

**Formative Assessment in Music Education:
Research on the Transfer of Knowledge During a Short Training Course**

By

Dr Cédricia Maugars
Observatoire Musical Français
Université Paris Sorbonne Paris IV
France

Abstract

This article summarises the main results of French research focused on teacher assessment in music in secondary schools, conducted between 1999-2004. The enquiry suggests reasons why short courses do not change established ideas about assessment in music. Increased time for reflection and self-criticism during training courses would facilitate the development of greater understanding and modification of practice. Knowing what the participants expect from the course may help to predict the extent of the transfer of information.

The present research describes conceptions that music teachers have of assessment while in training (before, during and after the training course). The training course is a short course, with a specific goal, limited in time and content. The first level is entitled: “Formative assessment in music teaching.” It is important to state that this research will not study the training itself but the conception the trainees have. The reactions of this cross-section of students form the core of this work.

We have attempted to describe the music teachers’ perceptions according to a questionnaire. More specifically, our reflection aims at understanding the paradox that music teachers face when assessing their students’ musical performance. Indeed, in previous research, we had noticed that teachers aim at being objective by using many criteria for assessment and markings, but rely more often than not on subjective judgments (accuracy, musicality, and interpretation).¹ Moreover, “assessment” is often associated with “test” (De Ketele, 1997). How

do music teachers assess performance? Do they mark it 14.5 or 16 out of 20? Is an error in musical interpretation considered a fault? (Astolfi, 1997).

Specific writings on musical assessment are rare in French. Therefore, we attempted to find the specific meaning of assessment in music teaching through the perceptions of teachers who attended a training course. We studied the specific questions of musical assessment in Anglo-Saxon works (Mills, 1991; Hunter 1999; Hewitt, 2003).

Several questions arose from the start of the investigation: what conceptions of assessment do teachers have? Did the training contribute to modifying them? More precisely, did their conceptions concerning musical assessment remain the same in time, i.e. in T0 (before the first day of the training course), T+3 months and T+7 months (after the second day of the training course)?

In other words, the research aims at determining whether the conceptions evolved over the months, and if the knowledge and/or techniques learned during the formative assessment were put into practice in their classes. It will establish the link between the expectations from the training course and the eventual changes in understanding.

Hypothesis

Regarding the assessment of a training course in formative assessment in music teaching, we have formulated three general hypotheses :

1. Music teachers' conceptions focus on marking.

Assessment is often synonymous to marking. The marking criteria are based on an "arrangement" specific to each teacher, which presents an area of strong resistance regarding personal identity and professional legitimacy. The music teacher is the only one to decide the marking, in order to have his or her authority recognized and to manage class discipline.

2. Conceptions of assessment are difficult to change after a short training course, for they are anchored in customs which are resistant to the formative process. “Profits from a day’s work are lost when there is more than six weeks time between two one-day training courses.”²

3. An interview with the trainees after the course could reflect the change in their conceptions. The quality of the change can be forecasted when the trainees’ thoughts and expectations about a training course are known.³ “No human action can be reduced to its observable components; it is always motivated and often refers to implicit values and interests. In many cases, beliefs have more power than objective facts.”⁴

In order to compare the evolution of the teachers’ conceptions, it was most important to have a parallel between the two days of the training course. Our protocol did not refer to a sample group. In fact, the program aims at observing a unique group with a pre-test and a post-test with no sample group.

Methods

The researcher’s protocol follows five phases (see table 1).

Several evaluation devices could be put into practice before, during and after the training days. The fusion of several methodological tools allowed us to have different levels of information in the questionnaire; the Q-sort and the interviews were not dealt with in the same way.

In order to understand music teachers’ conceptions, we looked for a tool that allowed the collection of both quantitative and qualitative data. Not only is the Q-sort of that kind but it is also a formative assessment tool which can be used for self-criticism.

The Q-sort⁵ is a tool to collect data which can include subjective opinions and personal thoughts more precisely than questionnaires or behavior scales (Stephenson, 1935). This tool offers any user the possibility to identify, through successive differentiations, his or her behaviour or position on the issue of musical assessment. In ideal conditions, it permits the researcher to gather subjective opinions emanating from authentic personal considerations in conditions which are specifically favourable to free speech. The Q-sort aims at offering the trainee the occasion to analyse his or her different conceptions of assessment according to implicitly or explicitly different choices. It will highlight the group's tendencies, consensus and contradictions regarding music assessment.

The Q-sort usually presents a series of items⁶ putting together different definitions of assessment. All the proposals were sorted at random. Our goal was to study how the teacher would react and respond. At two times during the training, the teacher will have to classify the twenty items on a scale of five points (-3, -2, 0, 2, 3) which range from the strongest acceptance to the most categorical refusal, according to the degree of importance he or she gave the written sentences. The number of sentences was established according to the curve of Laplace-Gauss. The distribution of the twenty affirmations was classified into five categories:

- Two items: affirmations I prefer.
- Four items: affirmations deserving to be considered.
- Eight items: neutral affirmations.
- Four items: bad affirmations.
- Two items: affirmations that should be completely rejected.

The second step in the work consisted of proceeding to four item equilibriums. The groups A, B, C and D are composed of five assertions each according to the behaviour of the teacher towards musical assessment.

In order to formulate degrees of importance, we took our inspiration from the Q-sort suggested by A. De Perretti (1998, p.127). The Q-sort created for this research was composed of twenty items sorted according to four conceptions of assessment (see table 2).

According to his or her conceptions of music assessment, the teacher copies the number of the item in the table's boxes of the Q-sort sheet according to the coefficient he or she chooses: - 3 "to be rejected," - 2 "bad," 0 "neutral," 2 "deserving to be considered," 3 "best liked."

The time allowed for answering did not exceed fifteen minutes, the goal being to answer spontaneously. The responses were followed by a rudimentary counting of the answers.

In order to analyse the answers, we gave a coefficient to each answer according to its ranking: - 3 "to be rejected," - 2 "bad," 0 "neutral," 2 "deserving to be considered," 3 "best liked." A total could therefore be determined for each affirmation. An analysis of the results was then taken as a basis for the affirmations, the conception of assessment, or the general tendency towards assessment. The Q-sort was a statistical tool by itself, since all the participants' answers were distributed according to the law of a Laplace-Gauss curve. Nevertheless, the present research did not include any statistical study, since the public we assessed did not correspond to a statistical cross-section.

To respect their anonymity, teachers were asked to draw a unique sign on top of their papers, so as to be the only one to recognize their responses. Therefore, during the interviews, the trainees could find their questionnaires and could amend their previous answers.

Individual Interviews

Through personal interviews we confirmed the answers, since it is often difficult to find the right words to qualify the practice of music assessment. Furthermore, we found it necessary to allow teachers to reflect on their pedagogies from a distance.

During the first interview, the teachers were requested to precisely explain their personal choices, and the rationale for choosing the two Q-sort affirmations with which they agreed completely, and the two with which they strongly rejected. They were then asked what they remembered from the documents which were distributed during the first training day.

The second interview focused more on the problems connected with music assessment. However, we decided to keep the same questions as during the first interview, so as to make a comparison between the two. Five axes have been chosen:

- The training course's intellectual attainments : what the teacher says he or she remembers from the training course and/or has put into practice
- Conceptions of the training course : expectations of the course, and experiences
- Conceptions of assessment : how the teacher defines the word "assessment"
- Conceptions of the practice of assessment: the teacher's opinion of assessment practices, and his marking criteria.
- Putting changes and resistance into words: what the teacher thinks he or she will improve, and reasons of resistance to changes.

Interviews lasted approximately one hour, and questions were asked randomly, so as to allow the teachers more freedom. Each interview was listened to and transcribed after the first and second days.

The data from all the interviews were separated and grouped by theme. These themes will

be introduced according to the frequency of their appearance, according to the importance teachers gave them, following the interpretative analysis model (Kaufmann, 1996). Subjective conceptions and meanings were classified according to the “phenomenological” approach (Cardinet, 1989).

Description of the Interviewed

Fourteen Q-sort were distributed : thirteen teachers agreed to leave their addresses and phone numbers, eight agreed to be interviewed - i.e. in chronological order, teachers n° 2, 7, 3, 9, 0, 6, 12 and 10. After the second training day, nine Q-sorts were distributed (see table 3).

Eight trainees attended the two training days and six were interviewed. We thus only kept six “couples” of Q-sort, those of the people who took part in both investigations. The description of the results will be based on this sample – i.e. teachers n0 0, 2, 3, 7, 10 and 12.⁷

Distribution of the Interviews Throughout the Year (see table 4)

Eight teachers participated in the first series of interviews and six in the second. This last sample of six trainees followed every step of the protocol (before and after the training course), permitting the measure of alterations of their conceptions of assessment.

This group consisted of four men and two women, ages 29 to 48, among which were five who qualified and one who passed the “Agrégation” competitive exam (see table 5).

One can notice the variations in age and number of teaching years the teachers registered for the training course possess.

Results of the Investigation

Tendencies of Conceptions in Q-sort 1

Table 6 displays the results of the first Q-sort distributed on November 10, 1998, which permitted the classification of the conceptions of assessment in four categories.

A first reading of the questionnaire's results and of the Q-sort shows that conceptions of assessment are paradoxical as shown in the variety of answers. Opinions focus on the assessed (C) – i.e. the student being assessed (teachers n°2-3-12). Other conceptions of assessment favour the assessor (A) (teacher n°10) and the relation between the assessor-assessed within the assessment (B) (teachers n° 0 and 7). On the contrary, the conception of assessment based on the assessor (A) – i.e. the teacher himself – is rejected (teachers 0-2-3-7). Also rejected is the relation between the assessor-assessed (B) (teacher n°12) and within the context (D) (teacher n° 10).

Tendencies of Conceptions in Q-sort 2

The second Q-sort showed that teachers approved an assessment based on the assessed (C) and rejected an assessment based on the assessor (A). However, the answer's distribution is not identical to Q-sort 1 (see table 7). The conceptions based on the assessor (A) (teachers n°0-2-3-7) and on assessment based on the context (D) (teachers n°10-12) were rejected by the same teachers as in Q-Sort 1. Conceptions based on the assessed (C) (four teachers, n°2-3-7-12) and assessment within the environment and the context (D) (one teacher, n°10) were favourable. In summary, two kinds of conceptions appear: from teachers n° 2-3-7-12 and from teachers n° 10-0. We then looked toward the interviews to confirm those two tendencies, and how specific music teachers' conceptions are compared to other subjects.

Prominent Use of Marking: Assessing Means Marking

Concerning the conceptions of the teachers who were interviewed, it appeared that assessment practices were synonymous with marking: "I give 10 to all so as to prompt them to sing," "assessing means marking," "I see marking as more precise." Assessment was seen as less precise or "objective" than marking: "it means judging of a level in a vague way" (teacher n° 0) or "seeing the totality, whereas marking means marking some precise thing, this being more musical, more scholastic" (teacher n° 2). Assessment was a specific moment in teaching,

materialized by a judgment whereas marking was time-bound: “marking takes place at a specific and precise moment whereas assessment is more general. I would include assessment in both judgment and marking” (teacher n° 3). “The student’s progress is not taken into account through judgement but through marking, which is given at the specific moment when he or she is examined.” All these conceptions were valid in regard to summative assessment. Teachers did not differentiate between assessment and marking.

Although teachers carry with them traumatic memories of their school days, they continue to reproduce a kind of summative and certificative assessment close to that of a music exam, which, given at a particular instant, is “something tormenting and emulating at the same time which forces them to react” (teacher n° 2). Few of the teachers interviewed consider marking as useless (teachers n° 3 and 12); they said it often allowed them to “spare” time. It was often thought that assessing was a waste, as it is separate from the teaching sequence, at its conclusion (teacher n°10).

The Marking Belongs to the Teacher

All the interviews demonstrated the deep-rooted conception of control on marking. Teachers were primarily responsible for their students’ assessment; they alone did the marking and did not practice co-assessment: “there is only one judge – me, and if assessment is to be done by others, it’s a whole management system one must put into place” (teacher n° 12). Marking means having power which was not delegated to their students. This allowed teachers to “take hold of the class:” “assessing is needed, it’s part of my teaching, it reinforces my authority because students are more inclined to silence” (teacher n° 2). Assessing was a means of pressure which allows the teachers to “survive”: “it would be very bad for me if there was no more marking because that means students would eat me” (teacher n° 7). Considering a future without marking in music: “it would pave the way for the end of music teaching as such; “it would soon become an option” (teacher n° 10).

The teacher is the only “judge,” whether students are examined orally or in writing. They are judged on musical criteria, but also and above all, on behaviour. The marking of the students and summative assessment remain the frame of reference: “I give a global marking according to the student’s behaviour and then I penalize him or her: I start with a positive marking and I take points off – if that one is never silent and talks all the time, I take a point off” (teacher n°7).

Assessment was neither explained nor justified to the students because it remained a mental activity in the hands of the teacher. They ended up admitting that they judged an overall impression rather than specific elements of the student or students’ performance. Teachers based their assessment of musical performances on their own experience and their aesthetic tastes.

Justification of Marking

Although some teachers were aware that marking alone is not valuable, they defended its “virtues:” it aimed to encourage students and is “the air of time” (four teachers). The reflection was often ambiguous: “a mark doesn’t mean anything, it’s a subjective thing that I try to turn objective, but from the moment I try to turn it objective, it becomes subjective by itself” (teacher n° 3).

Often facing this dilemma, teachers decide not to give any detail to their students: “once the mark is given, I don’t change. I examine four of them and in the end, I give each a mark. I don’t change it but I write down some remark, and when I calculate the average, if the mark shows 0.25, I’ll round at 0.5” (teacher n° 10). Teachers admit they mark the students’ behaviour, although it’s not “allowed” by the “official” policy: “I think about how the student behaves; it’s a whole. I admit it’s completely subjective” (teacher n° 3). “I’ll penalize those who do not listen, talk, turn around, blow in the flute anyhow, those who bother me while others are working” (teacher n°7). They admitted that the marking was very imprecise: “I don’t yet see

how I can give a mark. I don't want to do it because we're not creating super soloists. There are problems relative to the break of the voice, to the students themselves; they have a hard time trying to find out who they are" (teacher n° 7).

Among all the conceptions, two assessment profiles emerged, either focusing on a summative model (teachers n° 10 and 0) or oriented to giving value to the student (teachers n° 2, 3, 7 and 12). Let's start with the first profile which mainly judges the final performance.

Assessment Focusing on the Final Performance: Assessing Means Penalizing a Fault

Teachers n° 0 and 10 based their conceptions on a final performance assessment. Error became a fault which should be penalized: "fault is allowed; in other words, with reference to the moment when a mark is given, the fault is more allowed than the student given value to" (teacher n° 10). Music assessment in secondary schools was believed to be like a Music Academy exam: "Assessing is more like a sanction for the student at a given time; I think that for music, it should remain like that because finally, it's the same as what is done in the Music Academy, one is assessed at a given time" (teacher n° 10). An error in the performance of an instrument (note or rhythm) is connected with a fault that must be rectified, sanctioned or penalized.

Assessing Means Giving a Mark to an Examination

Teachers n° 0 and 10 considered that assessing meant giving a mark to a test. Their conception of assessment is summative: "The trainer or inspector tells us 'this is what I expect from you', according to the standard, to what the inspector thinks; if we don't do that, we're on the wrong track" (teacher n°10). Teachers were looking for unique criteria which would help them "mark correctly," and "know how many points to take off for each musical fault." They allowed the fault which was an error of rhythm or pitches: "What can the sanctions be? How to graduate them? ; I would have liked to know the extent of a sanction." They debit the mark with

points: “I think a mark should be given objectively ...I try to be just and take out half points for errors but if the student doesn't pay attention and makes a lot of mistakes, he or she quickly reaches 0” (teacher n°0). Teacher n° 0 starts from a perfect performance and takes points off. “Orthographic errors are taken into account in a dictation so, in music, when one plays the flute, the only things that can be taken into account are precisely notes and rhythm” (teacher n° 0). While assessing, these teachers made no differentiation between assessment and marking. Although they considered their marking to be just and objective, teacher n° 10 admitted he or she gave an approximate mark after the student had been examined. Nevertheless, they felt that marking was objective and admitted that they “cheat.”

These teachers mark on a scale of 1- 20 and took into account all the wrong notes and rhythms, although they thought value should also be placed on the students. The teachers of this profile tend to give a completely different meaning to “value” in some phrases such as: “the mark gives value to music teaching” which can also signify competition: “the student feels he or she is valuable in the subject in which he or she has the best marks, which gives him value” (teacher n° 10).

In this case, teachers gave a mark only to the work or to the music performance without taking the students into account.

Assessment Focusing on the Student: Assessing Means Assigning a Value

Teachers n° 2 and 12 had an assessment profile that gave value to the student: “I reward them with a mark, the famous mark that will take different goals into account: ‘you knew how to sing, you knew the song by heart, you respected the rhythm that you were taught’” (teacher n°12). Teacher n°12 never gave a mark under 5: “When the student makes a mistake, it's not a problem.” Error did not result in penalization, but was a normal step in learning. The mark was explained to the students and the teachers tried to justify it by all means: “It's not a “gross mark”; I won't tell them ‘you got 16’ but ‘you got 16 because you gained 4 points for knowing

your song by heart and not making mistakes” (teacher n° 12).

Many teachers hesitated between “giving value” and “rectifying” the error. Teacher n° 2 said: “If I have a role to play, it’s to give value to students in distress, before anything that has to do with music.” He or she performed formative assessment towards students: “to me, assessment means marking, which is formative because I form them by giving them examinations” (teacher n°2). The mark also had other goals: “the mark is a help; a crutch because there are times when I’m dead and making an examination is a good solution” (teacher n° 2). They referred to their experience as a student and remembered the fear when tested: “the assessment also focuses on behaviour towards the others; the mark also takes into account the fact of having butterflies in one’s stomach” (teacher n°2).

The Student Behaviour is Important for the Assessment

The teacher used assessment to manage their class, and have control over their students: “Students appreciate when their work is rewarded, but on the other hand, we need more time for music; we have to make tests but we lose time” (teacher n°7). Assessment was considered a reward for the student’s efforts. Teacher n°7’s conceptions show they cared for the student. Already during the first interview they said: “assessing means taking into account the students’ efforts.”

In difficult classes, when teachers faced students’ opposition to singing, they were challenged: “sometimes I face a class of 14-15 year- olds who refuse to sing or do anything when I try to have them sing because they don’t want to open their mouths or when they do, it’s to shout” (teacher n° 7). Teachers looked for other ways to mark in order to take into account the students’ behaviour equally as much as their work.

These conceptions showed that these teachers associated students with the act of assessment. They did not intend to assess only their work, but other potentials. They marked the

students' behaviour in the class: "if the student doesn't do anything, he or she must be penalized, but if I see that he or she makes an effort, although his concentration is not perfect, I'll give him a mark nonetheless; I'll bring him into the group because he or she has made an effort" (teacher n°3). Often, the teachers felt uncomfortable when they have to give a mark. Their approach to marking took into account the students' efforts.

A Multitude of Assessment Practices

Music activity (singing, playing the flute, written test) was not assessed in the same way: "I mark four activities out of 5 points, and for the flute it's easier to mark out of 5 than out of 20; a mark of 16 or 16.5 out of 20 is subjective" (teacher n°2). This teacher proceeded with individual tests for the flute; he or she didn't reveal the mark but gave an oral appreciation. "Each student's mark will be brought back to 5 at the end of the term. The student who has the most points will have a mark of 5, the one who has lesser points will have 4, and the one who has no points will have 2.5 or 3 points" (teacher n°2).

Assessment was connected to marking regardless of activity (singing, playing the flute, auditioning), whether oral or written, whatever the classes. The way assessment was marked varied: out of 10 or 20, with participation points, etc. Marking an instrument seems easier than the activity of singing.

Two written tests per term allowed the assessment of cumulative information: "it doesn't take too much time; it's true that one must be quick in marking, it's marked straight away" (teacher n°12). This type of testing was not different for the other subjects: "I always told them to pay attention, that they'll have a written test, that I give a written test per term, I give a mark for the flute and one mark for oral participation. I give a written test because, since we have official instructions, in my opinion, we like to know what's left" (teacher n° 3).

Teachers n° 0 and 12 examined their students each week: three or four students were

individually tested at random, on the flute and in singing once per term. Teacher n° 10 gave several marks in one term: he or she examined his students individually with on the flute, singing, and the drums or rhythm, and gave a written test. This teacher didn't take the student's behaviour into account for the marking: "I won't rise or lower a mark with regards to a student's behaviour in class" (teacher n°10). They judged the performance "instantaneously" (teacher n°10). They didn't assess improvement, but pointed out errors of rhythm, notes, legato articulation, breaths wrongly placed, and breaks between sounds. Apart from this, they considered the student's posture, tone color, and the knowledge of the text: "I frequently put marks under the average; if nothing comes out, I give a 3 or a 4" (teacher n°10). Assessment focused on the final performance.

Others teachers were opposed to individual marking in singing: "they are extremely stressed when holding a microphone in front of the class, especially when they're 14-15, when it goes against their "locked up" psychology" (teacher n°2). They preferred marking the students' collective behaviour while singing. Students' participation in the lesson was taken into account for everyday assessment, and it was a collective assessment when they sang: "for the singing, it's always collective. It's the same thing during the first term, the whole class is examined and participations points are given" (teacher n°2).

Difficulties in Assessing Music

Teachers admitted the subjectivity of music assessment: "it's even more subjective in music than in mathematics. It's almost as subjective as French, where correcting an essay is inherently subjective and we can't do anything about it" (teacher n°12).

Because of such difficulties, some preferred to avoid marking oral practice "because criteria are even worse; some sing out of tune without being aware of it, therefore I cannot allow it. That is not possible." However, they do not give up and instead tried to find solutions: "I must assess the singing so I must find a way to mark it" (teacher n°0).

Some teachers asked their students about their difficulties: “What’s wrong? Can you make a self-assessment?” Teacher n°3 used self-criticism with all the students in every class in order to help them improve their singing performance.

The teacher faces a dilemma because they strive for perfection in students’ performances and hope they feel uncontrolled musical pleasure at the same time: “I feel it is a burden to assess at any price; in the end, students are fed up when they’re being assessed about everything and anything. I think they should have some freedom to vocalize and sing only for their pleasure” (teacher n°0).

Some teachers were helpless when they had to assess pleasure in music: “I don’t know how to assess that, and am unsure what kind of mark I can give when a student has really taken part in it. I’ll give him an 18, while giving the others a little less, maybe 12.”

The Place for Error and Musical Perfection

Musical perfection cannot be reached: “I never give a mark of 20 because music is never perfect” (teachers n° 3 and 12). The same teacher encouraged his students’ musical pleasure and perfection and was strict in marking in order to be acknowledged and legitimized: “I won’t give 20 to everybody because I don’t want to lose my credibility ; we’ll try to make something good that remains scholastic because if I don’t do that, I’ll lose credibility in front of my colleagues” (teacher n°0).

Perfection in music performance was not the goal: “students seldom reach perfection- without incorrect notes, or difficulties with breath support for the flute for example” (teacher n°12). Nevertheless, the piece which was perfectly played remained a reference for instrumental performance: “it will never be perfect: some play very well, some average, and some less” (teacher n°12).

This group of teachers faced the paradox connected with music: “In music, we’re always on a tightrope: we need it to be spontaneous and scholastic at the same time. If we are too scholastic, there’s no more spontaneity and if we privilege spontaneity, it becomes chaos” (teacher n°2). They did not see how they could give a mark to musical pleasure: “I cannot give a mark from 0 to 20 to a student who has had pleasure listening to music” (teacher n°12).

Implied and Misunderstood Criteria

Criteria for which musical performance was assessed included technique, specific to rhythm and pitch. Teachers had difficulties explaining their marking criteria: “in my mind, I have some marks for notes that have been respected; their names, the fingering, the rhythm” (teacher n°7). At the same time, music teachers described many different assessment criteria and their equivalent marks which relied on subjective judgments. Accuracy, musicality, performance, pleasure, emotion, beauty, and students’ musical ability were not “objective” assessment criteria: “can pleasure be given a mark? A mark for pleasure can only be given through appreciation: a motivated student is a student that is interested” (teacher n°12).

Psychological criteria were not explained to the students: “I make a classification of the students once I’ve examined them.” There were a large variety of criteria: “phrasing, rhythm, musicality, delicacy, breaths, articulation while singing, intonation, and paying attention to accuracy” (teacher n°0). The mark was mentally calculated: “I take some points out in my mind, 2 or 3 points maybe” (teacher n° 10).

Assessing musical performances often consisted of giving a mark to a student without any criteria, but from a general impression: “we could only give a definition of all these marks, without doing any examination; in other words, by the end of the term, we could give a judging mark to the student but it’s all subjective, I have the feeling we would reach the same goal” (teacher n°3). The interviews showed that belief: “I think I would be able to give a mark even

without giving written flute or singing tests... I don't know if I'm objective because I find it very difficult to give a mark" (teacher n°3). Musical criteria were purely intuitive and subjective. Although errors while playing the flute were penalized, teacher n°2 admitted he or she based their judgment on a general impression and on class behaviour: "I don't count anything anymore, well, that's my impression" (teacher n°2).

Only assessment through marking was observed. We have shown that marking was deeply rooted in music teachers' assessment practices. Regardless of the profile considered, final assessment remains summative and controlled by the teacher.

Stability in Conceptions: Comparison of the Tendencies

Answers to questionnaire 1b allowed us to note the changes that occurred on assessment conceptions within 6 months. The study of each individual teacher brought to light the evolution of each teacher's tendencies (see table8).

We noticed that all of the teachers' tendencies remained the same between the two Q-sorts, but that some variations appeared from one teacher to another: favourable conceptions for the assessment of the assessed (C) and unfavourable conceptions for the assessment of the assessor (A). The results of the two Q-sorts were analyzed more in-depth according to the types of profiles which have been determined. These results from the Q-sorts showed that the teachers' conceptions focusing on the assessor (A) remained the same after the second training day.

Profiles of the Assessment Concepts

The interviews made clear that two out of six teachers had a distinctive assessment profile: teachers n°0 and 10 had conceptions focusing on the assessor and the result, whereas teachers n° 3, 2, 7 and 12 had conceptions which focused more on the student.

Profiles Focusing on the Result and the Teacher (teachers n°0 and 10)

Responses from teachers n°0 and 10 show their reluctance to change anything about their assessment practices. They remarked that they hadn't changed their minds about their conceptions of assessment, which they often repeated during the interviews. This shows that they are not likely to change them. These two teachers worked in a rural zone, but hadn't taught for the same number of years: teacher n°0 was starting and teacher n°10 had been teaching for 10 years.

During the first Q-sort, teacher n°10 was favourable to an assessment focusing on the "assessed himself"; during the second, they preferred to focus on the "relation between the assessor-assessed" and still rejected "an assessment taking the context into account." This discourse mirrored what he or she answered in the Q-sort: they hadn't changed their mind between the two training days. During the interviews, teacher n°10 often repeated that he or she had neither changed their conceptions nor practice of assessment. They differentiated between teaching and marking, but would not change their mind on the matter of assessment: "I will not change my mind. I've been asking myself the question for several years; I find it difficult to give individual marks to the students and we lack time to do all the sequences we'd like to do." They continued to mark their students individually when orally examined "Why do I give 16 or 17? It's true that we could discuss this," "I like giving 10-11 years old a mark of, let's say, 16, which will be a very good mark. Little by little, some students manage to get 19 at singing." The teacher ended the interview by saying "I will not change my mind."

Teacher n°0's assessment conceptions privileged the relationship between the assessor-assessed in the first Q-sort, whereas they focused on the context in the second. Their conceptions conformed to the answers in the Q-sort; in other words, on an assessment only controlled by him or her. They didn't change their mind about an assessment focusing on the "assessor" after 6 months. This teacher was the only one to give marks to the students, and noted that his marking did not need to be questioned. A comparison of his two interviews shows

no fundamental change in assessment conceptions. For example, they gave the following definition: “assessing means giving a global appreciation, whether we consider all the criteria together or separate, and making an average- three “good,” two “fairly good,” it’s quite precise but not as precise as a mark.” On the second day, they also considered that marking was more precise than assessing: “assessing globally, with words... I see marking as something more precise.” This teacher did not change their approach to marking: “I’ll go on marking the same way as when I started teaching.”

The assessment conceptions of these two teachers placed more importance on marking and on the role of the teacher. We will not examine whether the other teachers in Paris suburban schools have more flexible profiles regarding the way they assess.

A Profile Focusing on the Student and his Progress

Although teachers n°3, 2, 7 and 12 didn’t change the way they assessed, they experienced the assessment patterns and tried to improve. They had unique profiles which will be described individually.

Teacher n°3 differs from the other teacher-trainees. He or she falls between the profile focusing on the result and that of focusing on the process. They had typical behaviours, had been teaching for more than 10 years in the suburbs, and was the only one who has passed the “Agrégation” exam. This teacher did not change assessment process focusing on the student. Their approach to marking remained the same in November when he or she stated: “I seldom give a 20, I tell them it can never be perfect in music, it can’t be,” and in April: “it’s to show them that music is fragile, it’s not a scientific, rational thing.” This teacher was preoccupied by the quest for objectivity. They stated that they didn’t change the way they assessed but, on the contrary, will be stricter: “honestly, since the first training day I’ve been thinking and practicing but nothing really changed, except the fact that I’ll penalize those who do not sing.” By the end of the second interview, this teacher wanted to change and penalize the students who do not

participate in vocal activities: “I give a global mark that is surely over-evaluated, but those children who do not sing will be penalized... I think there is a risk of putting individuals at the same level when giving a global mark to singing, but assessing means marking to me.” Unlike teachers n°0 and 10, teacher n°3 tried a system of self-criticism in which the students had to formulate the criteria to assess their singing. They first explained the criteria orally, and then listed them on the board so that the students kept them in mind. In the end, they realized that students sang better when practicing self-criticism. This teacher faced the paradox of giving a mark for pleasure and remaining in a scholastic system at the same time: “I tried performing music, having us spend a nice time together, having us do something that sounds good, like working specific phrasing. Not assessing, because it’s not a competitive exam- in assessing there is the idea of comparing.” Nevertheless, he or she remained doubtful and didn’t find any precise ideas or concrete solutions to assess music pleasure. They were unsure how to give a mark to pleasure: “nevertheless, we have to write down a mark on the end-of-term reports. I try to give them a taste for it, the pleasure of it- how can you assess pleasure? I don’t know, it’s very hard but I other... the student can compare himself with the others.”

Perhaps this unique profile can be explained by fluctuation between a will to listen to the students and to control of the marking.

We connected the profiles of teachers n°12 and 2 for their similar conceptions of assessment, although they hadn’t been teaching the same number of years (one was only starting, the other is close to retiring). Moreover, they were not teaching in the same conditions, as teacher n°2 taught in a secondary school qualified as “difficult,” and teacher n°12 in a secondary school considered to be “no problem.”

In the two Q-sorts teacher n°12 was favourable to an assessment focusing of the assessed. They were opposed to a conception of assessment focusing on the “relation between the assessor-assessed” in the first Q-sort, and to an “assessment on the context” in the second one.

Although this teacher tried to put into practice the assessment patterns learned during the training course, their assessment conceptions were connected with marking. Teacher n°12 seized the opportunity of our being present to try to put a system of self-assessment in place with his students based on a portfolio research. He or she expected to learn ways to assess singing from the training, and also ways to assess instrumental works, and discuss assessment with the other trainees. After the first day, the teacher didn't put the new information into practice: "I didn't use it and I don't see how to." After the second day, they put this experience into practice for the first time. They practiced assessment by marking, which suited them. The teacher thinks he or she should continue on this way of assessing "but it's always the same thing, we must end up with a mark." They were open to change: "I think one should properly think the question over when one has never proceeded with such an assessment." However, the second interview showed the return to previous habits: "although the substance won't change, the way I'll present things will." This teacher wasn't convinced with the way students could be assessed: "until now, I haven't found any end-of-term report in which to put letters or simply an appreciation; I think it would be looked down upon." Moreover, he or she found it hard to question themselves after twenty five years of teaching: "I'm not sure of the way colleagues assess the children, somehow, it perplexes me. Maybe I don't have a good ear for music, maybe I pay more attention to the individual than to the class."

Assessment conceptions of teacher n°2 focus on the assessed in both Q-sorts. Teacher n°2 didn't change their conception of "an assessment focusing on the assessor himself." His or her answers were coherent, and the way of assessing focused on the assessed. This teacher's assessment conceptions were similar to those of teacher n°0: "assessing a student could mean seeing him in his entirety; marking, meaning marking something very precise." (teacher n°2). They didn't look for perfection in performance but aimed to have students discover the pleasures of music. Their marking criteria didn't change between the two interviews: When I

give 2.5 to someone, it means it's not that good, someone who gets a 3 is average, someone who gets 3.5 it's more accurate, 4 is good, 4.5 is very good and 5 is flawless." They considered their practices as summative, but didn't feel like changing them: "In fact, I think it's final because the marking is out of 20. Anyway, we're asked to mark that way, one can go against it but I don't feel like it because it's kind of easy." This teacher was aware of the paradox between the quest for objectivity and the marking criteria which were subjective: "If there is a bad feeling between a teacher and a student, there's a lot of subjectivity present- we qualify as objective something that is completely subjective, regardless if the student likes us or not, or likes the subject." However, he or she kept control over marking: "I just realized something: I always want to have ascendancy over my class, total control, and if I delegate, nothing will prevent the person to give points to everybody." They were willing to remain objective to the extent that they were as fair as possible to the students. We may wonder if the teacher will change next year. He or she thought that for the end of this school year, it's too late. They hadn't changed their practices but remained open to the possibility because they have just started to teach: "I'm in a frame of mind that makes me think that I'll change next year, but not now; I don't need it here because I don't have awful classes." They will adapt themselves if the students lack discipline.

Out of the six teachers that we met, teacher n°7 appeared to be flexible and willing to adopt a formative assessment. Teacher n°7 is a female teacher with more than twenty years of experience, teaching "difficult" classes in a Priority Education Zone (Z.E.P.) secondary school. She is always looking for solutions, as she is confronted daily with scholastic failure, discipline and violence in her classes.

Teacher n°7 is more favourable to a "relation between assessor-assessed" although her conception is in favour of the assessed in her second Q-sort. She rejected an assessment focusing on the assessor in both Q-sorts. Her answers to the Q-sort changed, and her discourse

was open to change in the future. Her assessment was of a summative kind and the marking allowed her to manage class discipline. However, she was the only teacher who was convinced of the benefit of formative assessment from the beginning of the training. “During the training I learned to explain how to proceed when working on singing.” The training met her expectations: “I was pleased with it. It was positive because I found things I had been looking for. I was satisfied.” She had always wondered about the problem of assessment “so the students are aware of the marking and how I end up with the mark.” She seized the opportunity of our presence to experience the brainstorming card: “What does one need to do to sing well?.” Nevertheless, she said she would improve her assessment patterns next year: “I’ll try to do it next year, although this year, even if I haven’t handed the paper over, I tried to see all the classes. I had them write at the end of their copybooks what they should do to sing correctly, whatever their level.” The assessment cards would help her. She thought the assessment patterns must be built up with the students: “My goal with this card is that they understand more my way of marking.” She faced many questions regarding the patterns of self-assessment: “is a child aware of having respected the pitches? Those who sing out of tune, they don’t hear it, it’s an evaluation from the teacher; I’m the one who hears if they have respected the pitches or not.”

An interpretation of these profiles showed that, generally speaking, conceptions hadn’t changed in six months, which was confirmed by the interviews. However, if at first sight conceptions favour the assessed (what is also confirmed by the interviews), we noticed a difference between what each teacher thought was ideal and what he or she really did in practice. This didn’t appear in the results of the Q-sorts, so other methodological tools seem to be needed.

To conclude, the teachers we interviewed adopted specific profile characteristics, and displayed the diversity of music assessment practices. Variables such as the number of teaching

years (teachers n°12 and 7), gender (teachers n°0 and 7), type of school (teachers n°12 and 7), or status of the teacher (teacher n°3 compared with the other teachers) are not enough to explain the stability of the conceptions and the reluctance to practice formative assessment. Discourses are dominated by the connection of assessment with marking controlled by the teacher.

The results of the investigations tended to show that the trainees did not change their conceptions between the two training days: “among what was suggested, there may be things that I won’t keep, for sure, there are too many elements to make an assessment, it’s too complex” (teacher n°12). The teachers thought that the student was the one who sought self reflection to understand why he or she received a bad mark. The person who has the most teaching experience admitted it was more difficult to change than when you have just started to teach. Still, we can say that neither the age nor the number of years of experience explained the behaviours we described, which reproduce the habits of the scholastic system that formed them.

Yet, we can say that it’s a group of teachers who question themselves and look for training: every year they asked to be registered for a training course. Then, for what reasons do these teachers appear to have changed so little their marking practices in class during the training (apart from what was asked by the trainer)? What other reasons can explain that deep reluctance?

Transversal Analysis: Reasons Connected with Training

Only the interviews allowed us to understand the trainees’ real expectations and motivations. These two factors appeared to be important for the transfer of skill.

Motivations for Registering

It is the word “assessment” in the title of the training course that first caught the attention of the trainees; the word “formative” coming in second place. Teacher n° 0 focused

on the word “assessment” in the title of the training because “it was the only training that talked about assessment.” Upon registering, trainees thought they would collect “ideas for their examinations.”

The teachers who registered for this training course were willing to change the way they assessed by varying their practices: “it’s OK but I have to diversify my inclinations.” They thought they would make a new start with this training. Teacher n°12 wished to “better” at assessing his students, and to talk with the others about assessment: “How does one assess singing and instrumental work? How can teachers succeed in giving marks ? I came to the training course thinking that I would learn methods that would have allowed me to teach things to the students by examining them” (teacher n°2). Training was a good occasion to talk with colleagues, to break the professional isolation and to get out of the school, since music teachers are often the only ones at their school.

Finally, by registering for the training course teachers were able to take stock of their situations, and get “concrete” information, and to know if the way they are assessing is correct, which in turn allows for self reflection. “If I registered for the training course, it was to get some ideas from here and there” (teacher n° 10).

Expectations about the training

We can say that music teachers’ expectations about the training influenced the practice of the assessment tools that were suggested (binomial, one or more ...). The teachers expectations were varied and included : the possibility of discovering new assessment tools (teachers n° 3, 10 and 12), methodological reflection (teacher n°2), and getting more knowledge about assessment (teachers n° 0 and 7). Expectations were also very specific as, for example, learning methods in giving marks based on creativity, and finding a marking scale for the different types of errors (teacher n°10).

Expectations about “formative assessment” meant getting concrete models of examination in order to mark the singing more objectively, pedagogical methods, or, in other words, “recipes.” They admit they needed visual information, “concrete” is an adjective that constantly came back: “I expected cooking recipes to be more efficient when assessing.” (teacher n°2). They lacked ideas, substance, lessons, and pedagogical tools.

Although the training didn’t fulfill their expectations, teacher n°10 was convinced that “it would be a good thing for the students to know how we mark them” (teacher n°0). “I wanted to have methods to assess my students and wanted to know what kind of examination I could give according to the class” (teacher n°2). The fact that their expectations were not fulfilled could explain why the teachers had difficulties in following through with the training and consequently, explaining the stability in their assessment conceptions : “I question myself enough so it’s a pity that every time I attend a training we talk about the opposite of what I expected” (teacher n°2). “May be I misunderstood” (teacher n°0).

Analyzing the teachers’ discourse “leitmotifs” showed that professional changes could be connected with the context and training expectations. When the substance of the training didn’t fulfill the teachers’ expectations, they had no reason to change their assessment.

Opinions Regarding the Training

The teachers enjoyed the talks and briefings during the second training day: “we were more confronted with our personal experiences; comparing our difficulties and talking about our problems and how to resolve them is always a way of improving” (teacher n°10). The substance of the second training day was seen as less theoretical than the first day. The teachers preferred the reference to concrete situations than the part concerning theoretical notions about assessment. They found the second day more interesting “because we worked on sequences from some teachers” (teacher n°10).

Moreover, the assessment patterns were criticized: “during the training, there were patterns; it was crazy, I felt there were parapets everywhere, I think there should be more freedom to allow the children to have more spontaneity because sometimes they surprise us” (teacher n°3).

However, these critics showed that the training’s objective was not understood: “it was extremely theoretical ; I came here hoping I would learn how to asses my students with more rigor” (teacher n°2). The trainer’s vocabulary surprised some teachers: “the trainer talked to us in words of his own; after, he or she adapted to our ways but at the beginning, it seemed a little masterful, maybe too much” (teacher n°7). This explained why the vocabulary used to qualify assessment – formative, summative, self-assessment, pedagogy by goals, differentiated pedagogy, referee, referent, etc – that was used during the training was difficult to define during the interviews: “Everything has been looked at in a general way, assimilation needs time” (teacher n°7).

In particular, the notions of summative and formative assessment seemed misunderstood: “The difference between formative and summative is that we force children to answer a question, the lesson’s substance, whereas summative assessment means giving an authoritative lesson and asking the students to work on this or that for a given date and if they don’t, they’ll get a 0” (teacher n°2). Formative assessment seemed to mean that students would be informed on the criteria on which they will be judged. “To me, assessment means marking and formative because it forms the students, I form the students by giving them examinations” (teacher n°7). Teachers had a hard time trying to make a distinction between formative and summative assessment: “‘Summative’ is a mark given at the end of a work whereas ‘formative’ is judging in a more general way” (teacher n°3). The notion of “formative assessment” had not been understood, or has been interpreted as the clarification of the marking criteria for the students (teacher n°3).

If these factors seem to influence the evolution of the conceptions of assessment, it is not easy to find a unique compromise between opinions, expectations, and conceptions. Still, a positive factor appeared crucial: the interviews between the training days.

Attainments of a training about formative assessment

During the interviews, each teacher could explain what he or she had put into practice in their assessment. Whatever their ages and experiences, few teachers changed their conceptions and assessment practices. So, what remains of what they learned about formative assessment? The following are some positive images about formative assessment that the teachers remember. The training allowed them to discover:

- The brainstorming technique: “What is needed to sing well? What is wrong?”
- The principle of co-assessment: “It was the interactive part of the lesson, I insisted a lot on self-criticism, but although it doesn’t end up as an assessment with a mark, it’s an attempt for the group to make progresses in the song’s performance” (teacher n°3).
- The student’s self-questioning and interaction during the lesson: “How can one assess individually and give a mark? I do not know how to mark individual singing” (teacher n°3).
- The explanation of assessment criteria: “you explain the way you proceed regarding the work of singing, even if it wasn’t written” (teacher n°7).
- The tools for assessment: “there was the document, which was interesting, but we already had the programs” (teacher n°12).

The teacher’s opinion of their knowledge gained showed how he or she would implement these practices in class. Often, they said they had acquired assessment patterns, but they did not adopt the process of requiring the students’ cooperation: “I would rather impose it on them from the start” (teacher n°10), “students will not proceed with self-assessment, they’ll tell me they were concentrating, so I’m the one who decides, ‘you were not concentrating’ so I’ll give you a

2.” Then, we come back to the same story of marking” (teacher n°10).

At the end of the training, they found themselves explaining the way they mark more precisely (teachers n°3 and 12). However, this is not the goal of formative assessment! Their ideal assessment process remains directing and is based on deep representations: “ideally, I should examine the students differently throughout the year so that they don’t know in advance how I’m going to assess them; I don’t know if it’s positive because they would be in an uncomfortable situation the whole time” (teacher n°2). They assessed the students without knowledge. This is far from transparent co-assessment or self-assessment in a music class.

Discourse Regarding the Transfer of the Tools in Class

As a matter of fact, few teachers said they put the knowledge they acquired from the training into practice. Only two of them took the occasion to put into practice what they had understood during the training, and only one applied self-assessment based on the brainstorming technique (“What is needed to sing correctly?”) (teacher n°7). Is this an isolated experience, or the beginning of a more important change? It cannot lead us to conclude to a change in practice.

Help to Allow Conceptions to Emerge

The device allowed us to put out paradoxical attitudes from the music teachers which might explain their reluctance towards formative assessment.

We noticed contradictions between an ideal representation which focused on the student and a representation turned towards summative assessment and marking. Both attitudes could be found in the same teacher during the two steps of the investigation.

The “ideal” conceptions emerged as part of the music teacher’s opinions. In the questionnaires, teachers said they changed their practices in marking and felt the training had allowed them to change some of their professional practices regarding assessment. Yet, they felt

equally “doubtful” and “skeptical” towards formative assessment (teachers n°2, 3, 10 and 12).

Whether “convinced by” (teacher n°7) or “militant for” (teacher n°0) formative assessment, the teachers thought it is useful to contemplate the complexity of music assessment practices (teachers n°0, 2, 3 and 7). Most of the teachers felt their methods of assessment changed since the beginning of their career (five teachers). Teachers n°2, 12 and 7 said that assessment allowed them to question their teaching. They considered that assessing their students meant assessing them often (five teachers) or “occasionally” (teacher n°0). They wished that all their students (or most of them) succeeded in music (teachers n°0, 3, 7 and 12). Others were more careful and thought that success could only be reached step by step (teachers n°2 and 10). Besides this, answers in the questionnaire showed that some teachers found it necessary to have a variety of assessment tools (“that I can be sure of,” teachers n°0, 3 and 10); others felt they were essential (teachers n°2 and 7,) and only one teacher replied: “why not?” (teacher n°12).

Nevertheless, their opinions on formative assessment did not mirror their real conceptions. Representations that emerged from the music teachers appeared clearly during the interviews; the music teachers’ conceptions seemed to focus on summative assessment and marking, which is always apart from the learning (“assessing means marking”). We notice that students were seldom mentioned during the interviews.

On the other hand, the mark represents authority of the music teacher in a secondary school. The teachers tend to judge the students’ behaviour more in class than in musical work. We can see that the teachers try to give meaning to their marking, which they defend by rationalizing it. This justification of the marking that relies on subjective elements of judgment appears in ambivalent and contradictory discourses. “Ideal” conceptions of assessment show a wish that students feel pleasure towards music and “real” conceptions claim the necessity of “scholastic” control.

Preconceptions of formative assessment

When teachers were not convinced of the validity of formative assessment, it was very difficult to have them change their minds: “I know there is a truth to a part of what he or she says, but I don’t find that more objective than what I’m doing” (teacher n°10).

What could be the reasons for this reluctance to change? The teachers’ arguments show several explanations:

There are plenty of preconceptions to justify reluctance. Students were mentioned as those who would be reluctant to change: “they don’t like to say that it was good, they don’t feel at ease. It’s because they’re not used to it, they’ve been used to be given a mark as a mouthful, it’s hard for them to free themselves” (teacher n°3). Some teachers fear that self-assessment will bring anguish to their students: “we feel that they are tense or strained. They frown, their performance is not as good as usual, and I’m afraid that knowing they’ll have to give themselves a mark will lead to a second strain” (teacher n°12).

Teachers were skeptical because they have not tried the experiment in their classes (teacher n°0). They admitted the challenge in adopting their practices to what they have learned during the training: “I really would like to get out of this system. However, on one hand, I’ve been completely formed like that and on the other hand, towards my colleagues, we’ve only got one hour, it’s difficult” (teacher n°2).

Another preconception is that formative assessment should only be adopted when teaching in a “difficult” secondary school. Formative assessment was only useful in classes in which traditional marking doesn’t work, and is adapted only to “difficult” classes. “From a general point of view, collective assessment and assessment of students in small groups or self-assessment are somehow connected with difficult situations” (teacher n°10). Formative assessment would be useful only in cases when the teacher’s authority is questioned. When

everything works well with traditional assessment methods, why change?

On the other hand, the level of musicianship would diminish with formative assessment : “I want them to have an acceptable level in singing knowing that I won’t give them individual markings, it’s a risk, they’ll sing in a group : out of three, one will fake singing, the two others will really sing” (teacher n°12). Some teachers feared that the subject “music” itself might be sacrificed: “the device of self-assessment means compromising musical teaching, it’s not that difficult to self-assess but we’ll end up with a subject completely devalued, in comparison to the others” (teacher n°10).

Although these teachers were convinced of the benefits of formative assessment, they did not want to lose their authority: “I just realized something: I always want to have power over my class- total control, and if I delegate, nothing will prevent the person from giving points to everybody,” “It may be true that having students participate in their own assessment might have a positive impact on their behaviour, but for the time being, I’m only trying to take care of my health. It’s extremely selfish” (teacher n°2)

Above all, they were afraid of losing control over their classes by letting the students’ take part in the assessment. Teachers possessed the power of giving a mark, and were not willing to allow the student to experience self-assessment: “they need a standard from an adult” (teacher n°3).

On the other hand, teachers often mentioned the school system, and the obligation to give average marks. They also explained the lack of time they had to expound upon the permanence of their assessment practices: “questioning one’s own assessment method means losing momentarily control of the class. We are in a period of change, and it’s dangerous because students might feel we are unstable. I use the word ‘dangerous’ on purpose” (teacher n°2). This may be why some of the teachers mentioned their fear that time spent assessing might take away from the time allotted for teaching: “next year, it won’t be better, I’ll try but I’m not

convinced with the results. I'll try to see if I spare some time, but if I lose time and if it doesn't bring me anything in terms of results, I'll drop it" (teacher n°12). The practice of feedback is separated, cut off. The teachers' essential concern is to spare time by assessing students collectively instead of individually.

Formative Assessment Put into Practice

The results of putting formative assessment into practice remain unknown. Although teachers showed their goodwill and tried to put self-assessment patterns into practice during next school year, "after, it depends on the efficiency and the possibility to put it into practice; I need to experiment with it to see if I can bear that supplementary load of work, and if it will prejudice my teaching" (teacher n°0). We can see their fear that assessment may take time from teaching. Some teachers criticize: "criteria on paper are very good, but remain extremely scholastic. I don't say that being scholastic is negative, as I always feel that in music, we go beyond that" (teacher n°2).

Moreover, the training course was held during the school year and teachers might have felt it difficult to change their assessment practices in relation to their students: "maybe next year..." They have settled into habits with their students and they feel that changing during the school year might discredit them. Do they fear losing their credibility?

The results of the interviews seem to confirm that a short training course does not have an impact on the change of conceptions. The answers to the Q-sorts and the questionnaires did not allow us to show if the teachers had changed their assessment practices. The training did not have a "real effect" leading to a radical change, or a deeper self-questioning from the teachers. The trainees say that the training hasn't changed anything in their habits regarding assessment, as what was said during the interviews remains unchanged from their habits. This confirms the notion that a training course on formative assessment needs time to be integrated in music teaching, in order to change deeply ingrained habits in the tradition of marking.

Limits of an Investigation on a Training Course About Assessment in Music

Statement on the Contribution of the Self-criticism Interviews

The pedagogical follow-up after the training was a real process of self-criticism. The “mirror” interview allowed each teacher-trainee to talk freely and analyze his evolution during the training, without any judgment on value, and without any control. It seems to be a good method of self-criticism. “It helps teachers to develop a more precise idea of their own methods, offering them a mirror-image of what they are doing.” The investigation of the training course about on music assessment existed on three levels: observation of the training days, observation of the teachers (before, during and after the training) and observation of the classes (before and after training). The self-criticism protocol seemed positive for the teachers. It allowed us to follow-up with the trainees regularly, and to control the effects of the training with them. It was the first time in continuous training that they were asked about their assessment practices, and the interviews were positive: “my talking with you led me to think things over and I realized I was too direct” (teacher n°2), “my talking with you made me improve” (teacher n°12). Assessment can only be performed with the teachers’ help and consent, as teachers do not get involved with what is imposed on them. They want to associate with and participate personally in the elaboration of projects which concern them. In order for all of them to accept and hold to the principle of assessment, we must first inform and convince them of how personally, socially and pedagogically useful assessment is.

Methodological criticism of the investigation

By knowing what the participants expect from the course, we may be able to better predict the extent of its influence. Teachers showed contradictory attitudes towards the definition of assessment. Teachers also misunderstood the “general” concepts of assessment. This experience allowed us to understand that the questions and items of the Q-sort were not adequate for the realities of teaching. Indeed, the investigation showed that teachers

proceeded with unique kind of assessment: “summative” assessment through marking. The different notions of assessment – summative and formative – had little effect on the music teachers’ conceptions, which is why we were partially satisfied with the first results of the investigation. We had little information for the reasons that influenced teachers to choose certain items of the Q-sort. The methodological tool showed weaknesses in its conception: items and categories must be more precise if they are to show the reality of music assessment. We were not able to adapt the general concepts of the vocabulary from the sciences of education to the musical field, and this created a deadlock. For example, the categories of items concerning general assessment used in the Q-sort (A,B,C,D) look artificial and were not distinctive enough. These items were taken directly from Q-sorts referring to assessment in teaching (De Peretti, 1998) and did not permit us to classify the assessment profiles.

Definitions looked very much alike:

The item “Admitting that an error is not a fault; replacing the notion of failure by the difficulty of overcoming it, the notion of fault by the notion of error” belongs to category C and the item “Bringing into favor that error is a necessary step of music learning” also refers to the concept of error but is registered in category D.

This is only an example of the limits of the Q-sort conception above; the separation of the items in the four groups seems arbitrary and unrelated to the reality of music teaching.

On the other hand, the concepts of the sciences of education can hardly be applied to the art of music without being altered. Many data from the questionnaires were not expounded upon because they were considered to be too “general.” For example, questions referring to conceptions as an assessor did not seem adapted to music teaching. Interviews gave us more information on the degree of reluctance to assessment from music teachers after training. At the time, we had to make methodological choices, and at the beginning of this research we looked for a qualitative and quantitative method which would allow us to gather personal

conceptions. The Q-sort corresponds to these criteria. However, it seems to us today that a clinical method would be more adequate and correspond more to the complexity of conceptions regarding music assessment. The Q-sort forced us to form categories and create a table with different points of view.

This research, based on definite criteria, attempted to determine assessment conceptions. Focusing on the interviews and observations at different moments, we have described the evolution of teachers' conceptions. The aim of this research was to bring to light the trainees' conceptions, and examine the factors connected with the assessment of a continuous training course. I am aware of the limits of this work, especially methodological, but the results of this research may initiate a dialogue about assessment.

Notes

¹ NUNZIATI, G., (1987). Evaluation des actions de formation n°2. (p.45). *Les cahiers de l'évaluation GREPS*

² "Process through which people, groups, community adopt new ways of acting and thinking such as reactions and emotive, moral, aesthetical and spiritual behaviours, deep and authentic " TOUPIN, L., (1995). *De la formation au métier : savoir transférer ses connaissances dans l'action*, (p.196). Paris, ESF.

³ Conference (1986). *Evaluer l'évaluation* Dijon. INRAP, p.33.

⁴ As Mr Seltner mentions it in De Peretti, 1998, p.497, the Q-sort is the "sorting of qualitative statements."

⁵ The numbers were those given to the teachers when going through the Q-sorts.

⁶ Item is defined as an affirmative statement.

Appendix A

Pattern of coding for questionnaire 1a

1 - What is the purpose of this training?

1-1 The reason for registering is connected with (tick maximum two boxes):

- A personal choice
- A need
- An obligation
- An interest
- Curiosity
- Other

1-2 Additional reasons for your participation: your goal is more to (tick maximum two boxes)

- To build pedagogical tools.
- To find out about new ways of assessing.
- To eventually use them.
- To answer personal questions.
- To use them in a collective project.
- At the request of your administration.
- Other

1-3 The place and the dates of the training were important factors:

- very much
- more or less
- little
- not at all

1-4 You expected to: (grade the following affirmations from 1 to 10)

- discover new assessment tools
- carry out a reflection on methodology
- find a new approach with regard to traditional assessment
- deepen your knowledge on assessment
- experience group work
- have a chance to break with habits
- get a new knowledge on assessment
- have the occasion to question yourself
- have the possibility to communicate with the group
- other

2- Assessment

2-1 How would you qualify your assessment practices?

- they are OK
- I need to moderate my tendencies
- I need to vary my practices
- I expect a « a new start » from this training

2-2: Assessing my students also means assessing myself:

- seldom
- occasionally
- often

2-3 Is it necessary to have a large variety of means and tools to assess students who are very different one from another?

- why not
- it's too difficult

- I don't believe in it
- of course
- it is essential

2-4 Would it be normal and necessary for all the students (or most of them) in each class to “succeed” in music?

- it is not possible
- I wish it
- it annoys me
- we can try to reach that
- it is poorly considered
- it is possible, progressively

2-5 Would you say that:

- marking promotes competition
- marking ruins students' relations
- marking is emulating
- marks ruin the relationship between students and teacher

2-6 According to you, is the general practice of marking of our time?

1) YES 2) NO

2-7 Out of the following six categories, what is the one I prefer as corrector?

For me, teacher, correcting can be (tick only one box):

- giving value
- giving references
- rectifying
- sanctioning
- normalizing

- choosing

2-8 Frankly speaking, how do I see myself now as an assessor? I consider I'm more (tick maximum three boxes):

- lenient
- mindful
- light
- indulgent
- focusing on the students
- sometimes irregular
- sticking to routine
- objective
- tolerant
- fair
- indifferent
- balanced
- focusing on the programs
- sometimes inaccessible
- casual
- harsh
- impatient
- very touchy
- too conscientious
- often enigmatic
- focusing on musical discipline
- sometimes aggressive
- tired

Appendix B

Pattern of coding for questionnaire 1b

A After the first day ...

1. With regards to formative assessment, you feel you are ...
 - 1) doubtful 2) convinced 3) militant 4) deprived 5) hostile
2. Do you feel you have changed your way of assessing the students since you started teaching?
 - 1) YES 2) NO

If yes, in which activity (ies): 1) a music activity, 2) all of them
3. Do you feel it is useful today to think through the complexity of music assessment practices? 1) Hmm 2) Why not 3) Of course
4. Do you think this training will allow you to change some of your professional practices?
 - 1) YES 2) NO

If yes, in which activity :

 - 1) Conception of lessons and putting them into practice
 - 2) Practice of assessment
 - 3) Use of the documents and of their results
 - 4) My pedagogical behaviour

Autre(s) type(s) de modification(s) Other kind(s) of change(s)

B Who are you?

1. Gender : 1) *male* 2) *female*
2. Year of entrance in the National Education: 0) *before 1977* 1) 78-82 2) 83-87
- 3) 88-93 4) 94-98.
- 3- Birth date: 1) 1970-75, 2) 1965-69, 3) 1960-64, 4) *before 1959*.
- First teaching year in this field: 0) *before 1977* 1) 78-82 2) 83-87 3) 88-93 4) 94-98.

Level : 1) MA 2) QUALIFIED 3) *AGREGE (competitive exam)*

Seniority : 1) less than 10 years 2) more than ten years

Teaching only music : 1) YES 2) NO

C Where do you teach?

1 Type of school (s) : 1) Secondary school 2) High school

2. In which place 1) Suburbs 2) ZPE

D What was your training?

1. Diploma: 1) CAPES/Licence 2) Master 3) PhD/"Agrégation" competitive exam

2. Music diploma : 1) YES 2) NO

If yes, which one(s)

3. Did you have another profession before teaching? 1) YES 2) NO

E Continuous training

1. Did you already register for a training regarding the Academic Training Schedule?

1) YES 2) NO

If yes 1) about music teaching 2) trans-subject trainings 3) both

Year : 1) before 1994 2) between 1995 and 1998

2. Other continuous trainings in other fields? 1) YES 2) NO

If yes 1) in the scholastic field

3. Do you have practice as a trainer in the National Education? 1) YES 2) NO

4. Do you have practice as a trainer outside of the National Education? 1) YES

2) NO

Table 1

The experience's plan

Before November 10th 1998	On November 10 th 1998	Between November 10th 1998 and April 15th 1999	On April 15th 1999	After April 15th 1999
Questionnaire 1a (see appendix A) and Q-sort	Observation of the training and questionnaire 1b (see appendix B)	First interviewing campaign and observations.	Q-sort for the end of the course.	Second interviewing campaign and observations.

Table 2

Q-Sort Items

A/ The assessor himself

Music assessment means :	Item n°
wondering on the meaning of what one is doing	1
giving a mark to each student according to his music ability	5
sanctioning the student's behaviour in class	9
know to wait even if it's a loss of time.	13
assessing one's self	17

B/ The relation between the assessor-assessed

Music assessment means :	Item n°
using one's power to reinforce one's own authority	6
observing students in the process of learning a knowledge or a know-how	10
comparing achievements with forecasted goals	14
wondering first of all about what happens during the didactical relation	15
taking stock with the student of what are his acquisitions at the end of a teaching sequence	19

C/ The assessed himself

Music assessment means :	Item n°
allowing the youngster to avoid failure	2
giving value to the youngster as an individual	3
admitting that an error is not a fault ; replacing the notion of failure by difficulty of overcoming it, the notion of fault by error	4
allowing the student to perform self-assessment	11
observing the student's spontaneous activities	16

D/ Assessment takes into account the environment and the context:

Music assessment means :	Item n°
allowing the learners to reach autonomy	7
connecting the youngster's performances with what is expected from him in society	8
outlawing elitism	12
taking into account social inequality	18
Bringing into favor that error is a necessary step of music learning	20

Table 3

Number of Q-Sort distributed

Date	Number of Q-sorts distributed
on November 10th 1998	14
on April 15th 1999	9

Table 4

Planning of interviews

November 10th 1998	First training day
December 1998	One interview (teacher 2)
January 1999	Six interviews (teachers 0, 3, 6, 7, 9, 12)
February	One interview (teacher 10)
April 15th 1999	Second training day
May	Five interviews (teachers 0, 2, 3, 7, 12)
June	One interview (teacher 10)

Table 5

Age, experience of teaching of Music teacher interviewed

Teacher n°	Age	Number of teaching years
0	29	7
2	30	3
3	34	10
7	43	21
10	37	10
12	48	26

Table 6

Results of Q-Sort 1

<i>Conceptions based on :</i>	<i>Favourable tendency</i>	<i>Unfavourable tendency</i>
A The assessor himself	10	0, 2, 3,7
B The relation between the assessor-assessed	0, 7	12
C The assessed himself	2, 3, 12	No answer
D Assessment taking into account the environment and the context	No answer	10

Table 7

Results of Q-Sort2

<i>Conceptions based on :</i>	<i>Favourable tendency</i>	<i>Unfavourable tendency</i>
A The assessor himself	No answer	0, 2, 3,7
B The relation between assessor-assessed	10	No answer
C The assessed himself	2, 3, 7,12	No answer
D Assessment taking into account the environment and the context	0	10, 12

Table 8

Comparison between data of Q-Sort1 and 2

Teacher n°	Favourable tendency on November 10 th 1998	Favourable tendency on April 15 th 1999	Unfavourable tendency on November 10 th 1998	Unfavourable tendency on April 15 th 1999
0	B	D	A	A
2	CD	C	A	A
3	C	C	A	A
7	B	C	A	A
10	A	B	D	D
12	C	C	B	D

References

- Astolfi, J.P. (1997). *L'erreur, un outil pour enseigner*. Paris: ESF.
- Begue C. & Sisca C. (1999/2000). *Comment sensibiliser les élèves à l'auto-évaluation dans les différentes activités du cours d'Education musicale en 6^o?*. Académie de Rouen. Mémoire de PLC2, Education Musicale.
- Cardinet, J. (1989). Evaluer sans juger. *La Revue Française de Pédagogie*. (88: 51).
- De Ketele, J.M. (1997). *L'évaluation approche descriptive ou prescriptive ?* Bruxelles: De Boeck (2nd ed.).
- De Peretti, A. (1998). *Encyclopédie de l'évaluation en formation et en éducation*. Paris: ESF.
- Hadji, C. (1992). *L'évaluation des actions éducatives*. Paris. Presses Universitaires de France.
- Hewitt, A. (2003). Levels of Significance Attributed to Musical and Non-Musical Factors of Individual Difference by Classroom Music Teachers. *Research Studies in Music Education*, 20, 48-59.
- Hunter, D. (1999). Developing Peer-Learning Programmes in Music: Group Presentations and Peer Assessment. *British Journal of Music Education*, 16, 51.
- Kaufmann, J.C. & Singly, F.D. (1996). *L'Entretien compréhensif*. Paris: Nathan.
- Le Diagon- Jacquin, L. (1999) L'évaluation en cours d'Education musicale. *L'éducation musicale*, Mai.
- Maugars, C. (2003). Des enseignants en musique s'auto-évaluent. Leurs conceptions sur l'évaluation: Revue de questions à propos d'évaluation formatrice en Education Musicale. Paris, Université de Paris-Sorbonne. Document de Recherche de l'O.M.F. Série Didactique de la musique, 26, Juin.
- Maugars, C. (2004). De la Notation à l'évaluation en Education musicale: des représentations d'enseignants. *Thèse de Doctorat, Université Paris IV-Sorbonne, non-éditée*.
- Mills, J. (1991). Assessing musical performance musically. *Educational Studies*, 17: 173-81.
- Nunziati, G. (1987). Evaluation des actions de formation n°2. *Les cahiers de l'évaluation GREPS*, 45.
- Terrien, P. (2000). L'évaluation en question dans l'Education musicale. *L'éducation musicale*, Sept- Oct.
- Toupin, L. (1995). *De la formation au métier : savoir transférer ses connaissances dans l'action*. Paris. ESF.

