

THE ASSESSMENT OF COMPETENCY AND THE IMPORTANCE OF JUDGMENT



GÉRARD SCALLON
Associate Professor
Faculty of Education
Université de Montréal

For years now and still today, evaluation of learning practices have adapted without too much trouble to measurement and evaluation concepts and even managed to grow despite some confusion in terminology. Evaluations have become a reality which encompasses a host of operations; all without calling upon the judgment of the evaluators. The practice of objective examinations, characterized by precise tasks such as answering multiple choice questions, is a firmly established tradition. The grades of several of these examinations are later blended into a total result which is analyzed and used in the “mechanics” of certification. The evaluation of complex performances such as writing a text and oral presentations, does not escape this practice. Even if tasks given to the students are more complex and elaborate than those in objective examinations, the mathematical processing of the results remains the same. Although the work is qualified in relation to criteria, the evaluation itself leads to quantifiable results and/or grades that will be tabulated sooner or later.

In a competency-based approach, evaluations rests on a methodology linked on one hand, to the nature of the problem situation or situational task¹ facing the student being evaluated and, on the other, to the analysis of results obtained during the course of study or at the end of training.

In everyday life, information is nothing more than sterile data unless it is based on judgment. Thus, a patient can be aware of the results of a cholesterol test, an investor the performance of his stock at the stock exchange, but the meaning of the data must be interpreted and in some cases, specialists must be called in. In light of these examples, we can say that quantifiable results obtained in examinations designed to evaluate performances are not judgments. They correspond nevertheless to evaluation practices in the field of learning, so it is up to those in charge within the education system to ensure results are interpreted. Conversely, in a competency-based approach, the process of evaluation must be a little more elaborate.

To infer the presence of a competency, complex tasks need to be designed at various moments during the training as well as at end of term. This series of complex tasks call into play the judgment of those in charge of the training or learning program. The judgment process clearly differs from the “measurement” process and largely exceeds the mathematical processing of quantifiable results.

¹ The authors are divided on the use of this. The expression “problem situation” is suitable for science and mathematics but less well suited to other fields. To write a text and to compose a melody cannot be considered problem situations. The word “task” seems more appropriate. The word “situation” can be ambiguous (an evaluation can correspond to what is happening in the examination room without referring to a task that the student must handle to demonstrate his competency). The expression “situational task” seeks to remove this ambiguity while respecting the widespread use of the word “situation”.

BEGINNING AT THE END

In a competency-based approach, we acknowledge the importance of judgment when it becomes necessary to attest, certify or recognize to what extent students have met the expectations of a program of studies. In matters relating to academic achievement, resorting to a passing grade as a benchmark for judgment is no longer valid.

A number of training and study programs target the development of more than one competency. To fully grasp the importance of an evaluation at the end of a long training period, we will assume that the student's mastery of each competency has been validated, that is to say, judged in a reliable and valid manner. But what of the way in which this judgment is expressed? Should a competency be judged on an “all or nothing” basis, or on a “graduated” score based on a rating scale? Table 1 illustrates results based on these two rating methods.

Table 1
CONTRAST BETWEEN TWO
COMPETENCY PROFILES USING
A RATING SCALE

	All or nothing judgment	Graduated score (from 1 to 4)
Competency 1	+	3
Competency 2	-	2
Competency 3	+	4
Competency 4	+	3
Competency 5	-	1
Competency n	-	2



Keep in mind that judgments made for each competency are based on the individual's handling of complex tasks. It is thus necessary to use tools that highlight the quality of this handling to assist us in making judgments, both from the perspective of the finished product (work achieved) as well as the capacity to call upon knowledge and skills.

The profiles seen in Table 1 must be taken into account in a process of certification (for example, when issuing a diploma). Although they are the object of written rules that vary from one establishment to another, evaluation practices on this subject matter are relatively unknown. The simplest solution would be to return to a mechanistic calculation and demand the acquisition of a minimum number of competencies, regardless of the competency. For example, to master six out of eight competencies! This model is considered "compensatory" in that the achievement of one component becomes a compensation for the failure of another. In addition, a more complex model is used in a great number of educational programs with mandatory and optional courses.

At the heart of the methodology used in the assessment of competencies lies the accomplishment of complex tasks and the handling of problem situations or situational tasks by the individual being observed.

Mastery of specific competencies can be required (conjunctive model) whereas the principle of a minimum requirement is applicable to other competencies that are grouped in subsets (compensatory model). Readers interested in this question can consult the writings of Louis (1999), Trumbull and Farr (2000) and Scallion (2004) among others.

As a matter of fact, basic knowledge and skills relative to the demonstration of competencies were not included in this instance. From a judgment perspective, taking them into account makes certification rules even more complex. Haladyna and Hess (2000) discussed this by presenting a "Sequential" model that requires the mastery of basic knowledge and know-how in individuals being evaluated before attempting to evaluate their competencies.

THE CONCEPT OF COMPETENCY

The certification of competencies at the end of a training program must be perceived as the result of a long series of events witnessed by teachers. As we will see, judgment holds a prominent role in the process. But before describing the evaluation approach whose principal objective is the tracking of each student's progress, let us review the concept of competency. In the opinion of a number of authors who echoed the work of Le Boterf (1994), such as Perrenoud (1995) and Roegiers (2000), competency is defined as: [...] the ability of the student to mobilize his own resources or to spontaneously use external resources to accomplish complex tasks within a same family of tasks.

This definition, like others used for inspirational sources, touches upon evaluation concerns. Concepts of complex situations, problem situations and tasks to be accomplished seem to be an integral part of the question. To be certified as competent, a student must have produced something tangible within a precise context and in such a way that the task accomplished is clearly distinct from knowledge and know-how. It is necessary for the individual being observed, to decide autonomously to put **what he knows and what he can do into practice**. The demonstration of this ability is not specific to one single problem situation but rather a whole set of situations within a same family. This family of situations is an important concept from both a training and an evaluation perspective. It is necessary to present students with several situations so they may manifest and exercise their abilities, thereby creating situations for observation of the competency. In a competency-based approach, the integration of these two training aspects will have to be better articulated relative to evaluation practices.

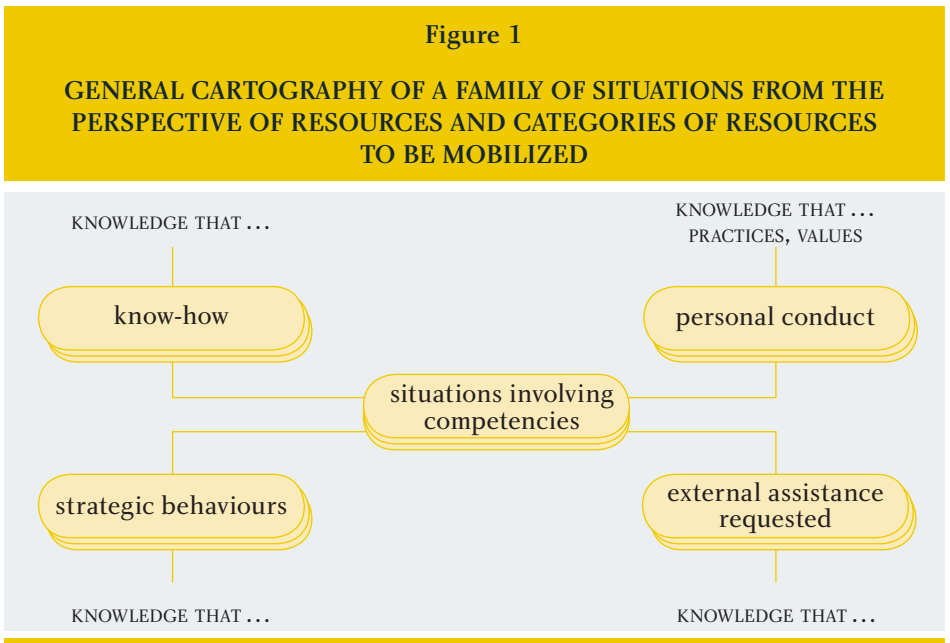
At the heart of the methodology used in the assessment of competencies lies the accomplishment of complex tasks and the handling of problem situations or situational tasks by the individual being observed. It is also one of the key points of the selected definition. To infer the existence of a competency, situations presented to the students must have specific characteristics such as, being connected to the targeted finished product. The situations must be realistic, i.e., authentic as in professional training. But the most critical component is the ability of the situation or task to galvanize the student into mobilizing many resources, those in his personal repertoire and those external to him but that he must spontaneously call upon on his road to autonomy (documents, resource persons, tools).

In evaluations, a limited number of "personal-resource" categories are required for an individual to demonstrate competency in the accomplishment of a complex task: Skills (knowledge), know-how, abilities and strategies (know-how deliberately chosen based on circumstances) and personal conduct. This last category of resources is not easy to manage insofar as we continue to confine it to the affective



sphere, as is generally the case. In a competency-based approach, personal conduct can appear as ingrained practices that add a certain quality to an individual's work: compliance with safety requirements, precision in calculations, a desire to cooperate and provide mutual assistance, a tendency to verify and carry out a reflexive review of work done.

Situations involving competencies should be analyzed from the perspective of resources that need to be mobilized. It is one way of validating situations that allow students to demonstrate their competency, i.e. their ability to call upon their own resources as well as external resources. This analytical approach demands judgment and understanding. To illustrate the approach used during the examination of one or several situations involving competencies (problem-situations or complex tasks), we are including a chart or "cartography" below.



PROGRESSION MUST PRECEDE ANY FORM OF CERTIFICATION

It is commonplace to affirm that our ability to wisely use what we know and what we can do is a process that develops (some would say "is constructed") gradually. According to what progression though? In music as in gymnastics, the pieces to be learned or routines to be carried out can be graduated in complexity and difficulty. This approach is not appropriate for all fields and can be difficult to apply in cases like reading, writing or mathematical problem-solving. A definition of competency that values the ability to "mobilize resources" suggests one approach from which to examine the students' progression in the development of one or more competencies, that is, the degree of autonomy students display in the use of the relevant resources for the handling of complex situations of a same family (Becker, 2002).

For example, let us look at the manufacture of wooden objects within the framework of manual labour. All kinds of projects can be carried out, anything from a storage box

to a birdhouse. To successfully complete each project, the individual must autonomously call upon his knowledge, know-how and personal conduct. He must be able to foresee requirements in materials and tools, read a plan, etc. Moreover, it is necessary for him to comply with the safety requirements and to control the quality of the work. Altogether, a great number of resources to mobilize, a vast amount of knowledge and know-how to complete the project wisely. However, the mobilization of each resource can be judged a "success" or "failure" for each project accomplished (each situation), taking into account the assistance needed to accomplish the task:

No mobilization, even with assistance (clues or incentive)	0
Mobilization with assistance	+
Mobilization without assistance (full autonomy)	+ ⁰

The symbols "0", "+" and "+⁰" (or any other code) can represent the progression of each student in a synthesis that shows the level of performance in many situations, a performance that is analyzed according to the principal resources to be mobilized. Table 2 provides an illustration of this.

Tracking the progression of each student in each situation is not without difficulty in collective or simultaneous instruction. As seen traditionally in the evaluation of complex work, rating the work does not require the student's presence. "Autonomy" as seen in Table 2 nonetheless requires the presence of the student or at least a certain knowledge of "how" each student performs. An evaluation methodology is required that will allow finished products to reveal the "mobilization" of specific resources thus facilitating the tracking of progression.



For example, precision in the assembly of a storage box can be seen as proof of the presence of quality control. On the other hand, the finished product of manual labour reveals nothing of compliance with safety requirements.

You will note that the judgment for the situations of Table 2, relative to the resources that must be mobilized, is a simple one: a 3-category checklist (no mobilization, mobilization with assistance and autonomous mobilization). The task of evaluation would be much harder if it were necessary to stringently apply an evaluation grid with criteria accompanied by refined levels of judgment. There is no emphasis placed on the analysis of performance in each situation. Rather, the approach suggested, although it may appear superficial at first glance, is supported by data collected from comparable situations. The approach can be adapted to suit several competencies, be they literary essays, financial assessments or nursing plans.

RETROSPECTIVE JUDGMENT

The tracking of progression based on several tasks or situations leads to another form of judgment that is not negligible: the retrospective judgment. Sooner or later during the progression of students, each competency requires an update or an assessment. Instead of adding another situation for students to formally demonstrate their abilities, it appears possible to use “evaluation memory” based on results achieved with situations already previously experienced. This is how the retrospective judgment comes into play, so called because it is supported by previously achieved progression.

Initially, an assessment is done at the appropriate time regarding the student's ability to mobilize each resource relative

to the targeted competency. This is a validation that the observed individual is “yes, capable” or “no, not capable” of wisely using a particular resource that is part of the definition of the competency. We therefore have a series of judgments expressed by a “yes” and by a question mark if there is doubt. A sequential series is interpreted as a profile such as the one shown in the last column of Table 2, under the heading “R” (Retrospective).

From the point of view of assistance and regulation of the progression of individual students, the profile we provided could remain unchanged given that it clearly shows strong points as well as areas that need work. From the point of view of provisional assessment or certification, a retrospective judgment will undoubtedly be reduced to a single rating expressed numerically (from 1 to 4) or expressed by a letter. Rules should therefore be put into place that link performance descriptions to ratings. For example, the highest rating, number 4, is assigned to students who demonstrate the ability to mobilize most of the necessary resources for the complex tasks connected to the competency. The next rating, number 3, means some weaknesses are allowed; and so on for the next numbers. These descriptions can be found in the paragraphs which constitute the levels of a global descriptive scale, a process described many times in writings on the subject of evaluation.

Table 2
SYNTHESIS SHOWING THE PROGRESSION OF A WOOD-WORKING APPRENTICE

	SITUATIONS					R
	1	2	3	4	5	
Reading of a plan	---	---	+	+ ⁰	+ ⁰	Yes
Planning	+	0	---	+	+ ⁰	Yes
Shaping	0	+	0	+ ⁰	+ ⁰	Yes
Assembly	+	0	+ ⁰	+ ⁰	+ ⁰	Yes
Reflective review	0	0	+	+	0	?
Safety requirements	---	+	+	---	+	?

---: Not relevant or non-applicable resource

R: Retrospective judgment supported by progression realized

What is really at stake in the judgment of teachers or any evaluators is the determination of the level and, consequently, the rating (from 1 to 4) for each student based on the profile taken from a retrospective judgment. With the use of descriptive ratings, we have now reached the most critical stage of the “judgment” adventure in a competency-based approach. There are many arbitrary factors in the determination of performance levels. To establish a rating that corresponds to the inferred competency of each student is without a doubt the most subjective portion of the evaluation process.



WHEN THE JUDGMENT ITSELF IS JUDGED

Quality control plays a major role in the field of evaluation and measurement. The measurements, the tools, even the evaluation approaches must be reliable and valid. The importance given to judgment in a competency-based approach makes it impossible to circumvent. Within the framework of tracking progression, situations must allow students to develop and apply their competencies, i.e. use their ability to mobilize knowledge, know-how and call upon personal conduct. By the same token, data collected on work done must lead to judgments that are clear so as to ensure the progression of each student and in certain cases, the necessary remedial or corrective adjustments.

Within the framework of validation and certification, quality controls are just as important. However they create a kind of impact that differs from the one implied in tracking progression. The seriousness of the decisions linked to certification or the issuing of diplomas, for example, requires particular caution even if only as regards the accuracy of judgments. The most current evaluation tool is tied to the judgment of one person who rates the work and performance of several students. Moreover, this tool does not reveal the degree of subjectivity inherent in the evaluation procedure. With the opinion of a second and even third person, it becomes possible to test the quality of the evaluation process as regards subjectivity. For example, Table 3 documents the evaluation of three students (J. C., P. A., M. N.,) by three evaluators (A, B and C).

At first glance, we see major discrepancies from one judge to another in the rating of students. Thus, we see that student M. N. has three separate ratings, a 2, 4 and 5. Using an evaluation tool that requires one judge only can mask the disparity seen above. Which one of the three ratings is most fair? With a greater number of students, this problem will not be as obvious as the example shown here. However, in such a case, it will be necessary to find a coefficient of agreement incorporating the degree to which different people can judge a set of productions. (Scallon, 2004) .

The importance given to judgment in a competency-based approach makes it impossible to circumvent.

Lack of agreement among evaluators may be a warning for us to sound the alarm about the process of evaluation that leads to key decisions. The concerns that arise are closely linked to key principles in the practice of evaluations regarding fairness, equity, transparency and more. Two kinds of corrective measures are required. On one hand, we should consider a revision of the evaluation tool, both the evaluation grid or the descriptive scale used for rating. On the other hand, we should begin training² those who are responsible for evaluations, through the use of standard copies whose characteristics and qualities are well-known.

The rating of student competency is not only about the evaluators. The situation or task presented during the evaluation is also a vital factor. Many studies have shown that even with equal competency levels, students do not accomplish tasks in the same

way. Completion is dependant on the statistical interaction of “subject x task”.

Table 3

FICTITIOUS CASE SHOWING THE WORK OF THREE STUDENTS RATED BY THREE JUDGES

	JUDGES		
	A	B	C
STUDENTS			
J. C.	3	2	3
P. A.	4	5	2
M. N.	2	4	5

This explains why a given theme designed to evaluate critical thinking or determine clinical competency in medicine can awaken different emotional reactions in different people and arouse a degree of commitment that is not the same for everyone, even in individuals of equal competency. Parkes (2001) wrote a well-documented article on the subject linking this phenomenon to a problem of transference. One possible solution is to resort to more than one situation to infer the existence of a competency. This relates to the “family” of situations discussed earlier.

All in all, situations designed to infer competencies should be examined from two different angles. First, these situations must galvanize the student into mobilizing several resources “to concretize” the evaluation process. Then, we can counter the undesirable effect of the statistical interaction of “subject x task” by resorting to several situations of the same family.

² Having “apprentice” judges evaluate productions already evaluated by “experts”, to compare their judgment against those of experts. Judges in training can then be provided with feedback.



IN CONCLUSION

The purpose of this article is to stress the importance of judgment in the evaluation of learning within the framework of a competency-based approach. Even if objective examinations are applicable to learning objectives linked to knowledge and basic know-how, the methodology is insufficient to infer competencies. Moreover, we must base our performance ratings on the quality of work and the ability of students to use what they know and what they know how to do (using all the components that are identified in the competency, that is, the ability to wisely use and mobilize various resources in the handling of complex situations). Lastly, the mechanics of totalling quantifiable results to reach a final grade is not appropriate in a competency-based approach; it cannot take the place of professional judgment, in particular the judgment of those in charge of training.

The importance of judgment can be seen within the first moments of progression of the individual. The mobilization of various resources in the handling of problem situations and the accomplishment of complex tasks is the result of a great deal of learning that must be tracked and regulated. The choice of situations both favourable to learning and able to provide indicators of the gradual mastery of a competency is most assuredly a question of judgment. Each situation that is designed must be linked to the global mapping of the desired competency that details all the knowledge, know-how and personal conduct to be mobilized.

The methodology for evaluating the mastery and progression of each competency remains to be developed. The principle of "autonomy" by which each student mobilizes various resources from one situation to another, has been kept

as the basis of progressive and retrospective judgments. The assessment can thus remain in the form of a profile identifying strong points and areas needing improvement. However, the assessment can also be summarized as an overall rating that is included in the report card. It is then necessary to resort to judgment tools with descriptive scales that provide qualifying statements for the competency in question. The drafting of this scale and its use require the use of judgment.

Nor does certification escape this process of judgment. Quite the contrary. When the competencies in a program of studies have been validated and certified as to their level of mastery or non-mastery, they must be integrated so as to meet the program requirements for success. At this stage of the evaluation process in a competency-based approach, the differences from the traditional processes are most obvious. The practices developed in this field are still relatively unknown and little used or appear to be based on an approach that is more qualitative than quantifiable.

There is a final point not to be overlooked as concerns the importance of judgment in a competency-based approach. The evaluation approach demands a lot from teachers and, sooner or later, training activities and improvement activities must be designed for them. Within the framework of a competency-based approach, those responsible for training intervene together with the students in the evaluation process, the latter group being a component that cannot be ignored in a post-secondary context.

Here too, there are an important distinctions from traditional evaluations. Examinations require a certain basic technology, including evaluation aspects like drafting questions and processing results which are not very conducive to student participation. However, it is possible to teach students to design their own self-evaluation tools in connection with the contents of the examinations (Scallon, 1999), but the procedure involves hard work. In a competency-based approach, the self-evaluation of students is crucial because the ability to make a judgment on their own progression is an integral part of competency, perhaps all competencies. In addition, the reflective review by the student on what he is in the process of achieving or what he has accomplished is the cornerstone of a great number of statements on competency. Without a doubt, the ability to make a judgment is in itself a competency to be developed in students and in fact, one of the goals of the training portfolio or file. ◀

BIBLIOGRAPHICAL REFERENCES

- BLECKERS, J., *Développer et évaluer des compétences à l'école: vers plus d'efficacité et d'équité*, Bruxelles, Éditions Labor, 2002.
- HALADYNA, T. and R. Hess, «An Evaluation of Conjunctive and Compensatory Standard-setting Strategies for Test Decisions», *Educational Assessment*, Vol. 6, n° 2, 2000, p. 129-153.
- LE BOTERF, G., *De la compétence. Essai sur un attracteur étrange*, Paris, Les Éditions d'organisation, 1994.
- LOUIS, R., *L'évaluation des apprentissages en classe. Théorie et pratique*, Laval, Études vivantes, 1999.
- PARKES, J., «The Role of Transfer in the Variability of Performance Assessment Scores», *Educational Assessment*, Vol. 7, n° 2, 2001, p. 143-164.



PERRENOUD, P., «Des savoirs aux compétences. De quoi parle-t-on en parlant de compétences?», *Pédagogie collégiale*, Vol. 9, n° 1, 1995, p. 20-24.

ROEGIERS, X., *Une pédagogie de l'intégration: compétences et intégration des acquis dans l'enseignement*, Bruxelles, De Boeck Université, 2000.

SCALLON, G., *L'évaluation formative*, Saint-Laurent, Éditions du Renouveau Pédagogique inc. (ERPI), 1999.

SCALLON, G., *L'évaluation des apprentissages dans une approche par compétences*, Saint-Laurent, Éditions du Renouveau Pédagogique inc. (ERPI), 2004.

TRUMBULL, E. and B. FARR, *Grading and Reporting Student Progress in art Age of Standards*, Norwood (Mass.), Christopher-Gordon, 2000.

Gérard SCALLON received his doctorate in theory of Education with a specialization in measurement and evaluation from the Ontario Institute for Studies in Education of the University of Toronto. He teaches evaluation of learning in the Faculty of Education at Université Laval since 1968. Author of two articles on formative evaluation (1988 and 1999), he also published several articles in scientific magazines. The advent of the competency-based programs aroused a keen interest in him for renewed practices in evaluation. This is the subject of his new book *L'évaluation des apprentissages dans une approche par compétences* published recently by Éditions du Renouveau Pédagogique inc. (ERPI).

gerard.scallon@videotron.ca