Assessment *Of* and *For* Learning in Higher Education: From the Traditional Summative Assessment to the More Emancipatory Formative and Educative Assessment in two faculties of the University of Porto

Manuel Firmino Torres[1] and Carlinda Leite[2]
University of Porto, Portugal

Theoretical framework

Higher education has become one of the most prominent topics of research in recent years, especially since it has been considered a strategic factor for economic development in which accessed new social and cultural groups as well as an increased number of students (Altbach & Engberg, 2006). The implementation of the Bologna Process has contributed to this situation, particularly in Portugal, forcing Universities to face significant changes and challenges, which imply the need to rethink and renew the existing concepts of teaching, learning and assessment (CEDEFOP, 2009; Boud & Associates, 2010). This transition scenario justifies studies that contribute to more in-depth knowledge about the changes that occurred in higher education, particularly regarding assessment and its association with learning.

Since the early 1970s, the link between assessment and learning has been studied. Now it is widely accepted that assessment is probably the single biggest influence on how students approach their learning (Ramsden, 1992; Gibbs, 1992; Brown, Bull, & Pendlebury, 1997) and clearly indicates that assessment has an impact in student learning (Boud & Falchikov, 2007). In this perspective, several authors argue that assessment has a significant impact on all aspects of students' experiences, satisfaction, outcomes and success (Ramsden, 2003; Richardson, 2005) and that both good and bad practices affect student engagement with learning (Prosser, Ramsden, Trigwell, & Martin, 2003; Coates, 2005). Therefore, assessment is significant to students' experience and engagement with the course (George, Cowan, Hewitt & Cannell, 2004) and is a key component of the learning process (Levia & Quiring, 2008).

From the conceptual point of view, we accept "assessment is probably the most important thing we can do to help our students learn" (Brown, 2004, p. 81). We also assume that there is a strong relationship between the way teachers assess learning and the way students organize themselves, get involved in their training and finally learn. Thus, it seems that it is necessary to use assessment methods appropriate to the aims and objectives and to the contexts in which the teaching-learning process occurs.

In that sense, as underlined by Levia and Quiring (2008), "adopting an appropriate method for the assessment of student learning is critically important because it impacts on how students learn and it helps to motivate students to focus on those knowledge and skill objectives that are deemed to be the most important".

Although research on assessment has intensified, the concepts utilized or the effects generated by assessment are not always consensual. It is suggested "facing assessment as a learning process, to be able to improve, opens other perspectives" (Le Boterf, 2005, p. 69). However, there are always warnings about the risk that may arise from the fact "the practice of assessment has moved from assessment of learning, through assessment for learning, to assessment as learning, with assessment procedures and practices coming completely to dominate the learning experience and 'criteria compliance' replacing 'learning'" (Torrance, 2007, p. 281).

This shift from a testing culture to an assessment culture is marked by a "trend for greater transparency and explicitness in all aspects of student assessment" (Elander, Harrington, Norton, Robinson & Reddy, 2006, p. 71), which may constitute an important advancement in this approach "of" and "for" learning. Though, as stated above, within this perspective it should be also recognized that transparency encourages instrumentalism (Torrance, 2007). However, "internationally over the past two decades, higher education institutions and educators have become increasingly committed to making assessment and grading more effective in promoting student learning (that is, in fulfilling a significant formative function) and to making less mysterious, more open and more explicit the grounds upon which student productions are graded" (Sadler, 2005, p. 193).

In a theoretical context there seems to be a transition from an assessment "of" learning to an assessment "for" learning. It is also possible to recognize that, although there is a clear conceptual distinction between summative assessment and formative assessment, some confusion remains regarding the procedures that can be included in the formative assessment.

Regarding *summative assessment*, most of the literature generally defines it as that which is used to assign grades and classifications and also to determine the student's academic development at the end of a specific time period. In this case, the learning assessment (Stiggins, 2002) would measure the mastery of a particular set of norms or contents (Perie, Marion, Gong, & Wurtzel, 2007), or verify the level achieved by students in relation to the learning objectives (Kealey, 2010).

In relation to *formative assessment*, several authors have examined this concept in order to contradict the exclusive use of summative assessment procedures. For example, at the end of the 1990s, Black (1999) showed some of the difficulties that characterize the traditional assessment practiced in higher education and pointed to the importance of self-assessment to allow students to assume responsibility for their own learning. This position defends the idea that a theory of formative assessment has been developed to establish a unifying basis for the diverse practices in this area (Black & William, 2009). However, in this field, according to Pinchok and Brandt (2009), the literature offers multiple (sometimes conflicting) definitions and it may be referred as a process to adjust teaching based on feedback about student performance (Popham, 2006) or as a set of tools to monitor student progress throughout learning (Stiggins, 2002; Dunn & Mulvenon, 2009). According to its objective, any assessment activity or tool can be classified as formative as long as it allows for informed and adapted teaching (Black & Wiliam, 1998; Perie et al., 2007).

Regarding these two ideas, there has been a debate for several years about whether assessment should be primarily summative, reporting on what has been achieved, or

formative, assisting in the learning process by providing information about the strengths and weaknesses of students. However, "assessment always had to contend with these competing purposes and probably always will" (Torrance, 1993, p. 333). In this sense, Sadler (1989, p. 120) argues "the primary difference between formative and summative assessment relates to purpose and effect, not to timing". We adopt this difference, between these two types of assessment as basic concepts in the design of the empirical study and for the data interpretation.

A deeper theory of formative assessment arose in the 80s, with francophone origin (evaluation formatrice), designated in this paper as educative assessment. First proposed by G. Scallon and later taken by other authors such as Bonniol (1986), Vial (1987) and Nunziati (1990), this idea was presented with distinctive formative characteristics, making the assessment process more emancipatory, which focuses on the view of the 'student's ensured regulation' (Abrecht, 1994, p. 49). Therefore, educative assessment "implies [the students'] conscious, systematic and reflected engagement in planning, organizing and evaluating their own learning" (Abrecht, 1994, p. 49). In this perspective, "the assessment methods express conditions of effective responsibility and emancipation"; thus "self-assessment practices are organized and implemented in situations of high student involvement" that are "defined by the teachers and students" (Leite & Fernandes, 2002, p. 65). According to this approach, it is important to highlight that educative assessment presupposes the creation of circumstances in which the student co-participates in the definition of the assessment methods and in negotiation, with the teacher, regarding the configuration of the assessment tools.

Hence, in logic of assessment "for" (and not "of") learning, there is also the concept of self-assessment practices. It is associated with "ideas of democratic education, with a formative and interactive orientation in which the students, themselves, are involved as authors of their own assessment" (Leite & Fernandes, 2002, p. 66).

In order to contribute to a formative assessment theory, Fernandes (2006) established a division between perspectives that derive from the francophone research traditions – where regulation of the learning process is the key concept (the author suggests: Bonniol, 1984; Cardinet, 1991; Gregoire, 1996; Perrenoud, 1998) – and the anglo-saxon research – where feedback has a primary role (the author suggests Black & William, 1998, 2006; Gipps, 1994, 1999; Gipps & Stobart, 2003; Shepard, 2001; Stiggies, 2004). Faced with this duality, Fernandes (2006) considered the "alternative assessment" as a process that is participative, transparent, integrated in teaching and learning, and aimed at regulating and improving. It is a procedure that focuses mainly on processes, without ignoring the products.

In this perspective, designated as *alternative formative assessment*, the author highlights the improvement and regulation of learning and teaching processes (formative assessment) and, simultaneously, pointing out an alternative to formative assessment inspired within a behaviorist approach (alternative assessment) (Fernandes, 2006). This point of view is different from the traditional concept of formative assessment.

In a behaviorist perspective, formative assessment involves short-term goals, clear assessment objectives, and detailed feedback to students on what they have or have not achieved and what they must improve. On the contrary, in the social constructivist perspective, formative assessment (and what we are calling alternative assessment) includes a role for the teacher in assisting the student to comprehend and be engaged with new ideas and problems (Torrance, 1993). This perspective looks forward rather than backwards and implies "that the teacher/tester and student collaborate actively to produce a best performance" (Wood, 1987, p. 242). As stated by Torrance (1993, p. 336), "such an

interpretation of formative assessment takes us into a much more dynamic and challenging territory, and has the merit of identifying an integrated role for assessment in the process of teaching and learning itself, rather than at one remove from it, as a more behaviorist or graded test approach would seem to imply".

In summary, the concept of alternative formative assessment implies an integrated resource of the two perspectives that, as noted, emphasized "regulation" and "feedback". This position seems to go in the direction of what Sadler (1989, p. 122) argues when he states "formative assessment includes both feedback and self-monitoring. The goal of many instructional systems is to facilitate the transition from feedback to self-monitoring".

The position that guides us combines these last two learning assessment concepts. We assume that conditions should be created so that procedures evolve from a formative assessment initially with more feedback to a formative and educative assessment, more focused on self-regulation of learning, making students progressively responsible for the evolution of their training programs. Following this perspective, we support Cambell, Künnemeyer, and Prinsep's (2008, p. 289) idea when they state that "teaching staff aimed to help students become progressively more independent in their learning (…) through a decrease in the amount of support and guidance offered as students progressed through their studies". At the same time, we also agree that the "assessment has most effect when (…) students progressively take responsibility for assessment and feedback processes" (Boud & Associates, 2010, p. 2). What we argue is that as the learning develops it becomes possible to shift from an external assessment, with a greater role give to the teacher (through assessment feedback), to a more internal assessment (students' self-assessment). This internal assessment requires great responsibility and self-control from the student and facilitates their appropriation of the learning assessment devices, promoting a particular transition from regulation logic to emancipation logic. In sum, we adhere to the idea of combining an "alternative formative assessment" that assigns to the teacher a catalyst action over the teaching-learning-assessment process, with an "educative assessment" that allows the student a more emancipatory action. Having these concepts as a base, we conducted a study that tried to understand the pedagogical work models and the learning assessment most used in higher education in the transition context of the Bologna Process.

The method

In order to collect data about the types of assessment in use in higher education at the University of Porto, we developed a survey based on a questionnaire that sought to identify teacher and student perceptions regarding assessment. This questionnaire, validated by experts, contained 16 items, 4 for each assessment concept under review: summative assessment (SA), formative assessment (FA), formative and educative assessment (FEA) and alternative formative and educative assessment (AFEA). A 7-point Likert scale was utilized for responses, with 1 representing "Never", 4 "Sometimes" and 7 "Always".

The sample was composed of teachers and students from two disciplines from the 1st year and two others from 4th year of two Integrated Master's Programs: Informatics and Computing Engineering (MIEIC) and Psychology (MIP), aiming to determine the influence of these variables in the assessment processes. These two Integrated Master's Programs were selected for our study because they represent two distinct areas of knowledge (technological and social sciences).

The eight disciplines were also selected according to the following criteria: number of students (many and few) and average grades (high and low). The first criterion (number

of students) was chosen to allow verification of whether this condition has a notable role, facilitating or constraining, the implementation of assessment procedures. Moreover, the second criterion (average grades) was selected to enable an analysis to confirm if these results are associated with a major resource of formative assessment that, theoretically, is conceived to favor and support the learning process, intending to foster better learning outcomes.

A questionnaire (online and in person) was given to 21 teachers and 253 students and 274 surveys obtained. This questionnaire was created and developed by an extended analysis of the already existing and validated questionnaires in this area of study and, moreover, was also validated by experts. Furthermore, the validation of this questionnaire also included a previous application to some teachers and students to verify the clarity of the items and the medium time of response. Likewise, it was not applied a pretest because through the Statistical Package for Social Sciences resource we could realize an exhaustive statistical analysis to ensure the psychometric qualities of the items, allowing us to exclude items less stringent. Thus, the data was corrected in relation to the existence of excess missing values and for "uni and multi varied" outliers. In the end, we obtained a convenience sample of 261 respondents, appropriate for the population studied; 100% of teachers and 20% or more of students for all selected disciplines. In the end, we obtained a convenience sample of 261 respondents, appropriate for the population studied; 100% of teachers and 20% or more of students for all selected disciplines.

The items were subjected to a Principal Components Analysis with a Varimax orthogonal rotation. All items presented communality values over .40. The analysis extracted 3 components with Eigenvalues above 1, with a Kaiser-Meyer-Olkin value of sampling adequacy of .88. The identified components were:

- 1st component (35.04% of variance) is made up of items such as:
- The assessment methods respect the different learning styles and rhythms, allowing students to choose assessment periods.
- The assessment is co-participated, integrated in the teaching-learning process, aimed at improving learning and focused on the process without ignoring the content matter.

These items are part of the proposed dimensions *a priori* FA-FEA-AFEA (Cronbach's alpha = .863).

- 2nd component (17.01%) consists of items including:
- The content and interactive resources available encourage self-assessment.
- The assessment procedures create conditions for teachers to improve their teaching methods and students to self-regulate their learning.

These items are part of the proposed dimensions *a priori* FA-FEA-AFEA (Cronbach's alpha = .799).

- 3rd component (9.38%) is constituted of items such as:
- The grades are assigned based only on tests or exams after completing the content program or the academic term (semester).
- Knowledge is checked quantitatively and assessed just by the teacher.

These items are part of the proposed dimension *a priori* SA (Cronbach's alpha = .665).

The first component was designated as formative assessment I (FA-I), the second as formative II (FA-II) and the third as summative assessment (SA); with the first two correlated. This option proved to be adjusted since the obtained correlations were the following: FA-I*FA-II (r = 0.628, *p* = .000); SA*FA-II (r = -0.225, *p* = .000); SA*FA-I (r = -0.049, *p* = .428).

The fact that respondents aggregated items of three types of non summative assessment proposed *a priori* within two different components may prove difficult in distinguishing those assessment modalities (FA–FEA–AFEA). This could also be a problem because formative assessment is included in those three types of assessment and, therefore, constitutes a kind of element/concept that is common or aggregator of the representations of the respondents.

The differences in obtained results between FA-I and FA-II, comparing and analyzing their items, can be explained by the fact that half of the items of FA-I are focused on AFEA, which focuses more on this type of assessment. Indeed, the content of those items reflects a more emancipatory perspective as can be verified by the mentioned examples.

In conclusion, although the items of these two components assess the same dimensions, they do not analyse it in the same way or with the same depth, making it clear that the complementarity of different components (I and II) of FA may contribute to the clarification of the transition situation experienced by teachers and students in terms of these assessment procedures.

The results

Multivariate analyses were performed, initially having compared the previously identified components (FA–FEA–AFEA), to obtain a global characterization of the students' and teachers' perceptions. A repeated measures variance analysis revealed a principal effect for assessment ($F(2, 520) = 24.412$, $p< .001$, partial $\eta^2 = .086$). The *post-hoc* analysis with Bonferroni correction revealed significantly higher values for the dimension FA-II ($M = 4.44$; $SD = 1.24$), followed by the SA ($M = 4.06$; $SD = 1.56$) and FA-I ($M = 3.63$; $SD = 1.34$)

dimensions.
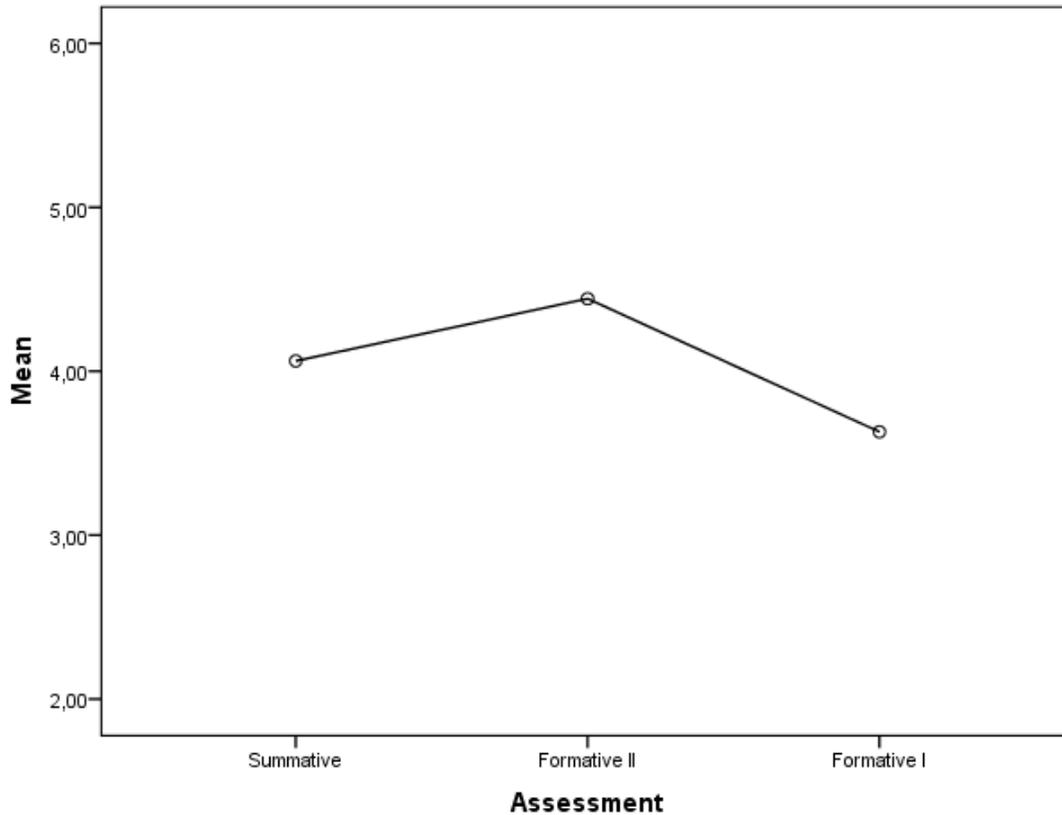
Figure 1: General perspective of the assessment results.

Comparison per Year: there was a main effect of year (F (2, 259) = 64.058, *p*< .001,partial η^2 = .198) and an interaction year*assessment[3] (F(2, 259) = 9.363, *p* = .001, partial η^2 = .035).

     Decomposing the interaction through variance analyses for both groups as well as through *post-hoc* comparisons with correction for multiple comparisons, we verified that although the main assessment effect is maintained in both groups (F(2, 332) = 23.445, *p* = .001 for the 1st year and F(2, 186) = 9.661, *p* = .001 for the 4th year) whilst in the 1st year, the values are significantly lower for the dimension FA-I in comparison to the dimensions FA-II and SA, that do not differ, for the 4th year the values are higher for FA-II, not differing for the dimensions FA-I and SA.
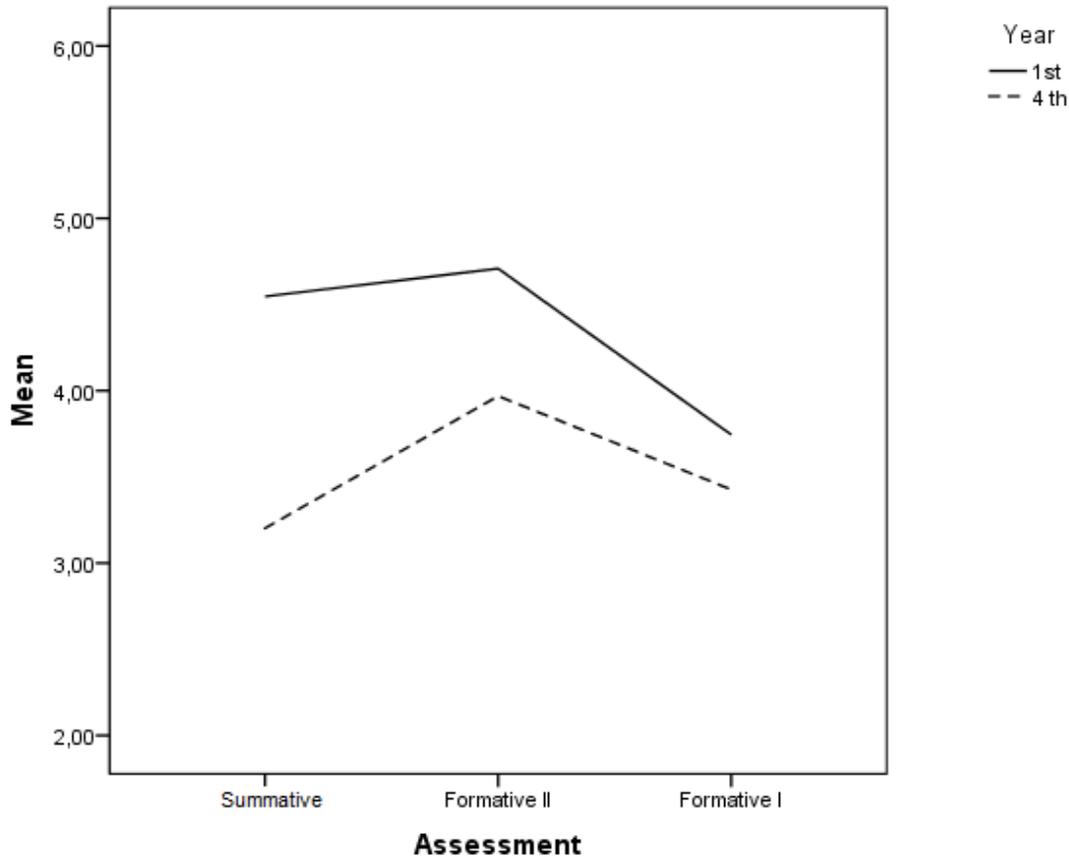
Figure 2: Comparison of the assessments by year of degree programs.

Comparison by Degree Programs: onlythe main effect already described for this variable is evident; that is, the values for the dimensions of assessment do not differ within the two degree programs (in other words between the two Integrated Master's Programs – MIEIC/MIP).

Teachers vs. Students: we verified the described effect of assessment and the interaction between respondent and assessment (F(2, 518) = 9.076, p = .001, partial η^2 = .034).

Decomposing the interaction, we confirmed that the assessment effect remains significant in both groups (teachers F(2, 40) = 11.002, p = .002; students F(2, 478) = 22.593, $p<$ .001). However, while the ASfor teachersis significantly lower than others, for students it appears that theAF-I is significantly lower, not differing in the remaining dimensions.
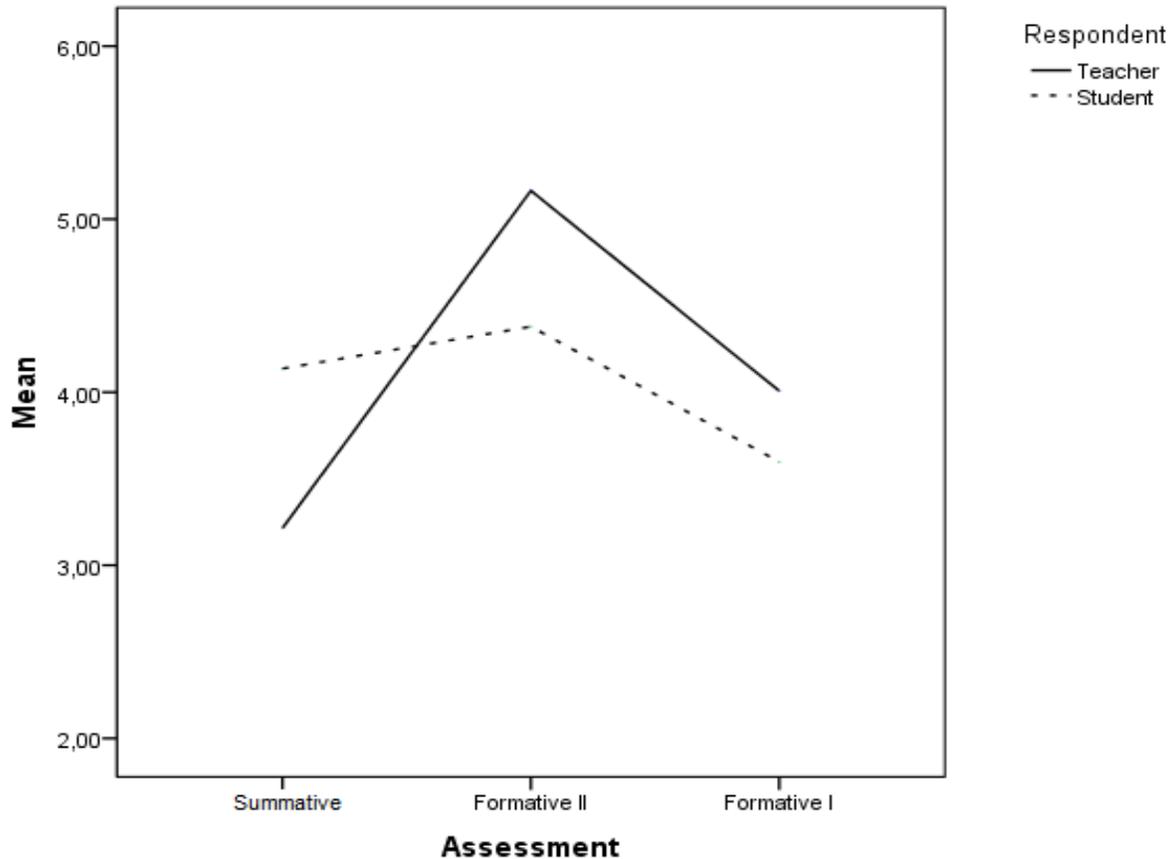
Figure 3: Comparison between teachers and students on assessment.

Number of students and Average grades: through the variance analysis with a number of students (many vs. few) and average grades (high vs. low) considered as an inter-subjects factor, we confirmed the effect of assessment (already described), an effect of the number of students ($F(1, 257) = 6.486$, $p = .011$, partial $\eta^2 = .025$), an effect of the average grades ($F(1, 257) = 4.427$, $p = .036$, partial $\eta^2 = .017$), an interaction between a number of students* average grades ($F(1, 257) = 45.008$, $p < .001$, partial $\eta^2 = .149$), an interaction between assessment*number of students ($F(2, 514) = 61.252$, $p < .001$, partial $\eta^2 = .192$) and also an interaction between assessment*number of student* average grades ($F(2, 514) = 5.365$, $p = .010$, partial $\eta^2 = .020$).

With the purpose of decomposing the second-order interaction, the disciplines with high and low numbers of students were analyzed separately.

In the disciplines with few students, there is a main effect of assessment ($F(2, 206) = 50.787$, $p < .001$), and an effect of average grades ($F(1, 103) = 29.462$, $p < .001$), but without an interaction between assessment* average grades. The dimension AF-II (M = 4.68, SD = 1.25) has significantly higher values, followed by the dimensions AF-I, (M = 4.07, SD = 1.26) and AS (M = 3.06, SD = 1.44).

In the disciplines with many students, there was an assessment effect, ($F(2, 308) = 33.207$, $p < .001$), an average grades effect ($F(1, 154) = 14.143$, $p < .001$) and an interaction between assessment*average grades ($F(2, 308) = 5.641$, $p = .008$).

The analysis of this interaction reveals that the respondents of disciplines with high and low grades do not differ in the dimension FA-I ($t(154) = 0.020$, ns), but the respondents of disciplines with low grades confer higher values to the dimension FA-II ($M = 4.44$, $SD = 1.21$) and to the dimension SA ($M = 5.01$, $SD = 1.24$) than the respondents of disciplines with high grades (FA-II $M = 3.87$, $SD = 1.17$; SA $M = 4.02$, $SD = 0.95$) (FA-II $t(154) = 2.657$, $p = .009$; SA $t(154) = 4.725$, $p < .001$).

Number of students = few
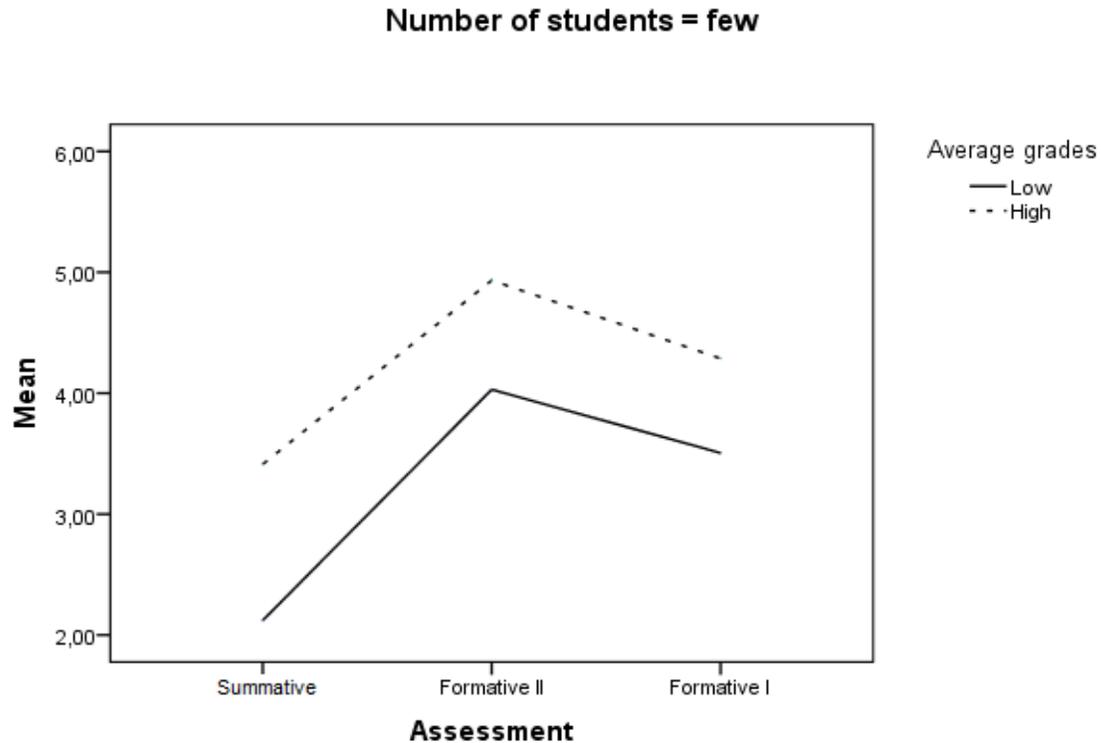
Figure 4: Comparison between high and low grades in disciplines with few students.

Number of students = many
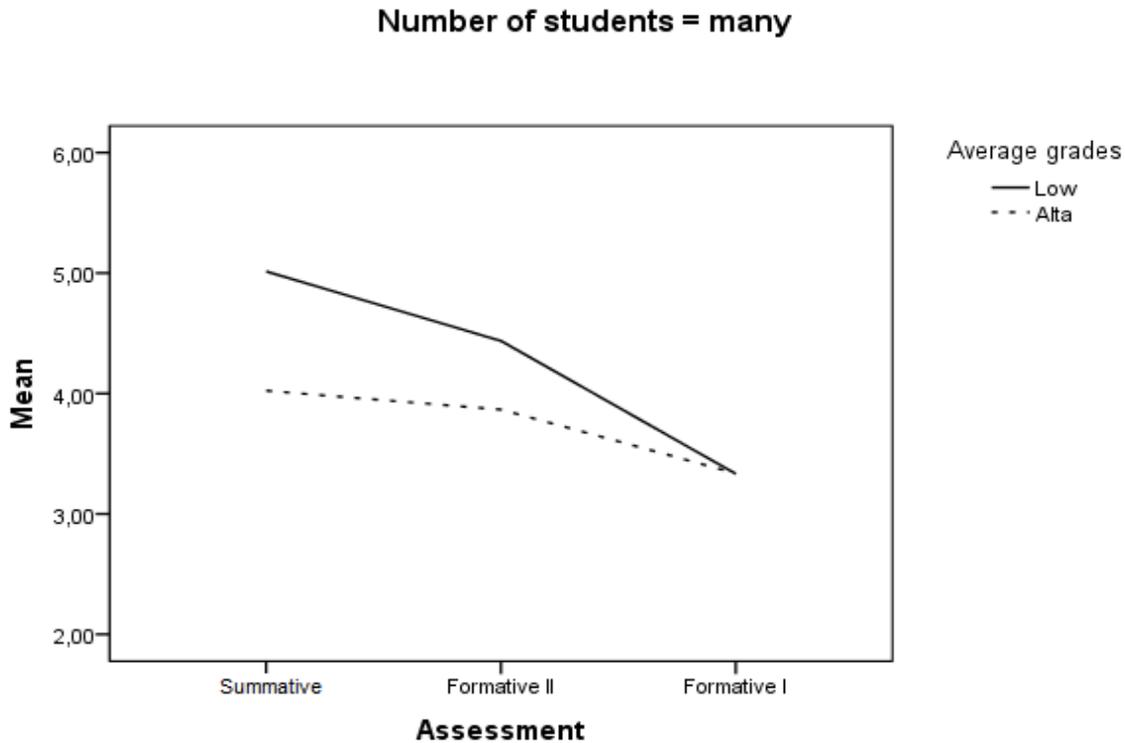
Figure 5: Comparison between high and low grades in disciplines with many students.

Discussion of results

Globally, assessment practices experienced are perceived by teachers and students as the following type (in decreasing order): 1st) Formative Assessment-II; 2nd) Summative Assessment; 3rd) Formative Assessment-I. This data most likely suggests a transition phase, not very distinct nor clear, in which different assessment methods co-exist, and the summative modality does not constitute the only or, in some cases, the main resource used.

The observed difference among formative assessment components that are more or less used may contribute to identifying a change that seems to be happening at distinct rates in relation to different procedures of formative assessment that are more or less emancipatory. This may also constitute an indicator of the multi-dimensionality of this concept, which includes formative assessment (FA), formative and educative assessment (FEA) and alternative formative and educative assessment (AFEA).

However, comparative analysis by class year revealed that in the 1st year (in which pedagogical work is focused on more introductory skills for higher education) SA and FA-II co-exist with the same intensity with FA-I less used. On the other hand, in the 4th year (in which pedagogical work is focused on more professional skills) FA-II is clearly the most used. In synthesis, at this level, the greater or lesser maturity of the students and the type of skills, initial or advanced seem to be aspects to consider relative to the assessment procedures adopted.

Regarding the perceptions of students and teachers while for teachers, SA is the least used and instead there is greater existence of procedures characteristic of FA-I and especially FA-II, for students SA it is the most used (along with FA-II) and FA-I the least used. Given

this difference in perceptions, it may be admissible to consider a possible influence of teachers' higher expectations for change, in relation to an area that traditionally has been their responsibility. However, it is necessary to highlight that desired effects do not seem to correspond to what is experienced by students. In this sense, for example, Goos, Gannaway, and Hughes (2011) conclude that there are differences in the perception about what teachers and students consider as relevant issues in assessment and how they should be approached. This author even states that given the apparent tensions between the assessment expectations of university students, their teachers, and senior academic staff responsible for quality assurance in teaching and learning, there is a need to investigate perceptions of the role of assessment held by each of these stakeholder groups. (Goos et al., 2011, p. 96)

Finally, the comparative analysis between disciplines with different characteristics (high or low number of students and high and low grades) allowed us to identify two very distinct situations:

- With low number of students, the type of assessment that is most clearly used is the formative (more FA-II than FA-I), with SA always recognized as the perspective with less applied procedures.
- Contrasting the former case, higher number of students results in major employment of SA, followed by FA-II (with a noticeable effect in the disciplines with low grades) and finally FA-I.

It is worth highlight that this is the only situation in which SA appears as the most used assessment procedure, which leads to the consideration that the existence of a high number of students seems to constitute a limiting condition in executing procedures inherent to FA (especially FA-I, more emancipatory). Due to this difficulty, the assessment strategies focus mainly on methods that aim to control the learning outcomes (and an effect of selection/grading of students), rather than an assessment of the processes used to generate those results and utilization of assessment to improve the teaching-learning situation. Much research has been conducted on class size and its effects on learning (Lindsay & Paton-Salzberg, 1987; Mahler, Neumann, &Tamir, 1986; Raimondo, Esposito, &Gershenberg, 1990; McKeachie, 1999). Large classes are often seen as a negative feature of modern day higher education, and there is some evidence that they can reduce students' learning outcomes (Raimondo et al., 1990; McKeachie, 1999). Biggs (2003) acknowledges that conditions such as class size and student diversity make good teaching more difficult than ever. Campbell et al. (2008) found that class sizes were identified as having considerable impact on student learning and lectures offer little opportunity for active student engagement. Moulding (2010, p. 151) shows that "research into the effects of large classes demonstrates that students are disadvantaged in terms of higher order learning because interactions between teachers and students occur at lower cognitive levels". Given these perspectives arising from the literature about disciplines with many students and its effects on learning, the data presented in this paper seems to emphasize the need to study also the effects on assessment.

Linking these perspectives with average grades, we verify that there is the same tendency. In this sense, within low number of students, FA-II is the assessment type more utilized, independently of the average grades, but with more intensity in high average grades. Further, summative assessment is mostly used in situations with a high number of students and low average grades, which can indicate that there is a need for formative assessment to improve better learning outcomes.

In summary relative to the transition context that takes place in higher education, it is worth noting that, in almost all analysis modalities, the results prove that the assessment

procedures utilized are already perceived as more of the FA-II type, although with different intensities. However, SA is still perceived as predominant in disciplines with many students and has a similar use to the AF-II both within the 1st year and in the perspective of students.

As for the assessment methods least used, SA was identified from the perspective of the teachers and in the disciplines with few students as the least used in the 4th year. In these situations there seems to exist facilitating factors or/and a more favorable vision in relation to more formative assessment.

With the exception of the former cases, the fact that FA-I is always presented as the less used form of assessment seems to show the perception of constraint factors or/and a less favorable view in relation to the possibility of adopting FA procedures with stronger emancipatory characteristics.

Conclusions

According to the obtained data, it is possible to claim that we are still faced with a certain continuity of the dominant pedagogical culture in which the teacher assesses with the objective of grading. Hence, even when the assessment type being utilized is summative, we can observe the co-existence of formative components of assessment in the disciplines that were studied. However, the use of more emancipatory methods of assessment does not become apparent.

In this situation, which seems to be natural and typical of a transition context in higher education under the Bologna Process, there seems to be some difficulty in enacting some of the formative assessment methods, mainly the ones with more emancipatory character, and particularly when the number of students is too high. Although, considering the transnationality of the Bologna Process it should be noted that this transition could be emerging in most of the 47 countries, what would indicates that these study results might be similar in other nations and/or universities. This assertion justifies the necessity of conducting this type of studies in a transnational context, to better identify the impact of this transition in the assessment procedures and, furthermore, to realize if these modifications are the reflection of an emerging transnational curriculum. Indeed, it seems to be of major importance to question if the transnational change implicated by the Bologna Process brought more homogeneity or, by the contrary, reinforce the existing heterogeneity between assessment practices in different countries and higher education institutions. On the other hand, this study allows us to conclude that the type of degree program does not interfere significantly in the respondents' perceptions because the values for the dimensions relative to assessment do not differ according to this variable. This fact lends itself for the hypothesis that the nature of the degree programs and the specificities that may exist in each of these Faculties at the University of Porto do not constitute relevant aspects for the type of assessment used. Furthermore, it seems possible to admit that in a curricular perspective, a more emancipatory formative assessment could contribute for the development of the students' autonomy, transversely to the different areas of study, almost as if it was a component of the "hidden curriculum" within the universities' degree programs.

To conclude, we can infer that the evolution of the current situation, which aims to articulate summative and formative perspectives in order to provide a more integrated assessment "with" learning and more productive "for" learning depends, among other things, on the adoption of new assessment procedures by teachers as well as the presence of conditions that benefit the use of self-assessment modalities by students.

In agreement with Rieg and Wilson (2009), we argue "given the link between instruction and assessment, it can be assumed that these same teachers lack knowledge of assessment strategies as well as instructional pedagogy" (Rieg& Wilson, 2009, p. 281) and found that "sadly as the researchers found in this study, even faculty members who are aware of effective instructional and assessment techniques are not always using them in their university classrooms" (Rieg & Wilson, 2009, p. 292).

Finally, this study provides relevant contributions in order to deepen the awareness of the types of learning assessments that are in practice within the Bologna Process. However, it will remain unclear what kinds of links between summative and formative assessment may help to foster better learning conditions in higher education, which justifies future research that also considers the specificities of more and less emancipatory components that this study identified at the formative assessment level.

Notes

---

[1] manuel@firmino.pt

[2] carlinda@fpce.up.pt

[3] In this presentation of statistical analysis results, we use the symbol (*) as "and" to designate an interaction between two variables.

References

Abrecht, R. (1994). *A avaliação formativa*. Porto: Edições ASA.

Altbach, P., & Engberg, D. (2006). *Higher education: A worldwide inventory of centers and programs*. Massachusetts: Boston College.

Biggs, J. (2003). *Teaching for quality learning at university*. Buckingham: SRHE and the Open University Press.

Black, P. (1999). Assessment, learning theories and testing systems. In P. Murphy (Ed.), *Learners, learning and assessment* (pp. 118-134). London: Paul Chapman Publishing.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan International,80*(2), 139–144.

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability,21*(1), 5-31.

Bonniol, J-J. (1986). Rechercheset formations: Pour une problématique de l'evaluation formative. In J.M. De Ketele *(*Ed.), *L'évaluation: Approche descriptive ou prescriptive?* (pp. 119-133). Bruxelles: De Boeck.

Boud, D., & Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education.* Sydney: Australian Learning and Teaching Council.

Boud, D., & Falchikov, N. (2007). Introduction: Assessment for the longer term. In D. Boud & N. Falchikov (Eds.). *Rethinking assessment in higher education: Learning for the longer term* (pp. 3-13). New York: Routledge.

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.

Brown, S. (2004). Assessment for learning.*Learning and Teaching in Higher Education,1*, 81-89.

Campbell, A., Künnemeyer, R., & Prinsep, M. R. (2008). Staff perceptions of higher education science and engineering learning communities. *Research in Science & Technological Education,* 26(3), 279-294.

CEDEFOP (2009). *The shift to learning outcomes: Policies and practices in Europe*. Luxemburg: Office for Official Publications of the European Communities.

Coates, H. (2005). The value of student engagement for higher education quality assurance.*Quality in Higher Education,11* (1), 25–36.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation,14*(7), 1–11.

Elander, J., Harrington, K., Norton, L., Robinson, H., & Reddy, P. (2006). Complex skills and academic writing: A review of evidence about the types of learning required to meet core assessment criteria. *Assessment & Evaluation in Higher Education,* 31(1), 71–90.

Fernandes, D. (2006). Para uma teoria da avaliação formativa. *Revista Portuguesa de Educação,19*(2), 21-50.

George, J., Cowan, J., Hewitt, L., & Cannell, P. (2004). Failure dances to the tune of insecurity: Affective issues in the assessment and evaluation of access learning. *Journal of Access Policy and Practice,1*(2), 119-133.

Gibbs, G. (1992). *Improving the quality of student learning*. Bristol: TES.

Goos, M., Gannaway, D., & Hughes, C. (2011). Assessment as an equity issue in higher education: Comparing the perceptions of first year students, course coordinators, and academic leaders. *Australian Educational Researcher,38* (1), 95-107.

Kealey, E. (2010). Assessment and evaluation in social work education: Formative and summative approaches. *Journal of Teaching in Social Work, 30*(1), 64-74.

Le Boterf, G. (2005). *Construir as competências individuais e colectivas*. Porto: ASA Editores.

Leite, C., & Fernandes, P. (2002). Avaliação das aprendizagens dos alunos: Novos contextos, novas práticas. Porto: Edições Asa.

Levia Jr, D., & Quiring, S. (2008). Assessment of student learning in a hybrid PBL capstone seminar. *Journal of Geography in Higher Education,32*(2), 217-231.

Lindsay, R., & Paton-Salzberg, R. (1987). Resource changes and academic performance at an english polytechnic. *Studies in Higher Education,12*(2), 213–227.

Mahler, S., Neumann, L., & Tamir, P. (1986).The class-size effect upon activity and cognitive dimensions of lessons in higher education.*Assessment & Evaluation in Higher Education,11*(1), 43–59.

McKeachie, W. (1999).*Teaching tips*. Boston: Houghton Mifflin.

Moulding, N. T. (2010). Intelligent design: Student perceptions of teaching and learning in large social work classes. *Higher Education Research & Development,29*(2), 151-165.

Nunziati, G. (1990). Pour construire un dispositive d'évaluation formatrice. *Cahiers Pédagogiques: Apprendre*, *280*, 48-64.

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). *The role of interim assessments in a comprehensive assessment system.* Washington, DC: The Aspen Institute.

Pinchok, N., & Brandt, W. C. (2009).*Connecting formative assessment research to practice: An introductory guide for educators*. Learning Point Associates.

Popham, W. J. (2006). Phony formative assessments: Buyer beware! *Educational Leadership,* 64(3), 86–87.

Prosser, M., Ramsden, P., Trigwell, K., & Martin, E. (2003).Dissonance in experience of teaching and its relation to the quality of student learning.*Studies in Higher Education,28*(1), 37–48.

Raimondo, H., Esposito, L., & Gershenberg, I. (1990).Introductory class size and student performance in intermediate theory courses.*Journal of Economic Education,21*(4), 369–381.

Ramsden P. (2003). *Learning to teach in higher education*. London: Routledge.

Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.

Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education,* 30(4), 387–415.

Rieg, S. A., & Wilson, B. A. (2009).An investigation of the instructional pedagogy and assessment strategies used by teacher educators in two universities within a state system of higher education.*Education,130*(2), 277-294.

Sadler, D. R. (1989).Formative assessment and the design of instructional systems.*Instructional Science,18*, 119-144.

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education,* 30(2), 175-194.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan,83*(10), 758–765.

Torrance, H. (1993). Formative assessment: Some theoretical problems and empirical questions. *Cambridge Journal of Education,23*(3), 333-343.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education,14*(3), 281–294.

Vial, M. (1987). *Un dispositive d'évaluation formatrice en expression écrite au college*. Marseille: CRDP.

Wood, R. (1987). *Measurement and assessment in education and psychology*. London: Falmer Press.