration scale</i> .
MINIMUM DEMONSTRATION (60%)	- Difficulty in determining the impact and utility of three or more elements the 2 nd barrier (phagocytosis, inflammation, fever, interferons, complement). OR - No demonstration of the ability to describe why one of the following elements is triggered: inflammation, phagocytosis or a fever. OR - Presence of four or more deficits listed in the <i>Partial Demonstration scale</i> .
PARTIAL DEMONSTRATION (75% 2-3 DEFICITS) (80% 1 DEFICIT)	- Difficulty in analyzing in context of the impact of the 1 st barrier. - Difficulty in determining in context whether the interferons or the complement were triggered. - Difficulty in analyzing in context the utility of one or two elements of the 2 nd barrier (phagocytosis, inflammation, fever, interferons, complement). - Difficulty in analyzing in context the interactions among the elements of the 2 nd barrier. OR - Incomplete demonstration of what triggers, in context, the fever, the phagocytosis or the inflammation.
OBVIOUS DEMONSTRATION (90%)	- Deficits in clarity or incomplete answers leading to inaccuracies. OR - Approximate mastery of one or several basic concepts that does not interfere significantly with the ability to integrate.
EXPECTED DEMONSTRATION (100%)	- Clear analysis, in context of how the 1 st barrier is crossed. - Recognition, in context, of the conditions that trigger 2 nd barrier elements in response to infection. - Clear analysis, in context, of the effects and utility of 2 nd barrier elements. AND - Clear analysis, in context, of the interactions among 2 nd barrier elements in response to infection.

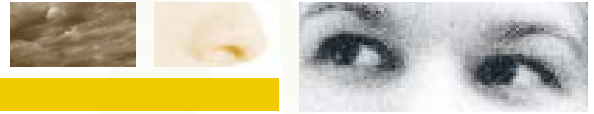
The grids were developed with the following objectives in mind:

- to clarify what is expected of the learner;
- to specify the criteria for rendering a judgement on required learning;
- to attribute a grade – which is in fact a rating – based on the nature and importance of what has and has not been learned rather than on the number of learned and unlearned elements.
- to design an efficient evaluation and feedback tool;
- to provide learners with the necessary information for taking over their own learning process;
- to take into account how the course fits into the program: in the case of final courses, more comprehensive evaluation grids could be developed.

Each grid should be interpreted in the following way:

- the lowest grade indicates that the required learning has not been demonstrated and is attributed a value of 0;
- the highest grade corresponds to what is expected – realistically – and is attributed a value of 100%;

- the grade corresponding to a 60% value corresponds to the minimum expectations for the KLO in question, expectations that should be driven by the threshold criteria defining success as provided in the broader framework relating to this KLO;
- the grade with a 90% value corresponds to a clear demonstration of the required learning, although with certain gaps relating to the quality of the performance (for example, unclear or vague explanations) or relating to a minor concept or content element;



- the grade with a rating between 70% and 80% corresponds to a demonstration that surpasses the minimum requirements, but remains incomplete, and the importance of what has not been learned⁵ is what justifies the attributed grade.
- the grade with a 30% rating corresponds to an insufficient demonstration, one that is below the threshold of success. For this particular course, given that there is as yet no double threshold and that, for the time being, the final grade is the sum of scores obtained for the various KLOs⁶, I wanted to prevent students who failed a number of KLOs from receiving a passing grade for the course. With this grade value of 30%, failing several KLOs leads almost inevitably to failing the course.

For those who would like to use such tools, a word of caution as to the number of grades that a descriptive evaluation grid like the I used, should have. Since a competency is a coherent whole, a grade should not correspond to the number of satisfactory performances carried out or to the number of mistakes committed, but rather to their nature and importance. Also, the number of grades in a grid should be determined according to the different types or the different kinds of demonstration of acquired learning that the learners' performances are likely to illustrate. Whatever the grid, there are bound to be performances that force a teacher to attribute an intermediary value that falls between two grades. Should such a situation persist, it might be wise to add a new grade. However, the fact of adding more grades only increases the difficulty of rendering a judgement and, ultimately, contributes to making arbitrary what should be based on an enlightened exercise in subjectivity.

IMPLEMENTING THE STRATEGY

In practice, this new evaluation strategy – new in my teaching practice – begins when I describe it to the students in the very first class, together with the presentation of the plan for the course. I also provide them with the evaluation plan matrix as well as the evaluation grids. During a performance that is the object of an evaluation, the relevant grid(s) are included in the exam questionnaire or the instructions sheet for the report. The feedback and the grade received are also directly indicated on the grids.

When a more recent performance by the learner demonstrates a superior level of acquired learning and he or she obtains a higher grade than previously for a KLO, then the grade is adjusted upward.⁷ This means that the teacher must use subsequent evaluation tools that include all the elements covered the first time and must set as high a level of requirements as were obtained in previous evaluations. It also requires that the course lend itself to a gradual construction of learning and to a repetition of opportunities for demonstrating its acquisition.

AN OVERVIEW OF REACTIONS

To date, a majority of the students to whom I have presented this evaluation strategy had never been introduced to anything quite like it. Although a number of them applauded the possibility of repeatedly demonstrating their acquired learning, students in general are sceptical, which is quite understandable. Some also fear a possible negative impact on their R scores.

Most colleagues have a similar reaction and it is not my intention here to either defend or condemn their opinions. They are also cautious and understandably so. Most of their questions relate to evaluation objectivity, grading and how to manage grades as well as to issues relating to teacher freedom.

VERY PROMISING RESULTS

Are the expressed concerns justified? It depends. It is not easy to assess the impact of this evaluation strategy on students' R scores. To date, their overall average results seem to have increased since I started implementing this evaluation strategy about five years ago. But it would be unwise to link the increase solely to the implementation of this strategy. At first glance, strong students obtain better grades than before, as do those who consistently work hard. On the other hand, those for whom college studies will represent the best eight years of their lives bring the group average down.

Having said that, students' reactions in general are very positive. I notice that many are really putting the KLO grids to good use. What is more, the number of grade-related complaints has diminished since the implementation of these grids. This is

⁵ We consider a performance to be incomplete when it fails to demonstrate one or more non-essential learning elements or when it reveals only partial acquisition of this learning.

⁶ This situation should change in the near future.

⁷ The opposite is not true, and this for many reasons. Let's just say that a more recent performance is considered more reliable; but it does not make the previous performance obsolete. In addition, the possibility of having an already-obtained grade lowered would place the learner in a stressful situation that would weaken the reliability of the evaluation itself.





a curious reaction from individuals who often consider evaluation to be objective, since evaluation is a process which calls for judgement, an act which is necessarily subjective, and with the KLO grids that subjectivity is clearly evident. However, such tools require that the reasoning underlying the attributed marks on such tools be documented and that leaves no room for arbitrary judgements.

Actually, the advantages of this evaluation strategy are even more far-reaching. With this method, learners have in their hands all the tools they need to effectively assume responsibility for their own studies. And many of them take full advantage of this! In addition, it promotes feedback: at any given moment during the course, learners are informed about what is expected of them in a way that enables them to assess their own mastery level of the competency and their own progress in constructing the required learning. Consequently, they are able to evaluate what they have and have not learned for each of the essential capabilities to be developed in the course and to recognize what remains to be done. Another benefit of this evaluation strategy is that it is an ongoing process! Concepts seen at the beginning of the course are not simply and indefinitely relegated to memory where they serve as fragile and unreliable foundations later on. On the contrary, the regular recurrence of these concepts and the strong need to keep them in mind throughout the course contribute significantly in a clear and verifiable way to increasing the quality of the learning which depends on them.

At first glance, strong students obtain better grades than before, as do those who consistently work hard.

From the teacher's perspective, the benefits are, in my view, major. For myself, I had been teaching the course for three years when I transformed my evaluation approach and I was overwhelmed by the depth of the new vision I had of the course after having developed the KLO grids. Since then, producing evaluation tools has become much easier and the amount of grading has been lessened. The only real inconvenience is the added task of grade management.

As for teachers' freedom, this way of proceeding determines the end but not necessarily the means. However, it seems to me that the end should be determined by enlightened consensus. Consequently, this evaluation strategy does not imply any new constraints, just those normally associated with a course. And since tools should be developed to be transferable, even though the initial development of the grids requires a lot of work, they create the opportunity for highly beneficial exchanges that often facilitate matters pertaining to fairness and equivalence between courses, a hotly debated issue.

CONCLUSION

A similar evaluation strategy has been implemented very successfully in other biology courses (such as *Physiology, Evolution and the Diversity of Life* and *Integration in Biology*). These very positive results have exceeded my expectations with regard to my initial objectives. They have attested to better mastery of competency components, more efficient feedback and regulation, a more coherent course, and easier development of evaluation and grading tools. However, since perfection is not of this world, we are not in any way suggesting that this type of evaluation strategy will solve all existing

problems. To this end, no doubt better results from this type of strategy will be obtained when students will no longer be trained in the evaluation of learning according to a strategy which is poles apart, as is still too often the case at the moment.

I am counting on the educational renewal at the secondary level to advance this cause. All in all, the fact of reducing the final grade of the course to the sum of grades obtained for each KLO strikes me as a way of changing the nature of the process at the end of the session; and that is why I am considering the possibility of soon introducing a double success threshold for this course. ●

Nicolas FAUCHER has taught Biology in the Natural Sciences, Dietetics and Nursing Care programs of Cégep Limoilou since 1999. He has served on two expanded MELS committees on educational renewal in the 2nd cycle of secondary education. He has occupied for three years, the position of Program Coordinator for the Natural Sciences program and, as a member of the Dietetics program Committee, he has also participated in developing the program's final exam.

nicolas.foucher@climoilou.qc.ca