Recherche d'une définition équitable de l'évaluation située : le cas de l'université d'Austin au Texas.

Claire Tardieu[*1]

[1]Langues, Textes, Arts et Cultures du Monde Anglophone (PRISMES) – Université Paris III - Sorbonne nouvelle – UFR du Monde Anglophone, 5 Rue de l'Ecole de Médecine, 75006 Paris, France

Résumé

This paper focuses on situated assessment from the point of view of both research and teaching. Are there any characteristics that are ethically preferable to others and, if so, which ones? What equitable definition of situated assessment, respectful of the issues, the contexts and, above all, the evaluated subjects, can we propose for higher education?

We will first examine the macro level of research in evaluation, and then the meso level of an institutional context, in this case, that of American higher education, illustrated by the University of Texas at Austin (UTA). Theoretically, we will refer to the existing literature on language assessment (Bachman, 1990, 2007, Alderson 2002, Spolsky, Hadji, 1997, Horner, 2010, Huver & Springer, 2011, Narcy-Combes, 2009, Tardieu 2009, 2014, and Gardner, 1997), with a special focus on assessment at the university in relation with grading practices (docimology). If we consider higher education settings, we will note that even though there may be recommendations and clear policies regarding the number of papers to be handed in by the students, and precise requirements in relation to grading, the content of the exams, the questions or exercises proposed by the teachers are not usually supervised. We will address the issue of reliability and validity in the context of the University of Texas at Austin (UTA). Are there clear general policies consistent throughout a department? Does the number of assignments vary from one course to another? What types of assignments are there? What are the requirements and how is the grading carried out accordingly? As for the validity issue, we should be reminded of Romainville (2014) who identified three different approaches and demonstrated that more often than not the students have to "guess" what is expected from them to get good grades. Although the university exams cannot claim to be valid or reliable like proper tests, they play such an important role in the academic life and decide on the success or the failure of the students that we should raise those questions. Are the goals clearly stated to the students? Do they have the opportunity to get more information from their teachers? Or even to revise their papers in order to improve their final grades?

This paper will more precisely address these issues regarding the organization of exams in terms of number of papers, requirements, and grading in the context of the English Department of the University of Texas at Austin for undergraduate students in the second semester of the academic year 2018-2019.

Methodologically, we will first analyze the online descriptions of all the English courses of the semester. This generic analysis will highlight the main features of assessment in the context of this university and enable us to answer our initial questions on reliability and fairness. Then,

---

[*]Intervenant

we will analyze more data from seven interviews of teachers who explain how they proceed to evaluate and grade their students with a special interest for some outstanding features. This second type of data analysis will permit us to answer the question of validity of the construct.

Références

Alderson, J.C. (2002). Testing proficiency and achievement: principles and practice. In, J. A. Coleman R. Grotjahn, and U. Raatz, (eds) *University language testing and the C-Test.* Bochum: AKS Verlag, p. 15-30.

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing.* Oxford: Oxford University Press.

Bachman, L.F. (2007). What is the construct? The dialectic of abilities and contexts in Defining constructs in language assessment. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C.E. Turner & C. Doe (eds.), *Language testing reconsidered,* pp. 41-72. Ottawa, CA: University of Ottawa Press.

Baudrit, A. (2007). *L'apprentissage collaboratif. Plus qu'une méthode collective ?*De Boeck Supérieur, Collection : Pédagogies en développement.

Ellis, R. (2003). *Task-based Language Learning and Teaching.* Oxford: Oxford University Press.

Hadji, C. (1997). *L'évaluation démystifiée.* Paris : ESF, Collection : Pratiques & enjeux pédagogiques.

Horner, D. (2010). *Le CECRL et l'évaluation de l'oral en anglais.* Paris : Belin, Collection : Guides de l'enseignement.

Narcy-Combes, J.-P. (2009). " La correction dans l'enseignement/apprentissage des langues : un problème malaisé à construire ". In *La correction dans l'enseignement des langues spécialisées, Les Cahiers de l'APLIUT, Pédagogie et Recherche,* Vol. 28, no3, pp. 26-38.

Nunan, D. (1989). *Designing Tasks for Communicative Classroom.* Cambridge: Cambridge University Press.

Romainville, M. 2014. " Motiver les étudiants à apprendre, plus qu'à étudier et à réussir ". Conférence SAPIENS, Université Paris 7 Denis Diderot, le 18 juin 2014.

Spolsky, B. (1981). Some ethical questions about language testing. In C. Klein-Braley & D. K. Stevenson (eds)*Practice and problems in language testing,* pp.5-30. Frankfurt am Main: Verlag Peter D. Lang.

Springer, C., Huver, E. (2011). *L'évaluation en Langues.* Paris : Didier.

Tardieu, C. (2009). " Corriger ou évaluer ". In *La correction dans l'enseignement des langues de spécialité, Les Cahiers de l'APLIUT, Pédagogie et recherche* vol. XXVIII, 3/2009, pp. 9-25. (http://halshs.archives- ouvertes.fr/halshs-00667700).

Tardieu, C. (2014). *Notions-clés de la didactique de l'anglais.* Paris : Presses de la Sorbonne Nouvelle.