

# Evaluer

Journal international de Recherche  
en Education et Formation

# Evaluer

Journal international de recherche en éducation et formation

## Rédacteur en chef

Marc Demeuse – Université de Mons – Belgique ([redacteur@e-jiref.education](mailto:redacteur@e-jiref.education))

## Rédactrice en chef adjointe

Annick Fagnant – Université de Liège – Belgique

## Webmasteur

Pascal Detroz – Université de Liège – Belgique

## Comité de rédaction

Réginald Burton – Université du Luxembourg – Luxembourg

Pierre-François Coen – HEP Fribourg – Suisse

Pascal Detroz – Université de Liège – Belgique

Rémy Goasdoué – Université Paris-Descartes – France

Marc Vantourout – Université Paris-Descartes – France

Nathalie Younes – Université Blaise Pascal – France.

## Relecture

Monique Jehin

## Secrétariat

Anne Sénécal – Université de Mons – Belgique

## Comité scientifique

Véronique Bedin – Université Toulouse Jean Jaurès – France

Alexandre Buysse – Université Laval - Canada

Alain Content – Université Libre de Bruxelles - Belgique

Marcel Crahay – Université de Genève - Suisse

Jean-Marie De Ketele – Université Catholique de Louvain – Belgique

Vincent Dupriez – Université Catholique de Louvain - Belgique

Gérard Figari – Université Pierre Mendès France de Grenoble - France

Anne Jorro – Conservatoire national des arts et métiers - France

Dominique Lafontaine – Université de Liège – Belgique

Romain Martin – Université du Luxembourg - Luxembourg

Pierre Merle – Université de Rennes 2 – France

Sophie Morlaix – Université de Bourgogne – France

Joëlle Morrissette – Université de Montréal – Canada

Lucie Mottier Lopez – Université de Genève – Suisse

Pascal Nidinga – Université du Québec à Montréal - Canada

Xavier Pons – Université Paris Est Créteil Val de Marne - France

Bernard Rey – Université Libre de Bruxelles - Belgique

Thierry Rocher – Ministère de l'Éducation nationale – France

Marc Romainville – Université de Namur – Belgique

Marie-Christine Toczek-Capelle – Université Blaise-Pascal de Clermont-Ferrand - France

# Sommaire

<i>Carmen CAVACO &amp; Christophe DIERENDONCK</i>	
Introduction	
Évaluation(s) et apprentissage(s) : des rapports d'interdépendance.....	5
<i>João BARROSO</i>	
L'évaluation des apprentissages et la régulation des politiques éducatives.....	9
<i>Marie-Christine TOCZEK &amp; Carine SOUCHAL</i>	
Le pouvoir des contextes évaluatifs.....	21
<i>Maria do Céu ROLDÃO</i>	
Évaluation et curriculum : à la recherche d'une nouvelle signification de l'enseignement et de l'apprentissage .....	37
<i>Walther TESSARO, François-Marie GERARD et Marcelo GIGLIO</i>	
Changements curriculaires : un levier pour les pratiques évaluatives des enseignants ?	51
<i>Domingos FERNANDES, Isabel FLALHO et Carlos BARREIRA</i>	
Perceptions et pratiques d'évaluation des apprentissages dans l'enseignement universitaire portugais.....	61
<i>Annick FAGNANT, Richard ETIENNE, Lucie MOTTIER LOPEZ et Marie-Noëlle HINDRYCKX</i>	
L'évaluation comme objet d'apprentissage et comme outil de développement professionnel dans le cadre de la formation des enseignants.....	77
<i>Christophe DIERENDONCK &amp; Carmen CAVACO</i>	
Evaluer : pour contrôler ou pour apprendre ? .....	99



## Introduction

# Évaluation(s) et apprentissage(s) : des rapports d'interdépendance

**Carmen Cavaco**

Universidade de Lisboa  
[carmen@ie.ulisboa.pt](mailto:carmen@ie.ulisboa.pt)

**Christophe Dierendonck**

Université de Luxembourg  
[christophe.dierendonck@uni.lu](mailto:christophe.dierendonck@uni.lu)

**Pour citer cet article** : Cavaco, C. & Dierendonck, Ch. (2017). Introduction. Évaluation(s) et apprentissage(s) : des rapports d'interdépendance. *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 5-8.

### 1. Introduction

L'évaluation *des* apprentissages a toujours occupé une place centrale au sein des politiques éducatives, des pratiques pédagogiques et des travaux scientifiques ayant l'évaluation pour objet d'étude. Mais au-delà des visées centrales de contrôle et de régulation des apprentissages, les pratiques d'évaluation ont également un potentiel formatif substantiel (Scriven, 1967 ; Allal, 1979; Hadji, 1994; Paquay, 2005; Jorro, 2007; Mottier Lopez, 2015) pour l'ensemble des acteurs impliqués (élèves et étudiants, enseignants, établissements d'enseignement, responsables des politiques éducatives). L'évaluation peut certes se caractériser par des dimensions diverses (méthodologique, théorique, épistémologique, déontologique), mais celles-ci s'avèrent souvent interdépendantes, ce qui pose en évidence la complexité de ce domaine.

Selon Ardoino & Berger (1989) « évaluer quelque chose ou quelqu'un c'est élaborer et proposer, à son propos, une appréciation ou de la conformité à des modèles donnés » (p.6). L'évaluation est « une pratique greffée sur une autre pratique, qui se distingue par un dispositif spécifique et un certain rapport aux valeurs » (Vial, 1999, p.1). Ces définitions permettent de dire que l'évaluation ne se réduit pas à sa dimension technique – basée sur des méthodologies, des outils et des référentiels – mais qu'elle comporte toujours une dimension politique incluant à la fois des valeurs qui lui sont indissociables et des zones d'opacité irréductibles (Ardoino & Berger, 1986, p.120). L'évaluation ne peut, par ailleurs, nier la complexité toujours présente dans les pratiques et les phénomènes à évaluer. Elle ne peut enfin se libérer de la présence et de l'influence de la subjectivité, de l'imprévisibilité, de l'interdépendance et de l'implication des acteurs et des contextes concernés.

Tous ces éléments posent avec évidence que les pratiques d'évaluation sont complexes et qu'il convient de considérer les rapports d'interdépendance qui existent entre l'évaluation et l'apprentissage en adoptant une perspective herméneutique qui reconnaît l'importance de la compréhension et de l'interprétation des pratiques, en considérant, d'une part, le sens que leur donnent les acteurs (Habermas, 1987) et, d'autre part, l'influence d'éléments sociohistoriques (Dilthey, 1992) qui conditionnent leur spécificité, leur singularité et leur complexité.

## 2. Structure du numéro spécial

Pour analyser les rapports d'interdépendance entre évaluation et apprentissage, il nous a paru important de mobiliser l'interdisciplinarité des Sciences de l'Éducation en sollicitant une variété de collègues issus de la Sociologie, la Didactique, la Psychologie, la Science Politique, l'Histoire et la Pédagogie. Ce numéro thématique de la Revue e-JIREF rassemble donc six contributions caractérisées par des filiations théoriques différentes mais centrées sur l'évaluation des apprentissages et les apprentissages qui résultent des pratiques d'évaluation. Il s'inscrit dans la foulée du 28<sup>e</sup> colloque international de l'ADMEE-Europe qui s'est tenu du 13 au 15 janvier 2016 à l'Institut de l'Éducation, de l'Université de Lisbonne. Les conférenciers et les participants des différentes tables rondes ont été sollicités pour produire des textes qui rendent compte, voire dépassent le cadre de leurs interventions orales respectives.

A partir d'analyses centrées sur les politiques, les contextes, les curricula, les dispositifs de formation et les acteurs concernés, les contributions examinent trois types d'articulation entre évaluation et apprentissage : l'évaluation *de* l'apprentissage, l'évaluation *au service de* l'apprentissage et l'évaluation *en tant qu'apprentissage*,

Au départ du cas portugais, le texte de **João Barroso**, intitulé « *L'évaluation des apprentissages et la régulation des politiques éducatives* », décrit le phénomène d'intensification et de diversification, pratiquement généralisé, des modalités et des objets de l'évaluation qui conduit à de nouveaux modes de gouvernance des systèmes éducatifs centrés sur les résultats obtenus. Ainsi, la régulation des politiques éducatives s'opère de plus en plus au départ d'évaluations externes ciblant les systèmes, mais également les établissements scolaires, les enseignants et les élèves. Cette analyse soutient l'idée que l'évaluation des apprentissages est devenue une dimension noyau de la construction de la forme scolaire et de la régulation bureaucratique et professionnelle en éducation. L'auteur reconnaît l'importance de l'évaluation de l'apprentissage dans le processus d'enseignement, mais remet en question les objectifs, les modalités et les effets de l'évaluation, à la lumière des changements des deux dernières décennies, observés quant aux modes de gouvernance et de régulation des politiques éducatives.

Dans leur article intitulé « *Le pouvoir des contextes évaluatifs* », **Marie-Christine Toczek** et **Carine Souchal** s'intéressent aux écarts de performance observés entre élèves lors des évaluations à visée de contrôle ou de certification. En évitant de mobiliser les arguments classiques mettant en cause les caractéristiques intrinsèques des élèves, elles convoquent les notions de menace stéréotypique et de buts de réussite pour montrer que les contextes évaluatifs dans lesquels sont placés les élèves exercent une influence significative sur la façon dont l'information est traitée par ces derniers et sur leurs performances aux évaluations. Plus particulièrement, les chercheuses constatent que, dans un climat de réussite axé sur la performance et la comparaison entre élèves, les performances en physique-chimie de lycéennes de seconde sont inférieures à celles de leurs condisciples masculins. Dans un climat

de réussite axé sur la maîtrise des apprentissages, les performances sont les plus élevées tant pour les filles que les garçons, à un niveau très proche. Enfin, dans un contexte où la situation d'évaluation n'est pas annoncée ou formalisée, les garçons sous-performent sensiblement alors que les filles affichent des performances moyennes très légèrement inférieures à la situation de climat de maîtrise.

Deux textes traitent ensuite de l'évaluation *au service de l'apprentissage* au départ d'une analyse examinant les liens entre l'enseignement, l'apprentissage, l'évaluation et les programmes. Le texte de **Maria do Céu Roldão**, intitulé « *Évaluation et curriculum – à la recherche d'une nouvelle signification de l'enseignement et de l'apprentissage* », porte sur la relation entre le champ de l'évaluation et celui du curriculum. Le présupposé de l'analyse est que le développement de la recherche en éducation a conduit à la spécialisation du champ théorique de l'évaluation, ce qui a provoqué une certaine distanciation de l'évaluation, par rapport aux champs d'étude du curriculum et de l'élaboration des curricula. Selon l'auteure, cela a provoqué un déphasage apparent entre le curriculum et l'évaluation en matière de recherche et de pratique scolaire et professionnelle. Le texte vise à donner une nouvelle signification, et, selon une vision intégrée, à certains des concepts-clés convoqués dans les deux champs épistémologiques. Il considère qu'il est nécessaire d'articuler et d'intégrer dans les curricula, les trois éléments que sont les contenus, les processus d'enseignement et l'évaluation, de façon à assurer la qualité et l'efficacité de l'éducation.

Le texte de **Walther Tessaro, François-Marie Gerard et Marcelo Giglio**, intitulé « *Changements curriculaires : un levier pour les pratiques évaluatives des enseignants ?* » porte sur la transposition des curricula dans les classes, en particulier au niveau des pratiques évaluatives des enseignants. Le texte dresse tout d'abord le constat que les curricula incluent rarement une composante évaluative, souvent sous le prétexte de laisser une marge de manœuvre et d'initiative aux enseignants. Si cet objectif paraît louable de prime abord, il alourdit aussi la responsabilité pédagogique des enseignants qui peuvent opérationnaliser de manière très différente la dimension évaluative de ces curricula en fonction de leurs contextes et de leurs exigences spécifiques. Pour les auteurs, il y aurait donc un intérêt manifeste à ce que les curricula comprennent une dimension évaluative (centrée sur la notion d'évaluation différenciée), mais qu'ils soient surtout construits en collaboration avec ceux qui sont chargés de leur implémentation. En lien avec les capacités transversales que l'on retrouve de plus en plus dans les référentiels, le texte plaide ensuite pour une clarification de la responsabilité de leur mise en œuvre dans les établissements scolaires et pour le développement de l'auto-évaluation des élèves en la matière. Enfin, à partir d'un exemple de recherche collaborative, le texte montre que les enseignants peuvent être à la source d'adaptations curriculaires, dans une relation dialectique renouvelée avec les autorités éducatives.

Le texte de **Domingos Fernandes, Isabel Fialho et Carlos Barreira**, intitulé « *Perceptions et pratiques d'évaluation des apprentissages dans l'enseignement universitaire portugais* », propose une problématisation des pratiques d'évaluation des apprentissages dans l'enseignement supérieur au Portugal et une analyse sur les représentations des étudiants et des enseignants sur l'évaluation, à partir d'une recherche conduite dans quatre universités. Cette étude conclut que la nature du travail pédagogique peut générer des environnements où l'enseignement, l'évaluation et l'apprentissage sont étroitement liés et organisés en vue de soutenir et d'améliorer la qualité de ce qui est appris. Cette recherche confirme l'importance de la distribution systématique et délibérée du *feed-back* aux étudiants et l'utilisation de l'évaluation formative pour promouvoir des processus d'apprentissage plus significatifs chez les étudiants. Ces résultats alimentent la discussion autour de l'évaluation *au service de l'apprentissage*.

Le texte d'**Annick Fagnant, Richard Etienne, Lucie Mottier Lopez et Marie-Noëlle Hindryckx**, intitulé « *L'évaluation comme objet d'apprentissage et comme outil de développement professionnel dans le cadre de la formation des enseignants* » s'appuie sur une analyse des conceptions et des pratiques d'évaluation en salle de classe. Il considère l'évaluation non seulement comme un outil de développement professionnel des futurs enseignants, mais aussi comme objet d'apprentissage spécifique. Pour les auteurs, il paraît essentiel de faire vivre aux étudiants le type de situations d'évaluation que l'on souhaite qu'ils mettent en œuvre dans leurs classes. Deux dispositifs de formation répondant à ce principe sont décrits par les auteurs à titre d'exemples. Le premier s'inscrit dans le cadre d'un cours consacré à l'évaluation et à la régulation des apprentissages, dispensé à l'université de Genève aux futurs enseignants du primaire. Il repose sur un portfolio qui poursuit notamment l'objectif de permettre aux étudiants d'expérimenter de nouvelles formes d'évaluation et de régulation formatives avec des interlocuteurs différents et des temporalités diverses. Le second dispositif concerne un cours de didactique disciplinaire en biologie dispensé à l'université de Liège aux futurs enseignants du secondaire. Entre mises en situation lors des cours théoriques et retours individuels et collectifs au sujet des pratiques observées en stage, il vise principalement l'apprentissage d'une posture réflexive par les étudiants ainsi qu'une ouverture aux autres formes de l'évaluation (*évaluation pour l'apprentissage et évaluation en tant qu'apprentissage*).

L'ensemble propose d'envisager les pratiques d'évaluation au-delà de leur fonction de mesure et de régulation de l'action qui, en définitive, est centrée sur un agir rationnel et un fonctionnalisme limité, transformant les acteurs en objets d'intervention. La perspective recherchée ici est plutôt de considérer l'évaluation comme agir communicationnel (Habermas, 1987) qui réunit des acteurs impliqués et qui renforce leur statut de sujets.

## Références

- Allal, L. (1979). Stratégies d'évaluation formative: conceptions psycho-pédagogiques et modalités d'application. In L. Allal, J. Cardinet & P. Perrenoud (Éd.), *L'évaluation formative dans un enseignement différencié* (pp. 153-183). Berne : Peter Lang.
- Ardoino, J. & Berger, G. (1986). L'évaluation comme interprétation, *Pour*, 107, 120-127.
- Ardoino, J. & Berger, G. (1989). Fondements de l'évaluation et démarche critique, *AECSE*, 6, 3-11.
- Dilthey, W. (1992). *Os Tipos de Concepção de Mundo*. Lusosofia net. Disponible en ligne: [http://www.lusosofia.net/textos/dilthey\\_tipos\\_de\\_concep\\_ao\\_do\\_mundo.pdf](http://www.lusosofia.net/textos/dilthey_tipos_de_concep_ao_do_mundo.pdf)
- Habermas, J. (1987). *Théorie de l'agir communicationnel*. Paris : Fayard.
- Hadi, C. (1994). *A Avaliação, Regras do Jogo. Das intenções aos instrumentos*. Porto: Porto Editora.
- Jorro, A. (2007). L'évaluation génératrice de développement professionnel ? In A. Jorro (Dir.) *Evaluation et développement professionnel* (pp.11-31). Paris : L'Harmattan.
- Mottier Lopez, L. (2015). *Évaluations Formative et Certificative des Apprentissages. Enjeux pour l'enseignement*. Louvain-la-Neuve : Bédica.
- Paquay, L. (2005). Vers quelles évaluations du personnel enseignant pour dynamiser leur développement professionnel et leur implication vers des résultats ? *Recherche et Formation*, 50, 55-74.
- Scriven, M. (1967). The methodology of evaluation. In R. Stake. *Curriculum evaluation*. Chicago: Rand McNally.
- Vial, M. (1999). Modèles et logiques de l'évaluation, Colloque International Ethique et qualité dans l'évaluation. Disponible en ligne : [http://www.michelvial.com/boite\\_96\\_00/1999-Modeles\\_et\\_logiques\\_de\\_l\\_evaluation.pdf](http://www.michelvial.com/boite_96_00/1999-Modeles_et_logiques_de_l_evaluation.pdf)

# L'évaluation des apprentissages et la régulation des politiques éducatives

**João Barroso**

Universidade de Lisboa

[jbarroso@ie.ulisboa.pt](mailto:jbarroso@ie.ulisboa.pt)

**Pour citer cet article :** Barroso, J. (2017). L'évaluation des apprentissages et la régulation des politiques éducatives. *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 9-19.

## 1. Introduction

L'évaluation interne des apprentissages constitue, depuis le début du XIXe siècle, un élément essentiel du processus de construction de la forme scolaire, basée sur l'enseignement simultané, et organisée en classes. L'évaluation servait à « classer » les élèves, régulant leur progression scolaire, afin de maintenir un minimum d'homogénéité parmi ceux qui assistaient à la même classe, pouvant avoir accès aux mêmes contenus des programmes. Plus tard, avec l'institutionnalisation des examens, l'évaluation externe est devenue un élément constitutif et structurant du système d'enseignement dans son ensemble. En plus de leur fonction rituelle et symbolique, les examens constituaient un instrument de gouvernance de l'État éducateur et un dispositif central de la régulation bureaucratique et professionnelle du travail enseignant et des écoles.

Aujourd'hui, l'évaluation des apprentissages s'est intensifiée et est devenue plus sophistiquée avec l'utilisation systématique des tests standardisés au niveau national ou international. En outre, elle commence à intégrer un ensemble de plus en plus large de modalités et d'objets d'évaluation résultant des politiques additives, allant des apprentissages aux buts et compétences, des élèves aux enseignants, aux chefs d'établissement et aux écoles.

L'existence de cet ensemble d'évaluations fait partie du processus d'émergence du soi-disant État évaluateur et s'insère dans un changement profond des modes de régulation de l'action publique : autrement dit, des modes de produire et d'appliquer les règles qui guident l'action des acteurs et le rôle que l'État assume dans ce processus, grâce à différents dispositifs de coordination, de contrôle et d'influence. De ce fait, la relation entre l'évaluation et les apprentissages peut être perçue à partir d'une approche basée sur l'analyse des politiques publiques et sur la problématique de la régulation (Barroso, 2005).

Ceci est le point de départ de notre exposé<sup>1</sup> qui a pour objectif principal d'analyser l'évaluation comme un instrument de régulation des politiques éducatives, selon le schéma suivant :

- Dans une première partie, nous allons mettre en évidence comment, historiquement, l'évaluation des apprentissages est devenue un élément central de la construction de la forme scolaire et de la régulation bureaucratique et professionnelle ;
- Dans une deuxième partie, nous allons nous concentrer sur l'actualité pour mettre en exergue deux grandes tendances dans le domaine de l'évaluation éducative : l'intensification de l'évaluation externe avec l'utilisation de tests standardisés et la diversification des modalités et des objets d'évaluation dans l'éducation ;
- Dans une troisième (et dernière partie), nous allons situer cette intensification et diversification par rapport à l'évaluation des apprentissages, dans le contexte de l'émergence d'une régulation post (ou néo) bureaucratique, en soulignant l'importance que les instruments de régulation acquièrent dans les nouveaux modes de gouvernance.

## **2. L'évaluation des apprentissages comme un élément central de la construction de la forme scolaire et de la régulation bureaucratique et professionnelle**

Comme on le sait, l'origine de l'école moderne est basée sur l'invention d'une organisation éducative spécifique pour l'enseignement collectif qui constitue l'une des marques essentielles de ce que Guy Vincent (1980) appelle la « forme scolaire » et d'autres, comme Tyack et Cuban (1995), la « grammaire de l'école ». Cette organisation, ayant survécu jusqu'à nos jours, a dans la « classe » un élément central, en tant que structure, qui permet de mettre en œuvre l'« enseignement simultané » (Barroso, 1995).

Selon la définition donnée en 1911, dans son *Dictionnaire de Pédagogie*, Buisson expliquait ainsi ce qu'était la « classe » :

*« Une organisation pédagogique rationnelle (...) selon la force moyenne des enfants dans l'ensemble des matières scolaires (...). Cette organisation permettra à l'enseignant de s'adresser à tous les élèves au même temps et de leur donner le même enseignement d'après la force moyenne des enfants qu'il a devant lui » (Buisson, 1911).*

Cette organisation était commune à l'enseignement primaire et secondaire, mais, au niveau de l'enseignement dans les écoles secondaires, la fonction homogénéisatrice de la classe couvre non seulement les élèves, mais aussi les enseignants et les connaissances, à travers la « concentration des matières », la « coordination des enseignants » et la « connexion de l'enseignement » (Barroso, 1995).

La classe, qui était d'abord une simple répartition des élèves, se transforme progressivement en un critère organisationnel pour hiérarchiser le service des enseignants et l'espace scolaire lui-même. En même temps, elle acquiert la valeur de « mesure » dans la progression des élèves

---

<sup>1</sup> Cet exposé est basé sur le soutien écrit de ma conférence au 28<sup>e</sup> colloque de l'ADMÉE-Europe, le 13 janvier 2016. J'y utilise surtout la réflexion produite dans le cadre des recherches menées sur l'évolution historique de l'organisation pédagogique et de l'administration des lycées au Portugal (Barroso, 1995) et, plus récemment, dans le cadre de deux projets de recherche avec d'autres universités européennes, sur les modes de régulation des politiques éducatives (Reguleducnetwork, voir, entre autres, Maroy, 2006 et Barroso, org. 2006) et sur la relation entre la connaissance et la politique en éducation (KNOWandPOL, voir les informations sur le projet à l'adresse <https://www.knowandpol.eu> et consulter, entre autres, Barroso e Afonso, org. 2011, Carvalho, coord. 2011).

(passer en classe supérieure) et dans l'organisation du temps lié au parcours scolaire (la classe devient synonyme d'année d'enseignement).

Avec cette architecture curriculaire, la classe devient le groupement nucléaire de l'organisation pédagogique de l'école primaire et le classement des élèves, l'une des principales fonctions organisationnelles pour son bon fonctionnement.

Dans ce contexte, comme l'indiquait le pédagogue français Campagne en 1872, dans son « Dictionnaire universel d'éducation et d'enseignement » :

*« Classer les élèves c'est les diviser en groupes susceptibles de recevoir avec profit les mêmes leçons. (...) Pour déterminer dans quelle classe un enfant doit entrer primitivement, ou quand il doit sortir de celle où il a été placé d'abord, on aura égard, en général, non pas à l'âge de l'enfant, non pas au temps plus ou moins long qu'il aura déjà passé à étudier, mais à son état actuel et réel d'instruction. (...) Le classement des élèves, pratiqué avec discernement, est une garantie d'ordre et d'harmonie dans une classe » (Campagne, 1886).*

Le classement des élèves devient progressivement un instrument de nivellement des groupes et une garantie de la constitution même des classes homogènes. L'évaluation des apprentissages sur laquelle repose ce classement acquiert ainsi une fonction administrative destinée à organiser l'école. Et plus l'organisation de l'école devient complexe, et plus grands sont le nombre et la diversité des élèves, plus cette opération s'avère décisive pour son bon fonctionnement.

Il faut ainsi relever que l'évaluation des apprentissages commence par avoir d'abord une fonction pronostique dictée par la nécessité de former, au début de l'année scolaire, des groupes homogènes d'élèves qui peuvent recevoir, en même temps, le même enseignement par le même enseignant. Sa principale fonction n'était pas de réguler les apprentissages, mais de réguler l'organisation scolaire et son administration.

## **2.1 Évolution : de l'évaluation à l'entrée à l'évaluation à la sortie**

Avec l'évolution du système d'enseignement et le renforcement d'une régulation bureaucratique, l'évaluation devient particulièrement sommative, sélectionnant les élèves, non pas à l'entrée, mais à la sortie de l'année scolaire, afin de les distribuer l'année suivante, par classe supérieure ou dans la même classe et, dans ce cas, en utilisant le mécanisme du redoublement (très contesté dès le départ pour des raisons pédagogiques, mais qui a jusqu'à présent survécu pour des raisons bureaucratiques et administratives).

C'est la fonction initiale des examens qui, dans son processus d'institutionnalisation, s'élargit petit à petit à d'autres fonctions, non seulement dans le domaine scolaire, mais aussi social et politique.

Au Portugal, on peut vérifier que, pour la première fois, et de manière explicite, l'utilisation de l'évaluation externe de l'apprentissage, comme instrument de régulation sociale et politique, coïncide avec le début du régime *Estado Novo* (le régime dictatorial qui a subsisté au Portugal entre 1933 et 1974).

En fait, le renforcement de la sélection et de la réhabilitation de l'examen sont les deux réponses les plus décisives de la politique éducative d'*Estado Novo*, pour essayer d'arrêter ce qui était considéré comme la perversion de la fonction sociale du lycée en tant qu'institution dédiée à la formation des élites et la base pour l'accès à l'enseignement supérieur. En même temps, cette politique était destinée à préserver son organisation, sa gestion et son exploitation, des dommages causés par la fréquence excessive et souvent sans la qualité souhaitée.

À cette époque, la défense de la sélection et des examens était soutenue par deux types d'arguments qui se mélangeaient et qu'il importe de distinguer, à savoir :

- Une raison pédagogique qui vise à éviter qu'un excès d'élèves puisse nuire à l'action enseignante et empêcher de profiter de l'organisation pédagogique utilisée pour l'enseignement collectif, basée sur le régime des classes, mentionné ci-dessus. Cette raison est visible, par exemple, dans les arguments utilisés par la loi pour justifier l'introduction de l'examen d'entrée au lycée (à la fin de l'enseignement primaire obligatoire) :

*« [Pour éviter que] dans l'enseignement secondaire entrent des élèves qui ne sont pas correctement sélectionnés et par ce motif réagissent de manière divergente à l'action de l'enseignant, ce qui rend l'enseignement collectif presque entièrement impossible (...) il est établi [par le présent décret] que l'admission à la fréquentation de l'enseignement secondaire dépend de la réalisation d'un examen d'entrée, visant à déterminer si les candidats possèdent certaines connaissances fondamentales et si, en plus, ils sont en mesure de répondre de façon congruente à l'action éducative qu'il est prévu d'accomplir grâce à cet enseignement » (Préambule du décret 25461, du 5 juin 1935).*

- Une raison politique et sociale qui vise à mettre l'école au service d'une technologie sociale - comme un ensemble de mesures ciblées afin d'influencer l'ajustement des attentes subjectives des utilisateurs de l'école par rapport à leur destination objective, en préservant la conformité de chacun avec sa condition sociale (Grácio, 1986).

Comme l'écrivait le ministre de l'Éducation, Eusébio Tamagnini, en 1936, pour justifier les mesures qu'il a prises à cet égard (y compris la susmentionnée introduction de l'examen d'entrée au lycée) :

*« Nos écoles ne peuvent pas s'inspirer d'un sentiment romantique de faire passer tous les jeunes par les mêmes cours, en les soumettant à une éducation uniforme qu'éveille leurs ambitions et leurs désirs de positions sociales équivalentes. Au contraire, le bien de la Nation exige des écoles de formation différenciées où chaque individu reçoit l'éducation en fonction de ses capacités et aptitudes. Le bien de la Nation et le service social exigent, par conséquent, que dans l'école soit pratiquée une sélection rigoureuse dès que possible » (Tamagnini, 1936).*

L'importance que l'examen atteint à partir de 1930 le fait devenir l'objectif premier d'une grande partie de la législation relative à l'enseignement secondaire pendant cette période. D'une part, on assiste à une tentative de créer des instruments qui, d'une manière rapide, permettent de trier correctement les élèves. D'autre part, les examens se transforment en des miroirs des lycées, ce qui oblige le proviseur à leur consacrer un soin particulier et l'administration à exercer un contrôle strict sur leur mise en œuvre.

De plus, la période des examens devient l'un des rares moments où il y a une interaction entre l'école et la communauté (souvent dans un sens négatif, en raison des accusations d'irrégularités ou d'abus de pouvoir de l'évaluation qui sont notoires dans la presse).

Bien que, au Portugal, avec la fin de la dictature (1974), les examens aient perdu la connotation idéologique qui leur avait été donnée par *Estado Novo*, ils ont gardé jusqu'à présent (implicitement ou explicitement) un ensemble de caractéristiques qui vont au-delà de la fonction de régulation et de certification des apprentissages, dont nous mettons en évidence ceux liés à l'organisation pédagogique, à l'administration scolaire et à la politique éducative : le classement des élèves pour assurer leur répartition entre des groupes homogènes ; un instrument de régulation du flux d'élèves tout au long de leur parcours scolaire ; la définition *a posteriori* des normes d'apprentissage ; un moyen moral pour induire chez les élèves le culte de l'effort, de la peur et de la punition ; un instrument de sélection scolaire pour pratiquer la technologie sociale.

Ces caractéristiques sont intégrées dans le modèle de l'État éducateur qui a prévalu pendant la majeure partie du XXe siècle au Portugal, mettant en évidence le rôle de l'évaluation des apprentissages (en particulier par le biais des examens) dans la régulation bureaucratique et professionnelle des politiques éducatives (Barroso, 2000).

## **2. Tendances actuelles : intensifier et diversifier les modalités et les objets d'évaluation**

Comme nous l'avons vu, l'évaluation des apprentissages avait, dès le début, une fonction de régulation qui a dépassé la dimension pédagogique et s'est élargie aux aspects organisationnels, administratifs et politiques. Aujourd'hui, nous pouvons dire qu'il existe, principalement, deux phénomènes : l'intensification et la diversification.

### **2.1 L'intensification**

Depuis la fin du XXe siècle, on a assisté, en particulier dans les pays anglo-saxons, mais s'étendant un peu sur toute la planète, et compte tenu de l'impact de la régulation transnationale, à la prolifération des tests standardisés comme une forme d'évaluation des apprentissages et de pilotage du système d'enseignement.

Les États-Unis sont l'un des pays où cette obsession de l'évaluation a été portée à l'extrême, après avoir été encouragée, dès le début, par le gouvernement fédéral pendant le temps de l'administration Bush (fils), associée au programme de réforme *No Child Left Behind* (2002), et sous l'administration Obama, associée au programme *Race to the Top* (2009).

Ces dernières années, cette croyance en la vertu des tests standardisés a été contestée aux États-Unis, même parmi les personnes ayant des responsabilités politiques dans leur diffusion, le cas le plus connu étant celui de l'ancienne secrétaire d'État Diane Ravitch (2010).

En dépit de ces critiques, la politique d'intensification des tests d'évaluation continue jusqu'en octobre 2015, lorsque le Secrétaire à l'éducation Arne Duncan demande de démissionner et le président Obama fait son « *mea culpa* », comme le rapporte la chaîne CNN le 24 octobre 2015 :

*After years of complaints from teachers, parents and students alike, the Obama administration on Saturday announced new guidelines toward standardized tests, saying kids spend too much time taking "unnecessary" exams in schools.*

*In a Facebook video message, President Barack Obama said he hears from parents who worry about "too much testing, and from teachers who feel so much pressure to teach to a test that it takes the joy out of teaching and learning."*

*"I want to fix that," he said.*

Après ces déclarations, le gouvernement fédéral impose une limite de « 2% du temps scolaire » pour l'application des tests et encourage les diverses autorités d'État à diversifier les instruments d'évaluation pour couvrir de multiples manifestations de la performance des élèves.

Bien qu'elle n'ait pas le même impact, l'intensification des examens et des tests était présente dans d'autres pays et avec des résultats similaires, en particulier en ce qui concerne un rétrécissement des curricula et l'utilisation des résultats de l'évaluation de la performance des élèves, en tant que mode de contrôle de la qualité des enseignants et des écoles.

Au Portugal, avec le gouvernement de droite, entre 2011 et fin 2015, nous avons eu quelques vestiges de cette politique de renforcement des examens. Rien de comparable en quantité avec ce qui s'est passé aux États-Unis, mais avec les mêmes objectifs : institutionnaliser une culture de l'évaluation, non seulement comme une forme de contrôle des apprentissages et de régulation du travail enseignant, mais surtout de définition curriculaire. Autrement dit, les instruments de mesure déterminent ce qui doit être mesuré, c'est-à-dire, les examens déterminent ce qui doit être enseigné (et ce n'est pas l'enseignement qui détermine ce qui doit être évalué).

## 2.2 La diversification

Le second phénomène qui caractérise actuellement l'évaluation des apprentissages est que celle-ci ne constitue plus le seul élément d'évaluation éducative. En fait, au Portugal, pendant les deux dernières décennies, de multiples processus d'évaluation ont commencé à se développer dans le champ éducatif, en se concentrant sur des objets différents, faisant référence à différents domaines de connaissance et ciblant des objectifs précis. Juste pour nommer les évaluations les plus fréquentes : l'évaluation des résultats scolaires des élèves ; l'évaluation externe et interne des écoles ; l'évaluation de la performance du personnel enseignant et non enseignant ; l'évaluation des systèmes d'enseignement ; l'évaluation et la garantie de la qualité ; etc.

La plupart du temps, ces évaluations sont déclenchées sans aucune articulation et les résultats obtenus avec l'application d'un dispositif particulier, lorsqu'ils sont analysés, ne prennent, en général, pas en compte les résultats obtenus par d'autres évaluations portant sur des champs connexes.

En ce qui concerne spécifiquement le cas portugais, il est possible de dire que cette « *évaluocratie* » (ainsi appelée par Correia, 2005 et 2010) a, pour l'instant, et surtout, un effet symbolique de création d'un véhicule de transition des cultures scolaires - entre la coopération et la compétition - et fait partie, comme nous le verrons dans la troisième partie, d'un changement plus large des modes de régulation des politiques et de l'action publique.

D'un point de vue de l'analyse politique, les études qui ont été menées sur ces différentes modalités d'évaluation (au Portugal, ces études portent principalement sur l'évaluation des écoles) montrent qu'il existe une tension évidente entre les fondements explicites justifiant les mesures prises et les effets recherchés par les autorités publiques chargées de la décision et de l'exécution.

Habituellement, l'évaluation est toujours justifiée comme faisant partie d'un processus global de rationalisation de l'action publique dont l'objectif est non seulement d'assurer la rationalité de la décision, mais aussi l'efficacité de l'action elle-même et de ses effets. En ce sens, elle est insérée dans le cycle normal de gestion (planification - mise en œuvre - évaluation - rétroaction), ayant pour but principal d'identifier les problèmes (au niveau des processus et des résultats) et de contribuer à leur résolution. L'évaluation est perçue, dans ce cas, comme un moyen de mieux gérer les ressources et d'améliorer la qualité et l'efficacité des politiques publiques.

Cependant, et en dépit de cette rhétorique, nous constatons que, dans la pratique, le caractère instrumental de l'évaluation est réduit et que les mesures sont prises pour nourrir bien plus une illusion de rationalité que sa mise en pratique. Comme le disent Floden et Weiner (1978), à cet égard, l'évaluation dans le domaine des politiques publiques fonctionne souvent comme un rituel dont le rôle est de calmer l'inquiétude des citoyens et de perpétuer l'image d'efficacité

et de responsabilité du gouvernement, créant l'impression qu'il est sérieusement engagé à poursuivre les objectifs.

Bien sûr, si l'on se réfère à l'action spécifique des acteurs dans le champ éducatif, nous constatons qu'il y a des situations très différentes et qu'il est possible de trouver, dans de nombreux cas, des preuves des effets positifs de l'évaluation (que ce soit des élèves, des enseignants ou des écoles). Ceci se constate, par exemple, dans le cas des élèves avec l'« évaluation formative » des apprentissages ou dans le cas des écoles avec l'« évaluation institutionnelle » et sa contribution à l'amélioration des écoles (Bolívar, 2016), et avec la dénommée « évaluation démocratique » (Afonso, 1998 et 2014).

En guise de synthèse, nous pouvons dire que nous vivons dans une période de politiques intensives dans le domaine de l'évaluation éducative. La coexistence de différentes modalités et objets d'évaluation implique une redéfinition de l'évaluation des apprentissages et de ses fonctions. Dans ce sens, l'évaluation des apprentissages ainsi que d'autres évaluations (des enseignants, des écoles, des cours, etc.), acquièrent un sens politique évident en tant qu'instruments de la gouvernance dans le cadre des principes de la nouvelle gestion publique (New Public Management), à savoir la définition de « buts » et de « repères » précis et objectifs, la construction d'indicateurs mesurables et fiables et l'évaluation rigoureuse des résultats. L'évaluation en général (quelle que soit sa modalité) devient ainsi un instrument de contrôle des dépenses, de répartition des ressources, de responsabilisation des gestionnaires et d'émulation de ses agents.

Voilà pourquoi une lecture intégrée de ces différents types d'évaluation est justifiée, situant leur apparition et leur développement dans le contexte de l'émergence de nouveaux modes de régulation des politiques publiques, « post (ou néo) – bureaucratiques ».

### **3. L'évaluation en tant qu'instrument de régulation post (ou néo) bureaucratique**

Comme nous l'avons vu dans la première partie, l'évaluation des apprentissages par des examens constitue, dès le début, un élément central de la construction et du développement de l'État éducateur au cours du XIXe siècle et d'une grande partie du XXe siècle. En plus d'être une pierre angulaire de la création d'une organisation pédagogique uniforme (enseignement en classe), les examens ont permis de renforcer le contrôle bureaucratique sur le curriculum et le travail des enseignants et des élèves. De cette manière, l'État garantissait, non seulement des règles obligatoires d'enseignement et d'apprentissage relatives aux programmes définis au niveau central, mais aussi, en fonction de ses propres critères politiques, la gestion des flux d'élèves tout au long des divers niveaux d'enseignement.

Comme nous l'avons vu dans la deuxième partie, à l'heure actuelle, l'intensification et la diversification des évaluations dans l'éducation donnent une nouvelle signification politique à ces processus d'évaluation, ne pouvant être comprise que dans le contexte de l'émergence de nouvelles formes de gouvernance et de régulation des politiques publiques.

En effet, nous sommes confrontés aujourd'hui, au Portugal comme dans d'autres pays, à une restauration du rôle de l'État dans la définition de l'éducation qui est fournie et qui ne passe pas vraiment par un changement de son importance relative, au niveau du financement et de la garantie des ressources. Il s'agit surtout d'un changement quant au rôle de l'État, de son implication et de l'importance de celle-ci à travers ses interventions dans la définition, la mise en œuvre et le contrôle des politiques éducatives.

Ce phénomène a été particulièrement visible dans les pays qui ont rejoint le projet *Reguleducnetwork* (Belgique, France, Hongrie, Portugal et Angleterre) et est identifié par ce que désignons comme la « régulation post-bureaucratique » (Maroy, 2006, Barroso, org. 2006). Ce type de régulation se caractérise principalement par le fait de ne pas reposer, à l'instar du modèle bureaucratique, sur la légitimité de la rationalité et de la loi, considérées, simultanément, comme des formes de coordination et une source de valeurs. La rationalité continue d'être appréciée, mais elle est surtout réduite à la rationalité instrumentale. Le contrôle *a priori*, par les normes, est remplacé par le contrôle *a posteriori*, par les résultats.

Cependant, l'émergence de ces nouvelles formes de régulation post-bureaucratique ne signifie pas que les formes antérieures de contrôle aient été abandonnées.

Voilà pourquoi nous assistons, au Portugal, comme dans beaucoup d'autres pays, soit à la coexistence de la « bureaucratie » et de la « post-bureaucratie », dans la définition et la mise en œuvre des politiques publiques de l'éducation, soit à l'émergence de formes explicites de néo-bureaucratie (Barroso, 2013).

Dans ce type de régulation (néo-bureaucratie), l'objectif principal est d'améliorer le fonctionnement bureaucratique de l'administration, compte tenu de la complexité croissante et de la diversité du système. Pour ce faire, les aspects morphologiques sont modifiés sans compromettre les aspects essentiels à l'exercice du pouvoir et de l'autorité de l'État et de l'efficacité des processus de contrôle.

Ainsi, nous entrons dans le domaine des « bureaucraties douces », selon Courpasson, des organisations dont les stratégies, en dépit d'être managériales, sont orientées vers la construction des politiques de centralisation.

*Le concept de « bureaucratie douce » cherche à exprimer l'émergence d'une centralisation politique des organisations, ciblant le développement de formes décentralisées pour mener son activité : les fonctions et les responsabilités deviennent de plus en plus décentralisées, mais les décisions politiques plus centralisées (Courpasson, 2000).*

Dans ce contexte, l'évaluation devient, dans l'éducation, un outil souhaité pour concilier l'autonomie et la responsabilité ou pour veiller à ce que la décentralisation de la gestion ne mette pas en cause le contrôle de la décision.

L'évaluation intègre ainsi un large éventail d'instruments basés sur les connaissances, spécifiquement produits pour exercer un effet de régulation, tels que, par exemple, les dispositifs d'évaluation, les référentiels de compétences, les contrats d'autonomie, les buts d'apprentissage, les « bonnes pratiques », etc.

Avec le développement des TIC, ces instruments sont associés à ce que Williamson (2005) appelle « *digital education governance* » (la gouvernance numérique de l'éducation). En fait, selon cet auteur, la gouvernance de l'éducation est fortement influencée par l'« objectif de la quantification de tous les types de comportements humain et social » ou, pour utiliser l'expression divulguée par Rose (1991), entre autres, la « gouvernance par les chiffres ». Ceci se reflète dans la construction et l'alimentation systématique de grandes bases de données, avec l'utilisation des technologies numériques que Williamson (2005) appelle les « instruments de la politique en temps réel ».

C'est ce qui se passe dans nos écoles, où les enseignants et les chefs d'établissement passent plus de temps à alimenter les plus diverses plates-formes informatiques de collecte et de visualisation des données en ligne que de réguler le domaine de l'évaluation, la gestion des ressources, l'organisation des horaires, etc. Cette « régulation par les instruments » permet

ainsi un contrôle indirect et, souvent à distance, des politiques éducatives, se trouvant présente à tous les niveaux (transnational, national, local) et étant responsable du renforcement des tendances néo-bureaucratiques mentionnées auparavant.

Enfin, il convient de noter l'importance de l'évaluation au niveau du pilotage et de la régulation des systèmes d'enseignement, en particulier l'évaluation des apprentissages par des tests standardisés, ce qui résulte également d'une réglementation transnationale accrue (notamment à travers les tests et les rapports PISA)<sup>2</sup>.

Carvalho affirme, à propos de l'évaluation internationale comparée de la performance des élèves, mise en œuvre par l'OCDE, par le biais du PISA (*Programme international de l'OCDE pour le suivi des acquis des élèves*) :

*« Les études ont montré la diversité des modes d'utilisation des données et des textes PISA dans les espaces nationaux, dans les univers de la recherche, de la politique et de l'administration de l'éducation (Carvalho et al, 2011) : pour élaborer des analyses des politiques en cours ; pour justifier des questions spécifiques de l'agenda politique ; comme source d'études sur les « réalités nationales » ; pour construire (ou développer) des instruments de régulation autochtones (notamment dans le domaine des politiques de reddition de comptes et / ou de garantie de la qualité) » (Carvalho, 2013, p. 70).*

La gouvernance de l'éducation, à travers la comparaison et « les chiffres », devient l'un des cas les plus flagrants de cette régulation transnationale par l'« effet PISA » (voir, entre autres, Barroso et Carvalho, 2008 et Carvalho, 2014).

#### **4. Conclusion**

L'objectif de l'approche présentée n'est pas de nier l'importance de l'évaluation dans l'éducation, mais plutôt de remettre en question ses objectifs, ses modalités et ses effets, à la lumière des changements actuels des modes de gouvernance et de régulation des politiques éducatives. Comme il ressort de la recherche et de la littérature existante, l'évaluation de l'apprentissage est une étape clé de l'enseignement et de la formation, ainsi que du processus et de l'amélioration de ses résultats. Cependant, lorsque nous analysons les objectifs, les processus et les effets des politiques de l'évaluation qui ont été prises, dans divers pays au cours des deux dernières décennies, nous constatons que, plutôt que la recherche de l'efficacité, ce qui est en cause dans cette utilisation massive de l'évaluation, surtout externe, c'est l'adaptation des procédures de coordination et de contrôle du système d'enseignement aux caractéristiques de la régulation post-bureaucratique et de la régulation transnationale émergente dans les nouvelles formes de gouvernance.

À la fin du projet *Regnuleducnetwork*, nous avons soulevé l'hypothèse de nous trouver face à l'existence de deux modèles dans ce processus, à savoir : celui de l'« État évaluateur » et celui du « quasi-marché » (Maroy, 2006).

Dans le premier, la régulation est effectuée principalement par la définition des buts et le contrôle des résultats, à travers la contractualisation et l'évaluation.

Dans le second, la régulation se fait principalement par la concurrence entre les opérateurs, publics ou privés, par le biais de la différenciation et la possibilité de choix des écoles. L'évaluation constitue, dans ce processus, non pas un instrument d'amélioration, mais un

---

<sup>2</sup> Une étude approfondie sur PISA comme instrument de « régulation basé sur la connaissance » (au niveau supranational et national), conduite par Luís Miguel Carvalho, a été développée dans le cadre du projet *Knowandpol*. Voir, entre autres, Carvalho, coord. 2011, 2013 et 2014.

instrument d'affichage des différences et de promotion d'une relation de compétition entre les écoles, afin que les consommateurs puissent faire un choix éclairé.

Comme l'affirmait Maroy à la fin du projet :

« Les politiques éducatives des cinq pays étudiés s'inspirent peu ou prou des modèles post-bureaucratiques évoqués. Le modèle de l'État évaluateur inspire ainsi la promotion simultanée de l'autonomie des établissements et de l'évaluation et, pour ce qui concerne les pays décentralisés, le renforcement des objectifs centraux et d'une standardisation curriculaire. Simultanément des traits plus inspirés du modèle du quasi-marché - la tolérance ou la promotion du libre choix, la relative diversification de l'offre scolaire de nature à rencontrer les demandes variées des usagers -, ont été également développés. Cependant, le degré d'intensité des politiques menées et le dosage entre ces modèles sont très variés » (Maroy, 2006, p.57).

Cette division est, comme cela se comprend, principalement analytique et la réalité des différents pays est plus hybride, avec la coexistence d'anciens et de nouveaux modes de régulation, ainsi que de logiques évaluatives diverses se chevauchant.

Certains cas, comme en Finlande, ont pour but de surmonter cette dichotomie, en mettant en cause le système d'« évaluation - compétition – concurrence » :

« (...) contrairement à de nombreux autres systèmes éducatifs contemporains, le système finlandais n'est pas infecté par une compétition de quasi-marché et par les politiques risquées d'évaluation par les tests. (...) La mesure et le contrôle des acquis sont très présents au sein de l'école finlandaise. Mais selon des modalités presque toujours conçues et réalisées en interne, par les professeurs eux-mêmes. Nous avons volontairement limité au strict minimum les tests standardisés, mis au point par des acteurs étrangers au système » (Sahlberg, 2011, p.39).

Nous sommes donc dans un processus de transition. Mais, indépendamment de son évolution, l'évaluation est un instrument de gouvernance qui ne nous quittera plus. Il est donc important d'adopter une vision holistique du processus et de prendre en compte la diversité des objets et des modalités d'évaluation existant dans le domaine de l'éducation. Pour ce faire, il est nécessaire d'adopter des approches différentes selon la pluralité des savoirs et des domaines scientifiques présents dans les sciences de l'éducation.

## 5. Références

- Afonso, A. (1998). *Políticas educativas e avaliação educacional*. Braga: Universidade do Minho.
- Afonso, A. (2014). *Avaliação Educacional: Regulação e Emancipação*. S. Paulo: Cortez Editora, 6<sup>e</sup> édition.
- Barroso, J. (1995). *Os liceus : organização pedagógica e administração (1836-1960)*. Lisboa : Fundação Calouste Gulbenkian e Junta Nacional de Investigação Científica e Tecnológica.
- Barroso, J. (2000). Autonomie et modes de régulation locale dans le système éducatif. *Revue Française de Pédagogie*, 130, 57-71. Disponible : [http://www.persee.fr/doc/rfp\\_0556-7807\\_2000\\_num\\_130\\_1\\_1053](http://www.persee.fr/doc/rfp_0556-7807_2000_num_130_1_1053)
- Barroso, J. (2005). Les nouveaux modes de régulation des politiques éducatives en Europe : de la régulation du système à un système de régulations. In Duterq, Y. (Dir.). *Les régulations des politiques d'éducation* (pp. 151-171). Rennes: Presses Universitaires de Rennes.
- Barroso, J. (Org.). (2006). *A regulação das políticas públicas de educação. Espaços, dinâmicas e actores*. Lisboa : EDUCA. Disponible: <http://hdl.handle.net/10451/5761>
- Barroso, J. (2013). Conhecimentos, políticas e práticas em educação. In Martins, Â. et al. (Orgs.). *Políticas e gestão da educação: desafios em tempos de mudanças* (pp. 1-23). Campinas, SP: AutoresAssociados.
- Barroso, J. & Carvalho, L. M. (2008). PISA : un instrument de régulation pour relier des mondes. *Revue Française de Pédagogie*, juillet-août-septembre 2008, 77-80.

- Barroso, J. & Afonso, N. (Orgs). (2011). *Políticas educativas: Mobilização de conhecimento e modos de regulação*. V. N. de Gaia: Fundação Manuel Leão.
- Bolívar, A. (2016). Como puede la evaluación institucional contribuir para mejorar la escuela? *Est. Aval. Est. Aval. Educ.*, São Paulo, p. 1-30, juil. 2016. Disponible : [https://www.researchgate.net/publication/304993684\\_Como\\_puede\\_la\\_evaluacion\\_institucional\\_contribuir\\_para\\_mejorar\\_la\\_escuela](https://www.researchgate.net/publication/304993684_Como_puede_la_evaluacion_institucional_contribuir_para_mejorar_la_escuela)
- Buisson, F. (1911). *Nouveau Dictionnaire de Pédagogie et d'Instruction Primaire*. Paris : Hachette.
- Campagne, E. M. (1886). *Dicionário Universal de Educação e Ensino. (Tradladado a português e ampliado nos vários assumptos relativos a Portugal por Camilo Castelo Branco)*. Porto: Livraria Internacional de Ernesto Chardon.
- Carvalho, L. M. Coord. (2011). *O espelho do perito: Inquéritos internacionais: Conhecimento e política em educação – O caso PISA*. V. N. de Gaia: Fundação Manuel Leão.
- Carvalho, L. M. (2013). Mutua vigilância organizada. In: *Educação - Temas e Problemas* 6, 12-13: 61-74. Disponible : <http://www.ciep.uevora.pt/revista/Completa/RevistaEDUCACAO.pdf>
- Carvalho, L. M. (2014). The attraction of mutual-surveillance of performances: PISA as a knowledge-policy instrument. In Fenwick, T. Ozga J. & Mangez E. (Eds.). *World Yearbook of Education 2014: Governing Knowledge: Comparison, Knowledge-based Technologies and Expertise in the Regulation of Education* (pp. 58-72). London: Routledge.
- Correia, J. (2005). Contributos para a construção de 'narrativas educativas' de esquerda. *Perspectiva*, Florianópolis, v. 23, n. 2, 407-426.
- Correia, J. (2010). La "evaluocracia": el papel de la evaluación en la legitimación y reconstrucción institucional de la educación. *Avances en Supervisión Educativa. ADIDE* n. 13, 1-9.
- Courpasson, D. (2000). "Managerial Strategies of Domination: power in soft bureaucracies". *Organization Studies*, 2000, 21/1, 141-161. Disponible : <http://www.carstensorensen.com/download/Courpasson2000.pdf>
- Floden, R.E., & Weiner, S.S. (1978). Rationality to ritual: The multiple roles of evaluations in governmental processes. *Policy Sciences*, 1978, 9, 9-1
- Grácio, S. (1986). *Política Educativa como Tecnologia Social – As Reformas do Ensino Técnico de 1948 a 1983*. Lisboa : Livros Horizonte.
- Maroy, Ch. (2006). *École, régulation et marché. Une comparaison de six espaces scolaires locaux en Europe*. Paris : PUF.
- Ravitch, D. (2010). *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic Books.
- Rose, N. (1991) Governing by numbers: figuring out democracy. *Accounting Organizations and Society, Vol. 16, No. 7*, 673-692. Disponible : <http://www.collier.sts.vt.edu/sts5424/pdfs/rose.pdf>
- Sahlberg, P. (2011). *Finnish lessons. What Can the World Learn from Educational Change in Finland?* New York and London: Teachers College Press.
- Tamagnini, E. (1936). *Alguns Aspectos do Problema Escolar Português*. Coimbra: Imprensa da Universidade.
- Tyack, D. & Cuban, L. (1995). Tinkering Toward Utopia. *Tinkering Toward Utopia. A Century of Public School Reform*. Cambridge: Harvard University Press.
- Vincent, G. (1980). *L'École Primaire Française*. Lyon: Presses Universitaires.
- Williamson, B. (2015). Digital education governance: data visualization, predictive analytics, and 'real-time' policy instruments, In *Journal of Education Policy*, Published online: 29 Apr 2015. Disponible : <http://dx.doi.org/10.1080/02680939.2015.1035758>



# Le pouvoir des contextes évaluatifs

**Marie-Christine Toczek**

Université Clermont Auvergne, ACTé  
[M-Christine.Toczek\\_Capelle@uca.fr](mailto:M-Christine.Toczek_Capelle@uca.fr)

**Carine Souchal**

Université Clermont Auvergne, ACTé,  
[carine.souchal@gmail.com](mailto:carine.souchal@gmail.com)

**Pour citer cet article :** Toczek, M-Ch. & Souchal, C. (2017). Le Pouvoir Des Contextes Evaluatifs. *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 21-35.

## 1. Introduction

La dernière enquête PISA révèle qu'en France, les élèves issus de milieux défavorisés ont quatre fois plus de chance de faire partie des élèves qui ont les plus mauvais scores. Seuls 2% de ce même groupe d'élèves se classent parmi les élèves les plus performants. En décembre 2016, cette information préoccupante a été relayée par l'ensemble des médias. Parmi les pays de l'OCDE, les résultats de la France sont les plus fortement corrélés avec le niveau socio-économique et culturel des familles. Cette corrélation, tristement récurrente, est présentée comme stable depuis 2006.

En d'autres termes, la France est présentée, depuis plus de dix ans, comme le pays de la reproduction sociale, la même conclusion est reprise par l'ensemble des médias s'intéressant à l'école : « *plus on vient d'un milieu social défavorisé et moins on a de chances de réussir à l'école* ». Plus précisément, les statisticiens de l'OCDE affirment que près de 40 % des élèves issus de milieux défavorisés sont en difficulté.

Par ailleurs, la Direction de l'évaluation, de la prospective et de la performance (DEPP, 2017) souligne qu'un autre groupe social pâtit d'une asymétrie de réussite dans le système scolaire français : il s'agit des filles. En effet, les statistiques de la DEPP montrent que peu de filles suivent les classes préparatoires scientifiques, peu de filles sont inscrites dans les filières relatives à la production ou à la technologie : elles représentent 29 % du nombre d'étudiants obtenant un diplôme d'ingénieur, 40 % parmi les docteurs en sciences. Et dans l'ensemble des pays de l'OCDE, les scores moyens en culture scientifique sont supérieurs pour les garçons.

En fait, sur la question de l'équité du système éducatif, on peut dire que deux groupes sociaux semblent particulièrement touchés chacun à leur manière : les filles, particulièrement sous-représentées dans les filières puis les carrières scientifiques à forte valorisation sociale et les enfants issus de milieux socioéconomiques défavorisés qui sortent plus rapidement et plus souvent sans diplôme de ce même système éducatif.

Ces différentes données régulièrement rapportées, maintes fois exposées nous invitent à penser que l'appartenance à certains groupes sociaux détermine la longueur et/ou la teneur des destins scolaires, des destins universitaires. Dès lors, on comprend fort bien le lien causal

qui peut apparaître « logique » entre ces différents paramètres, surtout dans une société où la réussite scolaire, synonyme de réussite sociale, occupe une place primordiale. Mais comment expliquer de telles inégalités de réussite ? Que dit la recherche ?

## **2. Comment expliquer ces différences de réussite ?**

La littérature scientifique est extrêmement abondante sur ces questions et propose plusieurs orientations théoriques. En effet, l'intérêt pour les inégalités de réussite dans le système scolaire est manifeste, l'étude de la littérature scientifique sur ce thème révèle un nombre très important de travaux. À titre indicatif, l'expression « inégalités scolaires » dans la base bibliographique internationale INIST en sciences humaines et sociales renvoie à 7329 références. C'est dire si la thématique intéresse les chercheurs, et ce, depuis plus de cinquante ans.

Dans cet article, nous n'envisageons pas de faire un inventaire de l'ensemble de ces travaux. Nous souhaitons exposer une perspective de recherche qui élucide les mécanismes en jeu en évitant de mobiliser des arguments mettant en cause des facteurs intrinsèques aux élèves qui pâtissent de ces inégalités scolaires.

Au demeurant, il est important de signaler que de tels arguments sont souvent invoqués. Pour expliquer réussites ou échecs scolaires, les capacités des élèves sont souvent mises en cause. Selon ce raisonnement -que nous estimons non suffisant pour expliquer ce qui se passe en classe-, ce sont les caractéristiques de l'élève qui détermineraient leur plus ou moins grande réussite dans une tâche. Dans cette perspective, nous serions conduits à penser que lorsqu'un élève réussit, c'est « tout simplement » qu'il en est capable et lorsqu'un autre élève ne réussit pas la même tâche, c'est en raison de ses capacités qui, en l'occurrence, lui font défaut (Huguet, 2005, 2006 ; Croizet & Neuville, 2004). En cas de difficulté, il convient alors de remédier à ce manque de capacités de diverses manières : en ajoutant des heures d'enseignement ou bien en proposant l'intervention d'un spécialiste par exemple. Il est intéressant de noter que cette relation, souvent pensée comme causale, entre réussite et capacités des apprenants est largement entretenue et illustrée dans le quotidien par de nombreuses expressions familières (p.e. « Elle est douée » ; « Quel potentiel ! » ; « il a moins de capacités que sa sœur » ... ; « il en a sous le pied ! »...). À ce sujet, l'ouvrage de Millet et Croizet (2016) étudie avec minutie ces phénomènes, et ce, dès l'école maternelle. Cette explication présentant la réussite scolaire et plus largement universitaire comme strictement dépendante des capacités nous aide à comprendre comment les situations d'évaluation peuvent être source d'anxiété ou de pression pour certains élèves...

Néanmoins, tout un pan de la littérature scientifique ne s'inscrit absolument pas dans cette perspective. Il s'agit des travaux conduits en psychologie sociale de l'éducation et sciences de l'éducation qui montrent comment la situation évaluative dans laquelle l'élève met en œuvre ses compétences est déterminante. Ces travaux révèlent que les contextes évaluatifs sont en mesure de réguler les performances de certains élèves, c'est-à-dire qu'ils sont susceptibles de les faire réussir dans certains cas et de les faire échouer dans d'autres. Plusieurs recherches montrent que le contexte de traitement de la tâche est déterminant pour l'activité cognitive de l'élève. Ces recherches ont identifié des mécanismes psychosociaux responsables des différences de performances. Elles soulignent l'intérêt de chercher à saisir les dynamiques psychosociales dans leur dimension temporelle et historique. Et ce, afin de ne pas se cantonner à une « épistémologie strictement cognitive », qui réduirait la mise en œuvre des performances et les apprentissages des élèves à « une relation tendanciellement bijective, où les propriétés intrinsèques des sujets et les propriétés intrinsèques des objets établiraient des

correspondances en dehors des instances et dynamiques sociales où le sujet est situé » (Huguet, 1992, p.218). Une telle perspective montre l'importance d'étudier l'impact cognitif des catégorisations et des conceptions des différents acteurs inscrits dans des situations d'enseignement et d'apprentissage : catégorisations sociales véhiculées par les contextes (p.e. dispositifs coopératifs, situations d'évaluation), catégorisations liées à l'histoire des sujets (échec ou réussite scolaire), catégorisations liées au sexe, à l'appartenance socioéconomique, aux conceptions de l'intelligence des différents acteurs de la situation ou bien aux buts d'apprentissage des élèves (Toczek, 2009).

En d'autres termes, ces travaux révèlent que la mobilisation de normes et valeurs intériorisées par les élèves ou par les enseignants au fil de leur histoire interfère avec le traitement cognitif de la situation d'enseignement dans laquelle ils sont insérés.

C'est précisément dans cette perspective que cet article présente des résultats de recherche permettant de comprendre les mécanismes en jeu. Les travaux présentés ici sont réalisés au sein de la cellule de base du système éducatif : la classe. En se fixant l'objectif d'étudier en quoi et comment le contexte même de la classe peut être en mesure de réguler le traitement de l'information des élèves (Toczek & Martinot, 2004 ; Dutrévis, Toczek, Buchs, 2015), ces travaux montrent que la classe est un lieu de multiples évaluations, positives ou négatives (Huguet, 2004 ; 2011). Evaluations à partir desquelles les élèves se comparent les uns avec les autres. Or, ces comparaisons ont des effets aussi bien, sur les états émotionnels et affectifs que sur les fonctionnements cognitifs et donc sur les performances des apprenants. La classe réunit les conditions d'émergence de représentations de soi et de comportements qui, pour une part au moins, sont le produit résumé de ce qu'une culture choisit d'honorer, de récompenser et de cultiver. Dans notre culture, et plus généralement dans les cultures nord-américaines et d'Europe de l'Ouest, l'emprise de la compétition est telle qu'elle contraint les individus à se comparer entre eux en permanence (Huguet, 2004; 2011). Dès lors, en prenant en considération le contexte de traitement de l'apprenant, ces études vont échapper à une psychologisation trop tentante ... Et ce, en interrogeant les mécanismes responsables des écarts de performances.

### **3. Les effets de situations évaluatives**

Dans cet article, il est important de préciser que lorsque nous parlons de situations d'évaluation, nous n'englobons bien entendu pas toutes les formes d'évaluation, ni toutes les logiques évaluatives. La situation d'évaluation se réfère à la fonction sociale de contrôle développée par Vial (2008) ou bien à une évaluation certificative (Allal, 1991). En d'autres termes, les situations évaluatives dont il est question ici ont pour objectif de vérifier la conformité d'une production d'élève selon une attente institutionnelle d'apprentissage.

### **4. Des contextes aux effets délétères**

Inscrites dans la perspective théorique exposée précédemment, plusieurs recherches développent les effets néfastes des préjugés de la réussite. En effet, nos sociétés vouent un véritable culte à la réussite (Croizet & Neuville, 2004 ; Croizet, 2011 ; Millet & Croizet, 2016). Nous devons réussir, nous devons tout réussir en permanence... aussi bien dans nos sphères de vie privée que professionnelle. Et l'école, comme d'autres sphères de la société, voue ce même culte à la réussite. Les élèves sont parfaitement clairvoyants de cette valeur rattachée à la réussite et à sa signification du point de vue de la connaissance de soi. Croizet (2011) explique fort bien qu'un concept est très lié à ce culte de la réussite, un concept « magique » : il s'agit du concept d'intelligence. L'intelligence est une qualité dotée d'une grande valeur sur

le plan social : lorsque l'on est intelligent ... on réussit. Ce concept exerce véritablement une fascination extraordinaire (Croizet, 2011). Dans la vie quotidienne, « *être intelligent, c'est Réussir* » ! C'est donc avoir des capacités ! Ce raccourci de pensée est très souvent activé, de manière implicite, à l'école et plus largement dans la société. De plus, dans cette même société, l'intelligence est vue comme une disposition stable. Or, cette conception est erronée et dangereuse à l'école. Selon Dweck (1999), c'est la conception que les individus ont de l'intelligence en général et de leur intelligence en particulier qui détermine dans une large mesure leurs performances intellectuelles.

A ce sujet, Souchal et Toczek (2010) montrent que les enfants issus de milieux socioéconomiques de faible statut ont tendance à considérer l'intelligence comme une caractéristique plus stable que leurs camarades issus de milieux socioéconomiques plus élevés. Cette étude se révèle être fort intéressante dans la mesure où elle montre que lorsque les élèves d'origine sociale différente sont placés dans un contexte de classe ordinaire, ils présentent des profils motivationnels différents en fonction de leur origine sociale. Les enfants montrent qu'ils ont forgé des représentations de l'intelligence et de la réussite qui diffèrent en fonction de leur appartenance sociale.

Afin d'illustrer cette perspective théorique, nous allons exposer plusieurs contextes évaluatifs régulateurs de performances. Plus précisément, les travaux qui vont être exposés dans un premier temps considèrent que le contexte situationnel de classe peut être, dans certaines conditions, vecteur d'une menace. Cette menace porterait soit sur l'identité personnelle des individus (Monteil & Huguet, 1991 ; 2001 ; 2002 ; Huguet, Brunot, Monteil, 2001), soit sur leur identité sociale (Steele & Aronson, 1995 ; Croizet & Claire, 1998 ; Spencer, Steele & Quinn, 1999) mais elle aurait dans tous les cas les mêmes conséquences délétères sur les comportements et les cognitions des individus, et donc sur leurs performances (Croizet, Désert, Dutrévis & Leyens, 2003).

Nous savons que les enfants possèdent des connaissances sur eux, notamment sur eux en tant qu'élèves, en partie construites à partir d'épisodes de leurs expériences antérieures et stockées en mémoire autobiographique. De telles connaissances appartiennent au concept de soi ainsi qu'à un sentiment particulier appelé sentiment d'efficacité personnelle qui désigne la confiance que les individus ont dans leur capacité à réussir une tâche donnée. Ces connaissances forment, selon Monteil et Huguet (2002) le contexte interne de l'élève. Or, ces auteurs montrent que, parfois, ce contexte interne peut entrer en résonance avec la situation que vit l'individu, situation alors appelée contexte externe. Selon Monteil et Huguet, en effet, il peut arriver que certains éléments de la situation agissent comme des indices permettant de récupérer en mémoire les expériences plus ou moins couronnées de succès – et donc émotionnellement positives ou négatives – que l'individu a déjà vécues dans cette même situation. C'est de cette interaction entre contextes interne et externe que naîtraient des conséquences cognitives pesant plus ou moins sur les performances.

Monteil et ses collaborateurs (Monteil & Huguet, 1991 ; 2001 ; 2002 ; Huguet, Brunot, Monteil, 2001) utilisent pour étudier cet effet une tâche particulière tirée d'un test utilisé en psychométrie et utilisant la figure de Rey-Osterrieth. Ce test s'effectue en deux temps : d'abord, le participant doit mémoriser la figure de Rey composée de plusieurs figures géométriques entrelacées, puis il doit la reproduire de mémoire. Sa performance peut ensuite être mesurée en fonction d'un barème qui alloue un certain nombre de points pour chacun des éléments de la figure initiale présent sur la reproduction effectuée par le participant. Ce barème tient également compte de la fidélité de la forme et de l'emplacement de chacun des éléments. Les participants des études citées étaient de jeunes adolescents âgés de 10 à 15 ans

et sélectionnés sur la base de leur passé scolaire de « bons élèves » ou au contraire de « mauvais élèves » en mathématiques et plus particulièrement en géométrie. Les chercheurs disposaient donc d'élèves ayant un contexte interne relié plutôt à des expériences de réussite dans le domaine de la géométrie et d'élèves jouissant dans ce même domaine d'un contexte interne plutôt relié à des expériences négatives. Les chercheurs ont ensuite manipulé le contexte externe des participants en aménageant la situation dans laquelle ils se trouvaient de façon à ce que celle-ci comporte – ou non – des indices activant ce contexte interne dissemblable. Ainsi, pour la moitié des participants, la tâche à effectuer était présentée comme une tâche de dessin alors qu'elle était décrite à l'autre moitié comme une tâche de géométrie. Les résultats montrent que, lorsque la tâche est présentée comme une tâche de géométrie, donc lorsqu'elle active le contexte interne des élèves dans cette discipline, les « bons élèves » réussissent mieux que les élèves connus pour rencontrer des difficultés dans cette discipline. En revanche, lorsque la tâche est présentée comme une tâche de dessin, les deux groupes d'élèves obtiennent des performances comparables : les élèves « faibles » obtenant dans cette condition de meilleures performances que les élèves « faibles » de la condition « géométrie ». Pour expliquer ces résultats récurrents, les auteurs ont recours à l'idée de charge cognitive. Ainsi, ils expliquent que la situation potentiellement menaçante amènerait l'individu à focaliser son attention sur le soi et que ce mécanisme aurait pour effet de diminuer les ressources attentionnelles disponibles pour accomplir la tâche (Monteil, Brunot & Huguet, 1996 ; Brunot, Huguet & Monteil, 2000).

Nous nous trouvons ici dans une situation où des indices situationnels entrent en écho avec les connaissances sur soi précédemment construites par un individu d'après son vécu personnel, son passé d'élève. Cependant, cette étude a également été répliquée en s'intéressant à une réputation d'infériorité provenant non plus du passé scolaire, mais de l'appartenance à un groupe social : le même protocole a, en effet, été proposé à des garçons et à des filles d'école élémentaire. L'hypothèse centrale de ces études était que les filles puiseraient dans le contexte situationnel de géométrie des indices entrant en interaction avec leur connaissance du stéréotype stipulant leur faiblesse dans cette discipline (Huguet, Régner, 2007 ; 2009). Cette interaction ne devait pas se produire dans le contexte plus neutre de dessin ou, selon l'étude, de test mnésique. Les résultats obtenus vont effectivement dans ce sens puisque les filles, qui obtiennent de moins bonnes performances que les garçons lorsqu'elles pensent faire de la géométrie, égalent et même surpassent ces derniers lorsqu'elles pensent réaliser une épreuve de dessin ou de mémoire. Ces études, bien qu'utilisant le même protocole que les études précédentes, sont sous-tendues par une problématique légèrement différente. Elles s'insèrent dans le champ de l'étude du phénomène de menace du stéréotype.

## **5. Le phénomène de menace du stéréotype**

Le phénomène de menace du stéréotype a été mis en lumière et nommé ainsi pour la première fois en 1995 par Steele et Aronson. Ces auteurs s'intéressaient à la population noire américaine et s'étonnaient de constater les moindres performances systématiquement enregistrées par les étudiants appartenant à cette population comparativement à celles obtenues par les étudiants blancs. Plus précisément, ces auteurs se sont intéressés aux étudiants noirs d'universités prestigieuses, ayant donc atteint un niveau d'excellence académique indéniable et ne pouvant donc pas être considérés comme désidentifiés de ce domaine académique ou comme possédant des connaissances sur soi négatives dans le contexte scolaire. Dans l'étude réalisée, Steele et Aronson proposent à des étudiants noirs et blancs de l'université de Stanford de réaliser un test verbal dont seul l'habillement diffère entre les deux groupes expérimentaux. Tandis qu'une première moitié des participants croit réaliser

un test diagnostique des capacités intellectuelles, l'autre moitié est avertie que cette tâche est un simple exercice de laboratoire. L'hypothèse des chercheurs est que le contexte externe – la tâche présentée comme diagnostique des capacités intellectuelles – active la connaissance du stéréotype d'infériorité intellectuelle des participants noirs et génère alors des conséquences cognitives conduisant à la baisse de leurs performances uniquement dans cette situation. Les résultats obtenus confirment cette hypothèse : les étudiants noirs de la condition « test d'intelligence » enregistrent de moins bonnes performances que les étudiants blancs de cette même condition, mais également que tous les participants de la condition « test de laboratoire ». L'explication donnée par les chercheurs est que les individus appartenant à des groupes sociaux stigmatisés par un stéréotype d'infériorité ressentent, lorsqu'ils sont placés dans une situation où ils sont susceptibles de confirmer à leurs yeux ou aux yeux des autres le stéréotype, un inconfort psychologique qui se traduit par une perturbation de leurs fonctions cognitives et donc, entre autres, par une diminution de leurs performances (Steele, 1997 ; Steele & Aronson, 1995). En fait, ce qui semble menacer l'individu stigmatisé est le fait que l'obtention d'une mauvaise performance dans cette situation stéréotypée pourrait être expliquée uniquement en faisant appel au stéréotype négatif. L'individu ne se sent donc plus pris en compte comme une personne avec des caractéristiques propres, mais comme un exemplaire de son groupe d'appartenance. Désert, Croizet & Leyens (2002) parlent à ce titre de *dé-individuation*.

Cet effet, initialement obtenu auprès d'étudiants noirs stigmatisés sur le domaine de l'intelligence, a également été depuis étudié auprès de nombreux groupes sociaux. En fait, il semble que la menace du stéréotype puisse atteindre tout membre d'un groupe stigmatisé par un stéréotype négatif. Ainsi, le groupe des femmes et le stéréotype d'infériorité intellectuelle dans le domaine scientifique qui le stigmatise ont-ils été tout particulièrement étudiés. Les résultats montrent généralement que les femmes placées dans une situation pouvant être interprétée à la lumière de ce stéréotype, obtiennent de moins bonnes performances que les hommes et que les femmes et les hommes placés dans une condition plus neutre.

Les individus issus de groupes sociaux de faible statut socioéconomique sont également sensibles à cet effet de menace du stéréotype : Croizet et Claire (1998) montrent en effet que les performances de ces étudiants à un test verbal sont moindres que celles des autres participants de haut statut lorsqu'ils considèrent ce test comme diagnostique des capacités verbales et que leur appartenance groupale a été rendue saillante par des questions concernant la profession et le niveau d'études de leurs parents (voir aussi Spencer & Castano, 2007 ; Harrison, Stevens, Monty & Coakley, 2006 ; Désert, Préaux & Jund, 2009).

Les groupes constitués sur la base des filières d'études suivies sont également la cible de stéréotypes plus ou moins négatifs et donc c'est tout logiquement que leurs membres peuvent également être touchés par cet effet de menace. Ainsi, Croizet, Dutrévis & Désert (2002) montrent que les étudiants titulaires d'un baccalauréat technique sont menacés face à un test verbal. Ils y obtiennent en effet de plus mauvaises performances lorsqu'on le leur présente comme diagnostique des capacités intellectuelles que lorsqu'ils pensent que ce test concerne le fonctionnement mnésique. Comme dans les autres études, les performances des participants non stigmatisés – *i.e* les titulaires d'un baccalauréat général – ne sont pas impactées par le contexte de passation.

En somme ce phénomène de menace du stéréotype peut être défini comme la crainte de confirmer ou d'être vu comme confirmant un stéréotype négatif à son groupe d'appartenance – et donc d'être réduit à ce trait négatif. Il entraîne une pression évaluative sur l'individu et

cette pression peut être suffisamment forte pour perturber le fonctionnement cognitif « normal » de la personne.

A ce stade, nous pourrions dire que les travaux de recherche présentés jusqu'alors révèlent trois points fondamentaux : (1) les productions cognitives en situation ne sont pas réductibles aux seules capacités de leurs auteurs ; (2) les représentations construites au fil du temps par les élèves à propos d'eux-mêmes sont déterminantes ; (3) les situations de réussite pour les apprenants sont plutôt des situations sans caractère évaluatif. En effet, de tels travaux montrent le véritable pouvoir des contextes évaluatifs puisque certains paramètres de la situation sont à même de réguler l'activité cognitive des élèves, leurs performances, leur réussite. Les contextes évaluatifs ne semblent pas garantir l'équité entre les élèves, entre les étudiants. Dès lors, « logiquement », sommes-nous conduits à poser la question suivante : ne faudrait-il pas renoncer à l'évaluation puisque, lorsqu'un élève ou un étudiant est « victime » d'un stéréotype et qu'il entre dans une situation évaluative, il anticipe son échec et des doutes sur soi, sur sa valeur l'envahissent ainsi que des pensées perturbatrices ? Ce qui entraîne une chute des ressources cognitives et des performances. Il n'est évidemment pas possible de supprimer toute forme d'évaluation car les enseignants ont besoin de pouvoir rendre compte des progrès de leurs élèves. Néanmoins, comment aménager une situation évaluative pour la rendre moins menaçante ? Est-il possible de concevoir des situations évaluatives équitables ? Telle est la question de recherche posée par Smeding, Darnon, Toczek, Butera, Souchal, 2013 ; Souchal, Toczek, Darnon, Smeding, Butera, & Martinot, 2014). Pour traiter cette question, nous (Souchal & *al.*, 2014) avons examiné plusieurs travaux qui ont cherché à comprendre si les pratiques des enseignants au sein des classes pouvaient induire chez les élèves certains types de buts de réussite.

## 6. Pour concevoir des situations évaluatives équitables

De nombreux travaux ont identifié des « climats motivationnels » découlant de pratiques orientées vers la maîtrise des compétences (climat de maîtrise) en comparaison avec des pratiques orientées vers la production de performances (climat de compétition). Carol Ames et ses collaborateurs (Ames, 1992 a et b ; Ames & Ames, 1984 ; Ames & Archers, 1988 ; Ames & Maehr, 1989 ; voir aussi Maehr, 1991) ont essayé de décrire plus finement le type d'arrangement. De manière synthétique, on peut dire que le climat de maîtrise est décrit comme encourageant l'autonomie, la coopération entre les individus et orientant leur attention sur le processus d'apprentissage. Le climat de compétition serait, quant à lui, défini comme suscitant la compétition entre les individus, en accordant une grande importance aux notes obtenues. Ils les amèneraient donc à considérer le résultat final, la performance produite comme l'élément le plus important.

Le caractère plus ou moins favorable aux apprentissages de ces différents climats a également été investigué. Ainsi, on sait que le climat orienté vers la maîtrise est un climat qui est généralement relié à une plus grande poursuite de buts de maîtrise de la part des élèves (Anderman & Midgley, 1997 ; Kaplan & Maehr, 1999 ; Midgley & Urda, 1995) et a un plus haut niveau de motivation intrinsèque (Ntoumanis, 2001 ; Cury, Biddle, Famose, Goudas, Sarrazin, Durand, 1996). Ce climat semble également corrélé avec l'utilisation de stratégies d'apprentissage adaptées et efficaces (Ames & Archers ; 1988) ainsi qu'avec l'obtention de meilleures performances scolaires (Midgley & Urda, 2001) même si cet effet n'est pas toujours obtenu (Anderman & Midgley, 1997). Les effets du climat orienté vers la performance ne sont, quant à eux, pas aussi clairs : ils sont souvent associés à la poursuite de buts de performance (Ntoumanis, 2001), à l'emploi de stratégies inadaptées comme l'auto-handicap ou l'évitement à rechercher de l'aide (Midgley & Urda, 2001 ; Urda, Midgley &

Anderman, 1998) ou encore au manque d'engagement dans des tâches difficiles (Sarrazin, Famose, Cury, 1995). Le recours à la tricherie semble également associé à ce type de climat (Anderman, Griesinger, Werterfield, 1998). Cependant, concernant les performances des élèves, les études ne montrent généralement pas de lien entre ce contexte et les performances obtenues et il arrive également que ce climat puisse être relié à de bonnes performances scolaires (Midgley & Urdan, 1995).

On le voit donc, le type d'évaluation annoncé aux individus est susceptible de rendre plus ou moins saillants certains types de buts. En fait, les différents types de buts de réussite adoptés par les individus diffèrent par le standard de comparaison choisi par ces individus : alors que les buts de performance reposent sur de la comparaison sociale en posant les performances de l'individu par rapport aux autres, les buts de maîtrise, quant à eux, semblent plutôt reposer sur un processus de comparaison temporelle, situant les performances de l'individu par rapport à celles déjà observées (Ames, 1992 ; Butler, 1987 ; Kaplan & Maehr, 2007). Il n'est donc pas étonnant, qu'en fonction des types de comparaisons suggérés par les résultats des évaluations, les individus puissent modifier leurs buts de réussite.

C'est donc par ce biais que nous avons cherché à créer des contextes qui, tout en restant évaluatifs, pourraient présenter des caractéristiques non menaçantes pour les individus stigmatisés. Pour traiter cette question, nous avons choisi de nous focaliser sur une question relative à l'équité entre les sexes. Nous l'avons déjà évoqué en introduction, selon plusieurs enquêtes internationales, bien que les filles dépassent régulièrement les garçons à l'école (OCDE, 2011), les filles sont encore devancées par les garçons en sciences et en mathématiques. En effet, même si elles sont en mesure de combler l'écart de genre observé historiquement en mathématiques (Else-Quest, Hyde, & Linn, 2010, pour un examen), leurs scores sont en moyenne 12 points inférieurs aux scores des garçons dans de nombreux pays de l'OCDE. En fait, il semble que les filles sous-performent les garçons quand elles sont placées dans des situations qui activent le stéréotype négatif au sujet de leurs capacités scientifiques supposées (Huguet & Régner, 2007; Keller & Dauenheimer, 2003, Appel, Kronberger, Aronson, 2011). C'est le cas lorsqu'un test est présenté comme diagnostic de capacités. Il s'agit d'un problème important dans un contexte éducatif où l'évaluation (c'est-à-dire les situations impliquant des tests qui sont utilisés précisément parce qu'ils diagnostiquent les capacités des élèves) est une pratique très courante et une étape nécessaire pour l'apprentissage. Il est intéressant de noter que la recherche indique également que la performance des garçons bénéficie de tests de diagnostic par rapport aux situations non diagnostiques (Walter & Cohen, 2003). Est-il possible, ensuite, d'utiliser l'évaluation à l'école mais d'une manière qui ne nuise ni à la performance des filles ni à celles des garçons ?

## **7. Une étude empirique originale**

Dans cette étude, nous soutenons que la composante menaçante de l'évaluation est due au fait que l'évaluation est principalement utilisée comme outil pour les objectifs de sélection et de performance. Les différences de genre ne devraient pas apparaître lorsque l'évaluation se concentre sur les objectifs de maîtrise. Plus précisément, nous avons tenté de mettre en place une situation évaluative se déroulant dans des contextes reprenant des caractéristiques du climat de maîtrise afin de protéger les filles de la menace potentiellement ressentie. En d'autres termes, l'idée était de créer des contextes évaluatifs non orientés vers la comparaison sociale, mais plutôt vers la comparaison temporelle, afin de protéger cette population.

Nous avons eu le souci de travailler dans des conditions écologiques maximales. Ainsi, afin de tenir compte du climat scolaire dans lequel les élèves évoluent depuis le collège -et peut-

être depuis l'école élémentaire-, hautement marqué par l'attribution de notes et afin de conserver un bon niveau de crédibilité aux conditions expérimentales évaluatives, il a été choisi de ne pas modifier celles-ci : une note a donc, dans tous les cas, été attribuée, note devant compter dans la moyenne et devant explicitement « évaluer » les compétences des élèves. Simplement, c'est l'objectif annoncé de l'évaluation qui a été manipulé. En effet, selon Ames (1992, p. 264, notre traduction) « La question ne concerne pas tant comment les étudiants sont évalués, mais, plus important, comment ils perçoivent la signification de cette évaluation. ». L'idée ici n'était donc pas de manipuler le type de feed-back distribué mais plutôt l'idée que s'en faisaient les participants et cela en donnant le même type de renforcements à tous, mais en renforçant chez certains le côté normatif de l'évaluation et chez les autres le côté formatif.

Notre étude s'est déroulée à l'occasion d'un cours de physique-chimie face à des lycéens fréquentant des classes de seconde. Elle incluait trois conditions : un groupe n'était pas évalué, le second attendait une évaluation notée dont la finalité annoncée était de comparer les lycéens entre eux (évaluation orientée vers la performance) et le troisième groupe attendait lui aussi une évaluation notée mais dont la finalité était d'aider les élèves dans leur apprentissage (évaluation orientée vers la maîtrise). Concrètement, les participants étaient répartis en trois groupes expérimentaux qui différaient uniquement par le type d'évaluation attendu à l'issue du cours. Tandis que les membres du premier groupe ne pensaient pas être évalués, les élèves du second et du troisième groupe étaient prévenus qu'ils passeraient une évaluation. Entre ces deux groupes, c'est l'orientation de l'évaluation qui était manipulée : pour les uns, l'évaluation était orientée vers la performance car elle était censée servir à comparer les élèves les uns aux autres ; pour les autres, elle était orientée vers la maîtrise car on expliquait que l'évaluation aiderait les élèves dans leur apprentissage. Dans les deux cas, les élèves pensaient recevoir une note qui compterait dans leur moyenne trimestrielle de sciences physique-chimie.

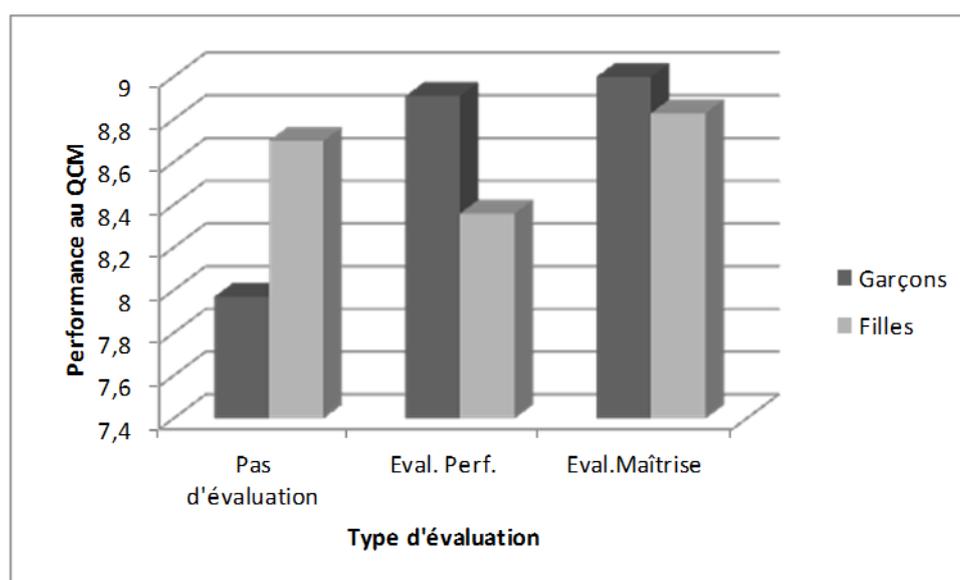
Deux objectifs étaient poursuivis par cette étude (Souchal & *al.*, 2014) : observer les effets des trois conditions évaluatives sur la performance des garçons et des filles et observer l'effet potentiellement bénéfique de la condition évaluative orientée vers la maîtrise pour tous les participants. Plus précisément, nous comptons tout d'abord obtenir au niveau des performances un effet de menace du stéréotype tout en restant dans des conditions les plus proches possibles de celles rencontrées habituellement lors de situations de classe normale, et ce, dans un domaine stigmatisant pour les filles, bien moins étudié que celui des mathématiques : la physique-chimie. Placées dans une situation évaluative indiquant que leurs compétences allaient être évaluées, les filles devaient se sentir menacées et donc obtenir de moindres performances que les garçons. Il était attendu que la situation évaluative orientée vers la maîtrise, quant à elle, permette aux jeunes filles de maintenir leurs performances. La situation non évaluative, par essence non diagnostique de leurs capacités, ne devait pas être menaçante pour elles et devait les amener à produire des performances aussi élevées que celles produites par les garçons.

Quant aux performances des garçons, s'il est connu qu'ils ne souffrent pas des situations évaluatives diagnostiques, on sait également qu'ils peuvent parfois même en tirer des bénéfices dans les domaines où ils sont ciblés par un stéréotype positif (Huguet & Régner, 2007; Seibt & Forster, 2004). La comparaison sociale n'est pas menaçante pour eux et peut même être stimulante. Covington (1992) note, par ailleurs, que les notes peuvent avoir un effet positif et stimulant sur les individus qui ont d'ordinaire de bonnes performances. Les garçons ont de plus une certaine préférence pour la compétition et les buts de performance proche (Duda, 1988). Tout cela amenait à penser que la situation non évaluative, ne

permettant aucune comparaison avec les autres, ne serait pas motivante pour eux. Nous pensions donc que les garçons obtiendraient leurs meilleures performances dans les deux situations évaluatives et que celles-ci chuteraient dans la condition non évaluative que nous pensions non motivante pour eux.

Filles et garçons devaient donc réussir de la même manière dans la condition évaluative orientée vers la maîtrise, tandis que la condition évaluative orientée vers la performance devait permettre aux garçons, mais pas aux filles d'avoir de bonnes performances et que la condition non évaluative devait permettre aux filles, mais pas aux garçons, d'avoir de bonnes performances.

Nos résultats sont présentés dans la figure ci-dessous.



**Figure 1 :** Performances obtenues au contrôle en fonction du sexe de l'élève et de la situation d'évaluation

Ces résultats sont intéressants, ils révèlent que les performances des garçons et des filles sont différentes dans les trois contextes évaluatifs. Les performances sont significativement différentes entre les filles et les garçons dans la condition « sans évaluation » et la condition « évaluation orientée vers la performance ». En revanche, il n'y a aucune différence significative dans la condition « évaluation orientée vers la maîtrise ».

Plus précisément, conformément à nos hypothèses, les lycéennes de notre étude obtiennent des performances égales, voire supérieures à celles obtenues par les garçons, uniquement dans des conditions non menaçantes pour elles. En revanche, lorsque l'évaluation est orientée vers la performance, c'est-à-dire lorsqu'on leur dit que leurs compétences sont mesurées et qu'elles seront comparées à celles des autres élèves, alors celles-ci chutent. Cet effet de menace du stéréotype, découvert tout d'abord par Steele et Aronson (1995) auprès des étudiants noirs stigmatisés sur le domaine de l'intelligence, a depuis été également mis en évidence chez les femmes dans le domaine des mathématiques (Spencer, Steele et Quinn, 1999 ; Ambady, Shih, Kim et Pettinski, 2001 ; Quinn et Spencer, 2001, Steele, 1997). Une minorité de travaux s'est intéressée à ce même phénomène au sein de disciplines connexes aux mathématiques où les femmes sont également stigmatisées : Smith, Morgan et White (2005) dans le domaine de l'informatique ; Bell, Spencer, Iserman & Logel (2003) en Ingénierie ; Smeding (2009) dans le domaine des mathématiques, sciences et ingénierie.

Enfin, citons Conway-Klaasen qui, dans ses travaux de thèse, s'est intéressée au domaine de la chimie, sans toutefois arriver à induire un effet de menace du stéréotype complet chez ses participantes féminines. Notre étude élargit donc la littérature relative au phénomène de menace du stéréotype puisque nos résultats l'étendent à un domaine peu étudié et pourtant stigmatisant pour les filles : les sciences physique – chimie.

Il est également intéressant de noter que les garçons, quant à eux, obtiennent de moins bonnes performances uniquement lorsqu'ils ne se sentent pas évalués. Ce résultat est conforme à celui que nous attendions. Ce résultat va dans le sens des études qui formulent l'hypothèse que la motivation des individus est à son paroxysme quand ils rencontrent un contexte dans lequel le but d'accomplissement saillant est congruent avec leur type de but préférenciel (Duda, 2001; Harackiewicz, Barron et Elliot, 1998 ; Standage, Duda et Ntoumanis, 2003). On sait que les garçons sont traditionnellement plus attirés par les buts de performance approche que les filles (Duda, 1988), ce que l'on observe également dans la présente étude. Les garçons seraient donc plus tournés vers les situations de compétition que les filles ; il est donc normal que leurs performances chutent lorsque le contexte évaluatif ne met pas du tout ce type de buts en jeu, comme dans la situation non évaluative de cette étude.

Conformément à nos hypothèses, la situation évaluative orientée vers la maîtrise est la seule des trois conditions testées qui permet à l'ensemble des participants de préserver leurs performances. Nous pensons que cette condition serait favorable à tous les participants car elle leur permettrait d'activer des buts de maîtrise, eux-mêmes favorables aux bonnes performances.

Même si cette étude présente encore de nombreuses zones d'ombre sur les mécanismes en jeu, elle permet d'ores-et-déjà d'identifier des situations comme étant plus ou moins favorables à la production de performance des femmes dans le domaine scientifique et semble-t-il, plus largement des groupes stigmatisés dans le domaine sur lequel porte le stéréotype. Elle permet également d'apporter des arguments en faveur des enseignants qui rapportent souvent naïvement qu'une évaluation sanctionnant les apprentissages est souvent nécessaire pour que certaines catégories d'élèves s'investissent dans les tâches proposées.

Elle permet également de désigner le climat motivationnel orienté vers la maîtrise comme protecteur des performances des individus stigmatisés même lorsqu'ils savent, par ailleurs, que leurs compétences vont être évaluées et sanctionnées par une note. Cela lui confère donc un large intérêt écologique.

## **7. En guise de conclusion**

Nous pouvons dire que nous sommes désormais certains que les performances des élèves ne sont pas réductibles à des capacités qui seraient intrinsèques aux sujets. Les performances cognitives semblent bien indissociables de leurs contextes sociaux et culturels de production.

Il est également important de noter pour conclure que les résultats exposés ici, concernant les effets des buts de maîtrise, sont consistants d'un point de vue scientifique (p.e., Smeding, Darnon, Toczek, Butera, Souchal, 2013; pour une synthèse, voir Darnon, 2014). La mobilisation d'objectifs de maîtrise lors de situations d'évaluation permet de réduire, voire de faire disparaître les écarts de performance qui préexistent. En outre, notons que ces résultats sont parfaitement cohérents avec les textes officiels du ministère français de l'Éducation nationale, notamment avec le nouveau décret du 31 décembre 2015 -JORF n°0002 du 3 janvier 2016-. Une évaluation bienveillante et exigeante, une évaluation positive

qui permet de rendre compte des progrès des élèves peut facilement être pensée et mise en œuvre en facilitant l'adoption de buts de maîtrise.

En somme, ce vingt-huitième colloque de l'ADMEE - Europe, en choisissant d'encourager, de souligner, de mettre en lumière des travaux à l'intersection de ces deux concepts « Evaluations et Apprentissages » nous a invités à une très belle note d'optimisme ... puisque les situations évaluatives caractérisées par des buts d'apprentissage se présentent comme une piste d'action éducative prometteuse aux problèmes récurrents de l'inégalité de réussite au sein des classes.

## 8. Références

- Allal, L. (1991). *Vers une pratique de l'évaluation formative : matériel de formation continue des enseignants*. Bruxelles : De Boeck-Wesmael
- Ambady, N., Shih, M., Kim, A. & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 12, 385-390.
- Ames, C. (1992a). Achievement goals and the classroom motivational climate. In D. H. Schunk & J. L. Meece (dir.), *Student perceptions in the classroom* (p.327-347). NJ : Erlbaum.
- Ames, C. (1992b). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261-271
- Ames, C. & Ames, R. (1984). System of student and teacher motivation: Toward a qualitative definition. *Journal of Educational Psychology*, 76, 535-556.
- Ames, C. & Archer, J. (1988). Achievement goals in the classroom: Student's learning strategies and motivation process. *Journal of Educational Psychology*, 80, 260-267.
- Ames, C.A. & Maehr, M. (1989). *Home and school corporation in social and motivational development*. Research funded by the Office of Special Education and Rehabilitative Services. Technical Report.
- Anderman, E. & Midgley, C. (1997). Changes in personal achievement goals and the perceived classroom goal structures across the transition to middle level schools. *Contemporary Educational Psychology*, 22, 269-298.
- Anderman, E.M., Griesinger, T. & Westerfield, G. (1998). Motivation and cheating during early adolescence. *Journal of Educational Psychology*, 90, 84-93.
- Appel, M., Kronberger, N. & Aronson, J. (2011). Stereotype threat impairs ability building: Effects on test preparation among women in science and technology. *European Journal of Social Psychology*, 41, 904-913. doi: 10.1002/ejsp.835
- Bell, A.E., Spencer, S.J., Iserman, E. & Logel, C.E.R. (2003). Stereotype threat and women's performance in Engineering. *Journal of Engineering Education*, 92, 307-312.
- Brunot, S., Huguet, P. & Monteil, J.-M. (2000). Performance feedback and self-focused attention in the classroom: When past and present interact. *Social Psychology of Education*, 3, 277-293.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79, 474-482.
- Conway-Klaassen, J.M. (2010). *Stereotype threat's effect on women's achievement in chemistry: The interaction of achievement goal orientation for women in science majors* (Thèse de doctorat, Université de Nevada, Las Vegas). Récupéré le 5 juillet 2011 de <http://digitalscholarship.unlv.edu/thesesdissertations/858>.
- Covington, M.V. (1992). *Making the Grade: a self-worth perspective on motivation and school reform*. New York, NY: Cambridge University Press.
- Croizet, J. & Claire, T. (1998). Extending the concept of stereotype threat to social class: The Intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, 24, 588-594.

- Croizet, J.-C., Désert, M., Dutrevis, M. & Leyens, J.P. (2003). L'impact des réputations d'infériorité sur les performances intellectuelles. *Revue Internationale de Psychologie Sociale*, 16, 97-124.
- Croizet, J., Després, G., Gauzins, M., Huguet, P., Leyens, J. & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, 30, 721-731.
- Croizet, J. C. & Neuville, E. (2004). Lutter contre l'échec scolaire dans la classe en agissant sur les préjugés de la réussite. In Toczek, M.C. et Martinot, D. (Eds) *Le défi éducatif. Des situations pour réussir* (pp. 55 -82). Paris: Armand Colin.
- Croizet, J. C. (2011). Le racisme de l'intelligence, (pp. 110-128). In Butera, F., Buchs, C. et Darnon, C. (Eds) *L'évaluation, une menace*, Paris: PUF.
- Cury, F., Biddle, S., Famose, J.-P., Goudas, M., Sarrazin, P. & Durand, M. (1996). Personal and Situational Factors Influencing Intrinsic Interest of Adolescent Girls in School Physical Education: a structural equation modeling analysis. *Educational Psychology*, 16, 305-315.
- Darnon, C. (2014). *Faire sa place dans le système : Enjeux relationnels, statutaires et idéologiques de la motivation en contexte éducatif*. Habilitation à Diriger des travaux de Recherches soutenue publiquement le 26 novembre 2014. Sous la direction du Professeur Delphine Martinot, Université Blaise-Pascal, Clermont-Ferrand.
- Désert, M., Croizet, J.-C. & Leyens, J.-P. (2002). La menace du stéréotype: Une interaction entre situation et identité. *L'Année Psychologique*, 102, 555-576.
- Désert, M., Préaux, M. & Jund, R. (2009). So young and already victims of stereotype threat: Socio-economic status and performance of 6 to 9 years old children on Raven's progressive matrices. *European Journal of Psychology of Education*, 24, 207-218.
- Direction de l'évaluation, de la prospective et de la performance (2017). *Filles et garçons, sur le chemin de l'égalité de l'école à l'enseignement supérieur*. Ministère de l'éducation nationale de l'enseignement supérieur et de la recherche.
- Duda, J.L. (1988). The relationship between goal perspectives, persistence and behavioral intensity among male and female recreational sport participants. *Leisure Sciences*, 10, 95-110.
- Duda, J. L. (2001). Achievement goal research in sport: Pushing the boundaries and clarifying some misunderstandings. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 129-182). Champaign, IL: Human Kinetics.
- Dutrevis, M., Toczek M. C., & Buchs, C. (2015). Régulation des apprentissages scolaires, (pp. 69-94), in M. Crahay & M. Dutrevis (éd.), *Psychologie des apprentissages scolaires*, Louvain-La-Neuve, de Boeck.
- Dweck, C. S. (1999). *Self-theories and goals: Their role in motivation, personality, and development*. Philadelphia: Taylor & Francis/Psychology Press.
- Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103-127. doi: 10.1037/a0018053.
- Harackiewicz, J. M., Barron, K. E. & Elliot, A. J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist*, 33, 1-21.
- Harrison, L. A., Stevens, C. M., Monty, A. N. & Coakley, C. A. (2006). The consequences of stereotype threat on the academic performance of white and non-white lower income college students. *Social Psychology of Education*, 9, 341-357.
- Huguet, P. (1992). *Catégorisations, insertions sociales et performances cognitives: Approche expérimentale*. Thèse de doctorat nouveau régime, Université Blaise-Pascal, Clermont-Ferrand.
- Huguet, P., Brunot, S. & Monteil, J.-M. (2001). Geometry versus drawing: Changing the meaning of the task as a mean to change performances. *Social Psychology of Education*, 4, 219-234.
- Huguet, P. (2005). Bon ou mauvais élève ? Dans M. Fournier & V. Troger (Eds.), *Les mutations de l'école* (pp. 77-84). Éditions Sciences Humaines.
- Huguet, P. (2006). Apprendre en groupe : la classe dans sa réalité sociale et émotionnelle (pp. 153-167). In E. Bourgeois & G. Chapelle (Eds.), *Apprendre et faire apprendre*. Paris : PUF.

- Huguet, P., Régner, I. (2007). Stereotype threat among school girls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, *99*, 545-560.
- Huguet, P., & Régner, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *Journal of Experimental Social Psychology*, *45*, 1024-1027.
- Kaplan, A. & Maehr, M.L. (1999). Achievement goals and student well-being. *Contemporary educational psychology*, *24*, 330-358.
- Kaplan, A. & Maehr, M. L. (2007). The contribution and prospects of goal orientation theory. *Educational Psychology Review*, *19*, 141-187. doi: 10.1007/s10648-006-9012-5
- Keller, J. & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, *29*, 371-381.
- Maehr, M.L. (1991). The psychological environment of the school: A focus for school Leadership, (pp. 51-81). In P. Thurston & P. Zodiates (Eds.), *Advances in educational administration*, Greenwich, CT: JAI Press.
- Midgley, C. & Urdan, T. (1995). Predictors of middle school students' use of selfhandicapping strategies. *Journal of Early Adolescence*, *15*, 389-411.
- Millet, M. & Croizet, J. C. (2016). *L'école des incapables ? La maternelle, un apprentissage de la domination*, Paris : La Dispute.
- Monteil, J.-M., Brunot, S. & Huguet, P. (1996). Cognitive performance and attention in the classroom: An interaction between past and present academic experiences. *Journal of Educational Psychology*, *88*, 242-248.
- Monteil, J.-M. & Huguet, P. (1991). Insertion sociale, catégorisation sociale et activités cognitives. *Psychologie Française*, *36*, 35-46.
- Monteil, J.-M. & Huguet, P. (2001). The social regulation of classroom performances: A theoretical outline. *Social Psychology of Education*, *4*, 359-372.
- Monteil, J.-M. & Huguet, P. (2002). *Réussir ou échouer à l'école : une question de contexte ?* Grenoble : Presses Universitaires de Grenoble.
- Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *British Journal of Educational Psychology*, *71*, 225-242.
- Organisation for Economic Co-operation and Development (OECD) (2011). *Education at a Glance 2011: OECD Indicators*, OECD Publishing. doi: 10.1787/eag-2011-en
- Organisation for Economic Co-operation and Development (OECD) (2016). *PISA 2015. Results in Focus - OECD 2016*
- Quinn, D. M. & Spencer, S. J. (2001). The Interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57*, 55-71.
- Sarrazin, P., Famose, J.P. & Cury, F. (1995). But motivationnel, habileté perçue et sélection d'un niveau de difficulté d'une voie en escalade. *STAPS*, *38*, 49-61.
- Seibt, B. & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, *87*, 38-56.
- Smeding, A., Darnon, C., Toczek, M.C., Butera, F. & Souchal, C. (2013). Reducing the Socio-Economic Status Achievement Gap at University by Promoting Mastery-Oriented Assessment. *Plos One*, *August 2013*, *8*, e71678.
- Smith, J.L., Morgan, C.L. & White, P.H. (2005). Investigating a measure of computer technology domain identification: A tool for understanding gender-differences and stereotypes. *Educational and Psychological Measurement*, *65*, 336-355.
- Souchal, C. & Toczek, MC. (2010). Buts de réussite, différences de performances liées à l'appartenance socio-économique des élèves : de nouvelles hypothèses explicatives ? *Les sciences de l'Éducation pour l'ère nouvelle*, *43*(1), 13-35.
- Souchal, C., Toczek, M.-C., Darnon, C., Smeding, A. Butera, F. & Martinot, D. (2014). Assessing Does Not Mean Threatening: Assessment as a key determinant of girls' and boys' performance in a science class, *British Journal of Educational Psychology*, *84*(1), 125-136. doi: 10.1111/bjep. 12012.

- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Spencer, B. & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research*, 20, 418-432.
- Standage, M., Duda, J.L. & Ntoumanis, N. (2003). A model of contextual motivation in physical education: Using constructs from self-determination and achievement goal theories to predict physical activity intentions. *Journal of Educational Psychology*, 95, 97-110.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629.
- Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, 69, 797-811.
- Toczek, M.C. & Martinot, D. (2004). *Le défi éducatif. Des situations pour réussir*. Paris : Armand Colin.
- Toczek, M.C. (2009). *Les situations d'enseignement : un programme de recherches centré sur l'analyse des déterminants et des impacts*. Habilitation à Diriger des travaux de Recherches soutenue publiquement le 16 novembre 2009. Sous la direction du Professeur Marcel Crahay, Universités de Liège et de Genève.
- Urduan, T., Midgley, C. & Anderman, E. (1998). The role of classroom goal structure in students use of self-handicapping strategies. *American Educational Research Journal*, 35, 101-122.
- Vial, M. (2008). *Se former pour évaluer : se donner une problématique et élaborer des concepts*. Bruxelles : De Boeck.
- Walton, G.M. & Cohen, G.L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39, 456-467.



# Évaluation et curriculum : à la recherche d'une nouvelle signification de l'enseignement et de l'apprentissage<sup>i</sup>

**Maria do Céu Roldão**

Universidade Católica Portuguesa

[mrceuroldao@gmail.com](mailto:mrceuroldao@gmail.com)

**Pour citer cet article** : Roldão, M.d.C. (2017). Évaluation et curriculum : à la recherche d'une nouvelle signification de l'enseignement et de l'apprentissage. *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 37-50.

## 1. Introduction

La problématique développée dans le présent document s'inscrit dans la question épistémologique de la spécialisation de certains champs appartenant au domaine de l'éducation ou des sciences de l'éducation, selon la matrice épistémologique dans laquelle la discussion s'insère. Dans ce cas, nous prendrons comme référence principale le cadre des sciences de l'éducation, car il est prédominant dans les approches théoriques et de la recherche au Portugal, en grande partie d'influence francophone. Cependant, l'approche des mêmes questions dans le cadre conceptuel du domaine de connaissance désigné comme éducation, plus enracinée dans des contextes de production scientifique avec une matrice anglo-saxonne, apporte au débat, et ceci de différentes façons, les concepts du curriculum et de l'élaboration des curricula, dans lequel le savoir relatif à l'évaluation est mobilisé et s'intègre. Cette analyse cherche ainsi à faire converger ces deux macro-lentilles d'analyse, sans perdre de vue les différentes approches qui caractérisent, en vue de la discussion de la question qui nous concerne, la relation entre le champ de l'évaluation et celui du curriculum, au sein du savoir de l'éducation et dans le champ praxéologique de cette même éducation.

Le développement de la recherche en éducation a conduit à la spécialisation du champ théorique de l'évaluation, dans ses différentes dimensions et objets, avec des avantages évidents pour la construction de la connaissance. Cependant, l'effet collatéral de cette spécialisation se traduit par une certaine distanciation de l'évaluation, par rapport aux champs d'étude du curriculum et de l'élaboration des curricula, en raison de son autonomisation en tant que champ d'étude spécifique.

Ce processus de nature épistémologique reflète la tradition praxéologique du travail scolaire, historiquement construit et donc, organisé en deux rythmes temporels, celui de l'enseignement et celui de l'évaluation. L'adoption, par les systèmes et les enseignants, d'autres visions par rapport à ces champs, notamment par la vulgarisation, surtout discursive, du concept d'évaluation formative, n'a toutefois pas brisé cette structure dichotomique du modèle scolaire, en contradiction avec le concept même de régulation et de progression, associé à des perspectives et des études plus récentes dans le domaine de l'évaluation et des curricula.

La discussion porte ainsi, en mobilisant la recherche dans les champs concernés, sur les volets où il serait possible, nécessaire et fructueux de donner une nouvelle signification, avec une vision intégrée, à certains des concepts clés impliqués dans ces deux champs épistémologiques, en particulier l'apprentissage, l'enseignement, le développement des curricula et l'évaluation des apprentissages.

## **2. Cadre socio-historique et conceptuel de la relation curriculum-évaluation**

Le débat théorique, portant sur le champ des curricula, influe sur les agendas et les intérêts qui sont en jeu dans chaque société, prenant une nouvelle signification à chaque époque et dans chaque contexte. Le curriculum est ainsi constitué en tant que construction sociale et historique, donnant lieu à ce que certains auteurs désignent métaphoriquement comme « arène » (Apple, 1997), dans la mesure où celui-ci traduit la négociation possible, pas nécessairement équitable, des pouvoirs, des besoins, mais aussi des intérêts impliqués.

Ce débat/combat autour du curriculum est évident dans les tentatives successives de nouvelles prescriptions curriculaires, qui très souvent se chevauchent et sont parcellaires, contradictoires, générant des incohérences. Ces prescriptions sont largement responsables de certaines des difficultés que les enseignants et les écoles ressentent pour pouvoir faire une gestion plus complète de l'élaboration des curricula dans le contexte scolaire, où nous situons le déphasage apparent entre le curriculum et l'évaluation en matière de recherche et de pratiques scolaire et professionnelle.

Bien qu'il s'avère inapproprié de parler de « réformes curriculaires » au Portugal, au cours des dernières décennies, nous reprenons quand même le concept de réforme curriculaire comme un moyen d'établir un système cohérent des apprentissages à offrir et à garantir aux citoyens, reposant sur un ensemble organisé de principes, de buts, de domaines de connaissance et de développement, organisés en fonction de la progression respective, articulés horizontalement et verticalement comme le tout qu'ils sont censés être. Au niveau du curriculum, il convient de noter que seules deux mesures politiques ont correspondu, au XXe siècle, dans le contexte portugais, au desideratum de la « réforme » que nous venons d'énoncer, à savoir :

- la réforme de 1973, aussi connue sous le nom de Réforme Veiga Simão, nom du ministre qui l'a conçue et lancée, a constitué la transformation la plus profonde et la plus innovante du système éducatif archaïque, qui était alors en vigueur au Portugal (ME, 1972). Cependant, cette réforme a été interrompue, lors de sa formalisation et de son développement, par l'événement historique du 25 avril<sup>NT</sup> qui a naturellement généré des dynamiques nouvelles et imprévues ;
- en 1989/90, la réforme de l'éducation, menée par le ministre Roberto Carneiro, a justement réorganisé et donné un sens à l'ensemble de nombreux changements essentiels, mais parfois aussi chaotiques, qui ont marqué les années 1974 à 1986, date de la publication de la Loi de base qui est toujours en vigueur (ME, 1988, 1989).

Ces deux ministres portugais resteront dans l'histoire des curricula et de l'éducation en général. En effet, dans les deux cas, la mise en œuvre de ces réformes globales du système éducatif se doit à leur vision rare de la signification des curricula comme un élément structurant intrinsèque du système éducatif formel, et de la nécessité conséquente de le construire comme un tout, encadrant l'évaluation dans le continuum du processus de

---

<sup>NT</sup>La révolution des œillets

développement des curricula. Ainsi, ces deux réformes ont établi, de pair avec la réorganisation globale de la structure et des contenus des curricula, des lignes directrices pour les pratiques d'évaluation respectives, marquées, par exemple, par l'introduction de la conceptualisation théorique des modalités d'évaluation, en particulier leur dimension formative.

Le réajustement difficile du curriculum à une scolarisation universelle massifiée, mais dont les sociétés ne peuvent pas se passer de la qualité, a soulevé les conflits et les tensions qui se font encore ressentir au niveau des curricula, rendant nécessaire de souligner que c'est la garantie des apprentissages curriculaires qui légitime l'école en tant qu'institution éducative. Cependant, l'idée du curriculum qui est présente dans la mentalité éducative dominante, se manifestant au niveau de la décision macro, méso et micro, de l'administration aux écoles et aux enseignants, privilégie sa lecture comme une simple prescription, ne saisissant pas sa signification sociale et politique essentielle.

La réflexion sur le curriculum, développée dans le présent document, est organisée autour des questions matricielles qui définissent le champ du curriculum lui-même et qui, bien que souvent négligées, doivent être situées en amont d'un certain nombre d'autres problématiques qui traversent la réflexion sur l'école, son (in) succès et la pression exercée pour améliorer les apprentissages. La relation entre le développement des curricula, à travers l'action d'enseignement offerte par l'école, et la compréhension et l'utilisation du concept d'évaluation qui traversent ces mêmes pratiques et conceptualisations, est discutée dans ce cadre d'analyse. Nous avons également tenu compte du fait que le dispositif curriculaire est structuré au sein d'une organisation particulière, l'école, et d'un système éducatif dont les logiques, différentes dans des pays avec d'autres traditions, deviennent des contraintes puissantes.

### **3. Curriculum, en quoi et le pourquoi de l'apprentissage**

Un curriculum correspond, grosso modo, en tant que construction socio-historique, à l'ensemble organisé des apprentissages (connaissances de différents niveaux, domaines et types, valeurs, techniques, etc.) que cette société estime, à un moment et un contexte donnés, avoir besoin pour survivre et que chacun de ses membres a besoin de s'approprier pour s'y intégrer de façon satisfaisante (Roldão, 1999, 2010). Dans ce sens, il est constitué comme un *parcours d'apprentissage face aux objectifs et aux possibilités fournies* (Goodson, 1988 ; Zabalza, 1992).

À la suite de l'analyse que nous avons mise au point dans des travaux précédents et que nous suivons en partie dans ce document, l'école, en tant qu'institution publique après l'industrialisation, et créée sous ce format, ne rend visible et n'institutionnalise qu'aux XVIIIe et XIXe siècles, comme un canon partagé formellement, ce *curriculum* qui, auparavant, n'était pas toujours nommé ainsi<sup>1</sup>.

Le curriculum est le noyau définissant l'existence de l'école. L'école est constituée historiquement comme une institution, lorsque la nécessité sociale de transmettre un certain nombre de savoirs de façon systématique à un groupe ou à un secteur de cette société, est reconnue. Cet ensemble de savoirs à transmettre de façon systématique constitue le curriculum de l'école. Le contenu des curricula varie et continuera à varier, selon l'évolution

---

<sup>1</sup> L'expression curriculum apparaît la première fois dans le titre de l'œuvre de Franklin Bobbitt publiée en 1918 (trad 2004). Le concept est ancré dans plusieurs développements antérieurs, en particulier liés à l'idée de discipline au XVIe siècle.

des besoins et des pressions sociales et, par conséquent, des publics que cette action de l'école estime souhaitable d'atteindre.

La nature *de ce qui est contenu (le contenu)* dans le curriculum mérite donc d'être analysée de manière critique, en tenant compte des circonstances, des besoins et des publics actuels. Pour cette raison, nous devons percevoir le curriculum, dans cette réalité socialement construite qui caractérise l'école comme une institution propre à chaque époque, et nous devons abandonner cette vision naturaliste du curriculum en tant que figure stable des matières qui sont enseignées par l'école ces derniers temps : ces matières n'ont pas toujours été les mêmes et le modèle ne sera pas toujours identique. Un curriculum est constitué avec tous les éléments auxquels il est attribué une finalisation en termes *de nécessité et de fonctionnalité sociale et individuelle*, qui est ainsi instituée (Roldão, 1999, 18).

Il devient ainsi essentiel d'interpréter au préalable la façon dont les connaissances, aujourd'hui incarnées dans le curriculum énoncé correspondent à ce double desideratum : (1) assurer la durabilité des besoins d'entretien ou de croissance de la société à un moment et un contexte donnés, (2) attribuer à chacun l'accès à cette connaissance, sans laquelle son inclusion sociale, son pouvoir d'intervention comme citoyen et son développement en tant que personne seraient menacés.

La première dimension à clarifier dans la discussion sur le curriculum (qu'est-ce qui est important à apprendre ?) est la légitimation sociale qui le fait émerger et changer. Et cette légitimation est le résultat (1) de la reconnaissance sociale des connaissances impliquées dans le curriculum, nécessaires à la vie économique et sociale, (2) de son association avec les dimensions du pouvoir personnel et social effectif, ainsi que (3) de sa potentialité, aidant à l'intégration de chacun dans le tissu social. Ainsi, il varie et change, parce que ni les pouvoirs en présence ne sont statiques, ni les besoins immuables. Aujourd'hui, la prépondérance attribuée à des matières telles que les mathématiques ou la science n'est pas seulement une mode ou un point de vue de certains responsables techniques ou décideurs politiques, elle découle aussi d'une matrice économique de soutien technique et scientifique, de plus en plus claire dans les sociétés actuelles, mettant en arrière-plan les dimensions des sciences humaines, avec les pertes culturelles et sociales inhérentes. Mais les acteurs et les secteurs liés à ces aspects ont un rôle dans l'attribution d'un équilibre à cette tendance hégémonique, en utilisant le pouvoir social qu'ils détiennent très souvent.

La légitimation résulte donc d'une interaction complexe des forces qui est à la recherche, dans le consensus et les ruptures successives qu'elle traverse, d'affirmations « d'utilité sociale » contextuellement affectées à ces compétences. Dans ce cadre, d'une plasticité socio-historique, le curriculum est à la fois un garant, bien que temporaire, d'une possible stabilité, et un agrégateur social puissant qui cimente un canon culturel commun, un facteur de la structuration réelle des sociétés et de leurs identités et cultures.

La correspondance entre le curriculum et un canon culturel est aussi marquée par la transformation historique et sociale. Le canon était très uniforme au temps de la pré-universalisation de l'éducation et au temps des États-nations culturellement bien délimités, en supposant la proximité presque totale du curriculum avec la culture dominante et les groupes sociaux de niveau supérieur. Dans un monde mondialisé et multiculturel et dans les périodes après l'universalisation de l'éducation scolaire, la force de ces groupes dans la négociation des curricula et la centralité de la culture dominante et de ses codes sont toujours présentes. Cependant, le jeu se déroule aujourd'hui dans un terrain de coexistence des cultures et d'interaction entre des partenaires d'origines multiples (Roldão, 2016).

On peut alors se demander si le canon culturel est unique de nos jours ou s'il est multiple. Puis, ce canon étant dans la plupart des cas multiple mais marqué par une culture qui exerce l'hégémonie, comment situer alors la légitimité des curricula ? Cette dialectique est précisément une des zones de la fracture curriculaire d'aujourd'hui, la source de nombreuses fautes d'éthique dans l'éducation, allant de l'exclusion expresse et tacite<sup>2</sup> de ceux qui sont socialement plus défavorisés dans l'échelle sociale, à la diversification précoce des parcours curriculaires correspondant à des alternatives plus ou moins valorisées socialement ou à la mise en place de groupes de niveau non-transformatif au sein de l'organisation scolaire (Roldão, 2003a ; Sousa, 2010). En gardant à l'esprit que la légitimation du « quoi enseigner » est exprimée dans son « pour quoi » personnel et social, il faudrait garantir, dans les curricula de base, (1) soit l'accès de tous aux connaissances et au domaine des codes canon de la culture dominante, (2) soit la réception des savoirs d'autres matrices culturelles. Des pays, tels que le Canada et la Nouvelle-Zélande, développent de bonnes pratiques dans ce domaine depuis longtemps.

Pour prendre des décisions sur le choix objectif des contenus curriculaires, il faut encore considérer la forme adoptée par cette sélection, visant les fins de légitimation mentionnées. L'option la plus classique, car la plus ancienne dans l'histoire de l'école, est la version encyclopédique, mobilisant une logique perpétuellement additive, qui introduit dans le curriculum un caractère cumulatif et additif de tout le savoir disponible en format nécessairement simplifié. Elle est clairement inefficace dans la production de la vraie connaissance ainsi que d'une faisabilité improbable à long terme. L'adoption de cette vision a favorisé, à l'école, une mise en forme stérile de la connaissance, encapsulée dans des unités d'enseignement pleines de sujets énumératifs (des programmes éternellement longs et de plus en plus vastes, compte tenu de la croissance des savoirs) qui s'épuisent au moment de leur évaluation finale, avec une faible appropriation traduite en usages ou transferts, générateurs de gains de savoir et de compétences (Roldão, 2003a).

La structuration des contenus des curricula prescrits, qui n'est pas sensée être encyclopédiste et stérile, ira (1) favoriser l'appropriation des dimensions structurantes de chaque *matière de la connaissance*, titulaire d'une élaboration historique importante, de ses concepts et des codes d'accès respectifs ; (2) promouvoir l'équilibre entre *les savoirs fonctionnels et les savoirs scientifiques et culturels* ; (3) assurer l'appropriation des *processus* et des méthodologies qui permettent d'avoir accès et / de construire la connaissance ; (4) permettre l'apprentissage des matières (alias « résultats », mais qui ne sont pas seulement ceux des notations...) par les élèves, que l'action d'enseigner devrait promouvoir systématiquement, comme une *appropriation et une utilisation intelligente de la connaissance*, dans toutes ses dimensions.

#### **4. Développer le curriculum : l'art du « comment », les options de l'enseignement**

Le développement des curricula correspond au processus de transformation du curriculum énoncé en un curriculum en action. En d'autres termes, le *curriculum prescrit* est inscrit dans le *curriculum mis en œuvre* à l'école par l'action des enseignants, ne se matérialisant que lorsqu'il est approprié par chaque élève, devenant le *curriculum réel*. C'est justement dans le cadre du *curriculum mis en œuvre* et du *curriculum réel* que « l'expertise » professionnelle de l'enseignant (Marcelo, 2009) et l'efficacité des buts, assignés aux contenus stabilisés dans le curriculum

---

<sup>2</sup> Rui Canário (2005) fait référence à l'exclusion *interne*, soit à l'école post-massification où elle a lieu, étant donné la façon dont l'institution, tout en accueillant tout le monde, fonctionne, c'est-à-dire, selon des formats organisationnels qui, au début, conduisent une partie importante de leurs élèves à l'échec.

prescrit, sont mises en jeu, en ce qui concerne les apprentissages effectifs réussis au niveau du curriculum réel.

Les phases de l'élaboration des curricula, qui illustrent ce « comment » du curriculum, ont été largement systématisées par Gaspar et Roldão (2007, pp 78-80) dont l'approche est ici reprise. La vision de ces auteurs est basée sur l'hypothèse que toute position théorique concernant le curriculum, en particulier l'adoption d'une rationalité pratique et critique, comme celle où l'analyse se situe, n'exclut pas, l'exigeant plutôt, l'opérationnalisation nécessairement technico-didactique qui, en toutes circonstances, donne corps et vie à la pratique du curriculum en situation. Dans cette perspective, les étapes classiques de l'élaboration des curricula sont mobilisées, étant les héritières de la tradition comportementaliste et néo-comportementaliste, mais analysées à partir d'une vision pratique et critique. C'est la logique employée pour organiser l'utilisation et l'achèvement de ce schéma d'opérationnalisation de l'élaboration des curricula qui bâtit les marques de la rationalité utilisée et pas le contraire. Il est ainsi adopté ce que nous désignons comme « dispositif commun » pour faire référence aux phases de l'élaboration des curricula qui sont par la suite systématisées :

- la conception du curriculum ,
- la mise en œuvre / opérationnalisation du curriculum ,
- l'évaluation du curriculum.

La phase de conception comprend l'analyse de la situation à laquelle le curriculum à construire se destine, une mise en place articulée des objectifs et des contenus et leur orientation vers les compétences à développer par leur biais.

La phase de mise en œuvre/opérationnalisation comprend la définition des stratégies d'action enseignante, qui correspondent à la construction stratégique des situations d'apprentissage conçues en fonction des élèves, et la façon d'évaluer prévue au cas où les objectifs d'apprentissage énoncés seraient atteints.

Enfin, la phase d'évaluation du curriculum implique, en plus de l'analyse des résultats d'évaluation de l'apprentissage obtenus dans la phase précédente, un processus d'examen/remise en question de tout le parcours réalisé, afin de s'interroger sur la pertinence et l'exactitude du développement de chaque phase et afin d'identifier et de reprendre les aspects qui ont contribué à l'échec de l'apprentissage visé : l'analyse de la situation a-t-elle été la plus correcte ? dans quelle mesure les objectifs et les contenus ont-ils été appropriés et ont-ils pris en compte la réalité des élèves ? est-ce qu'il y a eu des changements qui n'ont pas été pris en compte ? comment les stratégies étaient-elles ou n'étaient-elles pas appropriées aux objectifs et compétences visées ? les instruments et les activités d'évaluation correspondent-ils aux buts et aux stratégies travaillées ? Une telle évaluation ne s'épuise pas dans une fonction remédiable, mais elle suppose également que les situations réussies soient démantelées de façon critique afin de pouvoir identifier, situer et analyser les facteurs de succès et leur adéquation, dans le sens de les rentabiliser dans des situations futures.

Ce dispositif, qui naît des visions plus techniques du curriculum, peut toutefois être aussi constitué en une structure de base pour le développement des curricula reposant sur une matrice constructiviste. Les mêmes questions et les mêmes phases doivent être parcourues, mais suivant diverses logiques. La différence essentielle, en percevant le curriculum comme un plan à respecter (la rationalité technique), est son entière conception qui est préalable et déconnectée de la mise en œuvre et de l'évaluation. Celles-ci se produisent uniquement après la conception et ne la mettent pas en cause, ne cherchant qu'à améliorer son efficacité. Au

contraire, la conception tient compte de la réalité et des acteurs concrets qui y participent si une logique constructive est adoptée, souhaitant construire le curriculum comme un projet participatif, situé et analytique. Le processus de définition des objectifs et des stratégies est contextualisé et situé dans une réalité qui lui donne naissance, en interaction avec la conception initiale, pouvant et devant apporter des changements et des ajustements, dus à la réalité des élèves, et prenant en compte sa transformation tout au long du processus. L'évaluation, dans cette logique bien plus permanente et constructive, implique, à son tour, l'ajustement en permanence des procédures et la ré-conceptualisation des buts visés.

Le concept central, structurant un processus de développement des curricula, est donc son orientation stratégique, c'est-à-dire, l'organisation d'un parcours d'enseignement et d'apprentissage, guidé par des buts curriculaires clairs, devant être pensé comme le plus approprié à la population d'élèves en question, pour que ceux-ci puissent effectivement s'approprier des savoirs et devenir compétents dans différents domaines d'études. Un tel concept d'orientation stratégique suppose un processus d'autorégulation et d'inter-régulation permanent, permettant d'atteindre le succès souhaité. L'évaluation serait ainsi constituée comme son élément régulateur essentiel.

La réussite des curricula prend corps dans l'action d'enseigner et c'est sur cette même action que les interventions transformatives doivent avoir un accent particulier, selon la perspective stratégique de l'enseignement que nous défendons, en visant les questions liées au succès ou à l'échec des élèves de tous les milieux sociaux et culturels, qui définissent aujourd'hui les bénéficiaires du curriculum.

La conception de l'enseignement en tant qu'action stratégique continue présuppose l'abandon de la croyance simpliste dans le manichéisme pédagogique, les « bonnes » et les « mauvaises » méthodes, ou dans l'application des théories déterministes aux bénéficiaires, condamnés dès le départ, au statut de « bons » et de « mauvais » élèves. Ce qui est en jeu est d'abord le développement de l'action d'enseigner dans chaque contexte de façon « appropriée » à la situation, afin de maximiser le succès de celui qui apprend, face à chaque contenu du curriculum. Pour ce faire, il devient essentiel d'analyser, choisir, ajuster, différencier, pour pouvoir atteindre, avec une plus grande efficacité et une meilleure qualité, l'apprentissage curriculaire souhaité.

L'action d'enseigner est donc, et en soi-même, une action stratégique complète, guidée et régulée face au desideratum de l'accomplissement de l'apprentissage souhaité chez l'autre. L'enseignant n'a donc pas à se demander : « Comment vais-je organiser la présentation de ce contenu pour qu'il soit clair et perceptible » ? Mais, plutôt, « Comment puis-je concevoir et réaliser une ligne d'action (pouvant inclure la présentation du contenu, stratégiquement organisé et coordonné avec d'autres dispositifs), avec quelles tâches, quelles ressources, quelles étapes, pour faire en sorte que ces élèves apprennent précisément le contenu que je veux enseigner ? » (Roldão, 2009 : 57).

L'un des points critiques de l'inefficacité relative de l'enseignement dans le système portugais, selon les résultats d'études internationales, telles que l'enquête PISA, assez connue, où nous avons alors constaté des améliorations (OCDE, 2006, 2009, 2012, 2015), mais aussi observé dans les données des révisions de la recherche (Roldão *et al*, 2006), est l'incidence plus faible des pratiques d'enseignement dans le développement intentionnel des opérations cognitives plus complexes, en particulier celles qui exigent une pensée analytique. Cette fragilité de la pensée analytique, suggérée dans les activités d'apprentissage proposées aux élèves, peut être perçue dans de nombreux matériaux de travail et dans les manuels utilisés dans les écoles ou souvent, dans les tâches réalisées en classe, qui ont fréquemment tendance à se concentrer

sur la constatation au détriment, par exemple, des fondements soutenus ou de la formulation et de la preuve des hypothèses.

Les implications organisationnelles du travail curriculaire de l'école, élucidant l'initiative et les responsabilités clairement accrues de l'institution et des enseignants, sont relativement peu discutées au sein de la communauté des enseignants, en grande partie marquée par une culture normative et peu socialisée dans une culture de l'initiative, de l'autorégulation et de l'hétérorégulation. Généré au sein d'une institution de nature spécifiquement curriculaire, l'école, le curriculum est également, dans son *modus operandi* (qui lui donne réellement forme et régule son contenu et le mode de passage), une production organisationnelle, qui se conforme grandement et s'imbrique dans les mécanismes organisationnels que l'école configure (Roldão, 2000 ; 2003b).

Le débat sur les curricula autour du transfert tant proclamé des niveaux de prise de décision vers l'école et les enseignants manque d'une détermination claire de la place que les deux facteurs qui conditionnent absolument un tel transfert occupent : d'une part, la nature même des contenus des curricula et le travail lié à leur développement (c'est-à-dire, en *quoi* et *comment*, les méthodes didactiques, les modes dominants de la pratique enseignante, l'acte d'enseigner, et qu'enseigner, tels que vécus dans les écoles, opérationnalisent l'enseignement dans les dispositifs établis pour le travail des élèves) ; d'autre part, la structuration incontournable qui est donnée au curriculum par le squelette organisationnel de l'institution scolaire, étonnamment inchangée depuis près de deux siècles : l'organisation du temps, de l'espace, de la production scolaire, de l'acte, d'ailleurs des actes segmentaires, d'enseigner. (Alarcão, 2001 ; Roldão, 1999). Ceci est le nœud qui « lie » le curriculum à une organisation de l'école du XIXe siècle, mais « naturalisée » (des classes uniformes, la segmentation du savoir, le travail séparé de chaque enseignant, le rythme uniforme, l'annualité...), empêchant ou entravant les tentatives de transformation dans le sens de l'efficacité<sup>3</sup>.

La reprise au Portugal du débat politique sur la surcharge présumée des contenus et des matières dans le curriculum prescrit, semble s'amorcer, et positivement, visant à leur réduction ou diminution. Il est généralement admis qu'il y a une surcharge dans la prescription du curriculum. Mais, est-ce l'essence du problème ? En plus du cumul, de la contradiction et du remplacement à un rythme inadéquat des documents qui sont successivement prescrits, ces tentatives aideraient-elles, de fait, à clarifier ce qui est essentiel ? Il convient également de noter que le plus puissant agent de lecture curriculaire des documents officiels, quels qu'ils soient, se trouve dans les manuels, dans la mesure où ils sont constitués comme des définisseurs des options didactiques que les enseignants très souvent se dispensent de créer, suivant les manuels au lieu de les utiliser en tant que ressources, au service de leurs stratégies, décidées de façon différenciée et contextualisée.

Le curriculum est ainsi constitué dans cette lecture, ainsi que son développement approprié, le leadership et l'évaluation, reposant sur des modes de travail scolaire qui donnent forme à l'enseignement comme une action stratégique, devant être ajustée en permanence, améliorée

---

<sup>3</sup> Il convient de noter qu'il existe plusieurs situations qui rompent cette tendance : des projets d'intervention sont en cours, tels que, entre autres, le projet Fénix, engageant certains changements au niveau de la transformation de la structure « uniformisatrice » de l'organisation des curricula à l'école. (Vd. Matias Alves et Luisa Moreira, Projeto Fénix, 2011-2012).

De nombreuses pratiques d'organisation non « uniformisatrice » du travail ont également lieu, en particulier au 1er cycle. Elles sont développées de manière cohérente et inspirées en permanence par le Mouvement de l'école moderne.

et évaluée, comme la pierre de touche de l'amélioration des apprentissages, le dernier desideratum de l'école et le premier droit des citoyens.

## 5. Évaluation et curriculum, la nature curriculaire de l'évaluation

La relation entre l'évaluation et les curricula est placée dans ce cadre d'analyse en supposant, comme nous avons essayé de le justifier ci-dessus, la pleine intégration dans le processus de développement des curricula.

L'évaluation des apprentissages est encadrée donc par une double orientation théorique. D'une part, elle est constituée comme un champ d'élaboration des curricula, en ligne avec les théoriciens du curriculum, et avec un accent particulier sur la contribution des lignes de la matrice comportementaliste, très influentes dans l'éducation et la formation des enseignants dans les années 1970 et 1980 (Tyler, 1949 ; Taba, 1962 ; Kliebard, 1975). D'autre part, la genèse de l'évaluation comme un champ d'étude autonome émerge historiquement et, en partie de la docimologie, en tant que tentative de la mesurabilité des produits de l'apprentissage. La visibilité de l'évaluation dans sa forme certificative, c'est-à-dire à travers la moyenne de l'ensemble des notes de l'évaluation qui régit le progrès scolaire des élèves, explique sa position centrale dans le domaine de l'éducation, devenant peu à peu un champ de recherche spécifique au sein de la communauté scientifique de l'éducation (Alves, 2012 ; Allal, 1986, 2010 ; Fernandes, 2005).

Une telle séparation est également due à la logique immanente au modèle ou à la « grammaire » scolaire, qui est contraire à cette rationalité curriculaire : à l'école, le temps et le travail sont organisés autour de deux axes qui constituent deux temps, (1) le développement des contenus des programmes, « donner la matière » dans le jargon des enseignants, qui correspond à la grande partie du temps de cours, et (2) le temps des tâches d'évaluation sommative, en général considérées comme des tests, qui a traditionnellement lieu à la fin des trimestres et qui est orienté vers une moyenne - note finale, ponctuant les rythmes de la vie de l'organisation et la séquence du parcours scolaire de l'élève. De ce fait, une approche des domaines d'enseigner/apprendre et d'évaluer/noter est tendanciellement réalisée comme deux zones distinctes sur le plan conceptuel, mises au point à deux moments différents, ceci se projetant, à son tour, dans la séparation traditionnelle entre l'enseignement et l'évaluation, la renforçant (Roldão et Ferro, 2015).

La ligne théorique, issue des études liées au champ de l'élaboration des curricula, même si elle est amendée par des marques nucléaires, générées par les perspectives comportementalistes invoquées à partir des années 1950 et suivantes en tant qu'instruments des politiques tournées vers l'efficacité de l'éducation, constitue néanmoins le plus important générateur de l'articulation entre l'évaluation et l'éducation, en reconnaissant l'évaluation comme un élément intégrateur de la régulation du processus de développement des curricula. Sur la base strictement comportementaliste de Ralph Tyler (1949), ou de l'établissement par Hilda Taba (1962) des phases du *design* et de la mise en œuvre du curriculum, avec une régulation évaluative en permanence, ou de la pédagogie pour la maîtrise théorisée par Benjamin Bloom (Bloom *et al*, 1971), la notion d'évaluation établit une nouvelle conceptualisation. Les chercheurs suisses du groupe de Genève, avec l'introduction d'un travail abondant dans les années 1970 sur le concept d'évaluation formative (Allal, 1986 ; Perrenoud, 2003), ont contribué à la forte propagation du concept d'évaluation comme une régulation de l'apprentissage, traduite de différentes façons et en types différents d'actions évaluatives - diagnostique, formative, sommative, pronostique. De ce référentiel, la notion d'évaluation formative, comme une rupture de la séquence classique enseignement-tests, est

passée à la formation des enseignants et au discours et pratiques enseignantes. La nouveauté, à ce stade, du concept d'évaluation formative, a souvent été détournée dans les pratiques des écoles, comme si elle était synonyme d'adoption de plus de moments évaluatifs, de courte durée, interprétés comme une « évaluation formative », dans la plupart des cas avec une forte déviation conceptuelle, bien que la terminologie soit à présent installée.

Sur le plan politique, au Portugal, l'influence des perspectives comportementalistes des années 1970 et 1980 a obéi à la massification et était « prescriptive » d'une adoption de taxonomies pour l'évaluation, centrées sur une identification hiérarchique des opérations mentales nécessaires aux différentes étapes d'apprentissage, auxquelles l'enseignement devait se rapporter pour se guider et coter. Ces perspectives ont été intégrées dans les pratiques enseignantes, en particulier par le biais de la formation initiale et de la formation continue des enseignants au Portugal, dans les années 1970-1980. Elles étaient utilisées de façon faiblement réductrice, hautement quantifiée, prisonnière d'une rationalité technique aveugle, et d'une action pédagogique pauvre bien que techniquement raffinée, qui a déclenché de nombreuses résistances et réactions critiques (Roldão, 1983). Cette excessive technicisation a empêché une appropriation plus critique de la valeur de l'instrument analytique qui permettait la clarification et l'affinement des opérations mentales promues et nécessaires aux différents contenus des curricula, un aspect qui donnait, selon la perspective ici assumée, une valeur puissante pour améliorer l'enseignement et l'apprentissage. Ce conflit nous a fait assister, dans les années qui ont suivi, à un abandon non analysé de ce référentiel, soit quant à l'enseignement, soit quant à l'évaluation, qui demeurent ainsi des champs séparés, laissant l'évaluation des apprentissages, réalisée par les enseignants, dépourvue de repères de rigueur ou de toute matrice évaluative ou de développement des curricula qui pourrait l'intégrer.

Seule la production de tests d'évaluation finale au niveau national, à partir de la remise en place des examens externes et de la création d'un institut chargé de leur préparation (GAVE, IAVE- Ministère de l'éducation), s'est constituée comme un champ important de l'attention du public, à travers les examens, mais sans pouvoir témoigner d'une appropriation significative de ce référentiel d'évaluation, généralement avec des bases scientifiques solides, dans les instruments actuellement utilisés par les écoles. De plus, le domaine de l'élaboration des curricula, peu développé dans la tradition portugaise jusqu'aux années 1980, en raison aussi de l'impact de la formation de nombreux enseignants de l'enseignement supérieur dans des universités anglo-saxonnes (qui a caractérisé le *boom* de la formation supérieure au niveau de l'éducation au Portugal dans les années 1970 et 1980) entre dans les programmes de formation initiale des enseignants des universités et des écoles supérieures d'éducation. Mais, il est génériquement traité comme un apprentissage des techniques de planification, où l'étape de l'évaluation est moins bien travaillée que les autres. Certains plans d'études des cours de formation des enseignants comprenaient parfois des matières spécifiquement liées à l'évaluation. Mais, de par sa dissociation de la dimension didactique et de l'élaboration des curricula, l'évaluation est décidément reprise comme une fin en soi dans les pratiques des écoles.

Le renforcement de la gestion managériale et compétitive des dernières décennies a renforcé l'importance de l'évaluation des apprentissages, mais avec une perspective désintégré, non soutenue par un développement harmonieux des programmes, dans une poursuite aveugle de l'amélioration des résultats scolaires *en tant que tels*, de plus en plus perçus comme des buts en soi, selon une logique concurrentielle entre les écoles pour obtenir les meilleurs places aux *rankings*, sans prendre soin de l'amélioration corrélative de l'enseignement et de l'apprentissage pour tous, voire en encourageant une forte reprise des logiques d'exclusion des élèves qui réussissent moins bien. La concurrence découlant de ces politiques provoque

une exagération des dimensions de la préparation aux examens, renforçant une approche techniciste, dépourvue d'une connaissance pédagogique et curriculaire, et négligeant de plus en plus la fonction régulatrice de l'évaluation (Alves, 2012), la seule pouvant assurer efficacement l'amélioration réelle des processus d'enseignement, les promoteurs intentionnels de l'apprentissage, et par conséquent, des résultats.

## 6. Qu'est-ce qu'évaluer ?

Selon la perspective que nous défendons, enseigner signifie faire apprendre quelque chose à quelqu'un (Roldão, 2010, 2014). L'objet de cette action, qui définit la fonction professionnelle des enseignants, est l'apprentissage des élèves, créant ainsi une double transitivité : celui qui enseigne, enseigne un contenu et l'enseigne à certains individus inévitablement différents. La compétence à enseigner réside précisément dans cette étape de la promotion intentionnelle et stratégique de l'accès des apprenants à la connaissance à enseigner, inscrite dans le curriculum. Dans la perspective des auteurs, le concept de curriculum est compris comme l'ensemble organisé des apprentissages jugés nécessaires à tous, dans un temps et un contexte donné (Gaspar et Roldão, 2007).

C'est le développement de ce curriculum par l'enseignant qui facilite et favorise l'appropriation souhaitée des connaissances par les apprenants. Pour ce faire, l'enseignant a besoin d'avoir une connaissance professionnelle solide qui intègre, dans le sillage de la pensée de Lee Shulman, les dimensions de la connaissance du contenu, de la connaissance pédagogique et didactique du contenu, de la connaissance des curricula, de la connaissance de l'élève, de la connaissance du contexte. Ces dimensions convoquent et intègrent dans l'action pédagogique, à chaque acte professionnel précis, le fait de savoir enseigner (Shulman, 1987, Shulman & Shulman, 2004 ; Montero, 2005 ; Roldão, 2010, 2014).

Des études internationales ayant de l'impact mondial, telles que l'enquête PISA ou les documents structurants de la politique éducative basée sur la recherche (OCDE, 2005 ; UNESCO, 2015), insistent à plusieurs reprises sur le caractère central de la variable « mode d'enseignement de l'enseignant », en corrélation avec les meilleurs résultats des élèves, les variables appartenant au domaine économique, social et culturel étant contrôlées. Ces constatations impliquent un nouveau regard sur le mode d'enseignement, dans le sens des processus de développement des curricula mobilisés pour les différents élèves. Elles soulignent également l'importance centrale de la régulation et du feed-back dans ce processus en tant qu'évaluation régulatrice (Black & Wiliam, 1998 ; Hattie & Timperley, 2007), une dimension essentielle de la promotion effective de tout apprentissage, scolaire ou non scolaire.

L'évaluation est ainsi, selon Miguel Zabalza, l'une des phases fondamentales de l'élaboration des curricula et est centrée sur l'évaluation du processus et des résultats obtenus (Zabalza, 1992). Cette phase du processus de développement des curricula est le dispositif central de la régulation de tout le travail accompli dans l'éducation scolaire.

Cependant, l'évaluation scolaire se constitue, dans la culture des systèmes éducatifs, des écoles et des enseignants, en entité autonome avec sa propre vie et sa propre logique, qui interrompt et pervertit, dans une large mesure, l'intégrité du processus. Illustrant ce paradoxe, Zabalza donne comme exemple la situation suivante : « (...) j'arrive un jour en classe et je dis aux élèves que le lendemain « il n'y aura pas de cours » parce que nous allons faire une évaluation, que nous « reprenons » les cours normaux le jour d'après, alors là, je suis en train d'exprimer, en des termes clairs et précis, que l'évaluation à laquelle je procède implique un arrêt du fonctionnement normal des classes et que, en ce qui concerne le propre processus

éducatif, l'évaluation n'aura pas de conséquences directes (c'est à dire, les cours d'après-demain et les suivants, après l'évaluation, ne seront pas, en principe, différents des cours précédant l'évaluation et, l'étant, ce ne serait pas une conséquence de cette évaluation) ». (Zabalza, 1992, p. 224).

Cet exemple illustre le divorce persistant entre l'enseignement et l'évaluation que nous venons d'analyser. Il reflète l'absence d'une véritable culture curriculaire et exprime une perspective mécaniste de l'enseignement, non fondée sur l'analyse des processus d'enseignement-apprentissage scolaires en tant que processus de développement des curricula, reposant sur des critères de rigueur et de connaissance scientifique et clarificateurs de leur propre nature et dynamique. La recherche elle-même, exprimée dans l'organisation des cadres théoriques des sciences de l'éducation, reflète ce paradoxe par le développement d'un champ d'études dans le domaine spécifique de l'évaluation, déconnecté de son inclusion dans le processus, plus vaste, de l'élaboration des curricula. La spécificité et la complexité du processus et du concept d'évaluation permettent de comprendre que tout cela se produit dans le plan de la théorisation spécialisée, mais il est important de ne jamais perdre de vue que (1) l'évaluation fait partie intégrante de l'élaboration des curricula et qu' (2) elle est définie comme un processus et un système. Selon Zabalza (1992), « l'évaluation est un processus et se trouve dans un processus, c'est un système et se trouve dans un système » (p. 224). Il est important de comprendre qu'une telle pratique de la dissociation de l'évaluation à l'égard du processus d'enseignement correspond, dans la culture des écoles, à la superposition et souvent à la confusion entre les différentes fonctions que l'évaluation consolide, en ce qui concerne l'évaluation des apprentissages des élèves : la fonction de diagnostic ou d'analyse, la fonction régulatrice de la fonction certificative (Carrilho Ribeiro, 1990 ; Zabalza, 1992).

Il importe ainsi de déconstruire les pratiques d'évaluation les plus récentes, qui dissocient l'évaluation du processus d'enseignement, qui subsument l'évaluation dans la dimension certificative, avec l'omniprésence de la moyenne finale - notation, ne convoquant l'évaluation dite formative que comme une sorte de remède, ce qui est loin de sa signification théorique originale.

Toute discussion guidée pour améliorer les pratiques d'évaluation devra ainsi partir de deux lignes d'analyse :

1. Qu'est-ce que nous évaluons face à un ensemble de propositions ou de buts d'un curriculum donné ?
  - la clarification de l'objet (d'enseignement et d'évaluation)
2. Quelles sont les dimensions qui font partie de cet objet ?
  - la relation entre le contenu de la connaissance visée et les opérations cognitives d'acquérir / maîtriser / utiliser
3. Comment intégrer l'évaluation, dans son volet régulateur et sommatif, dans les processus de développement des curricula c ?
  - la clarification des processus d'enseignement et d'évaluation

Toute autre ligne d'analyse, telle que la discussion des instruments d'évaluation, la temporalité des moments d'évaluation et leur fréquence, les systèmes de notation, entre autres, ne peut se développer que comme résultat de ces deux dimensions - les contenus et les processus d'enseignement qui font l'objet de cette évaluation et que celle-ci reprend en tant que régulatrice.

Il importe donc de donner une nouvelle signification, dans une perspective curriculaire d'intégration, à ces deux axes de l'action d'enseigner, de façon à recréer, dans le curriculum et dans la recherche, une harmonisation qui est en grande partie absente, entraînant une perte de la qualité et de l'efficacité de l'éducation scolaire.

## 7. Références

- Alarcão, I. (2001). Escola reflexiva e a supervisão. Uma escola em desenvolvimento e aprendizagem. In I. Alarcão (org.), *Escola Reflexiva e Supervisão* (pp 12-23). Porto: Porto Editora.
- Allal, L. (1986). Estratégias de avaliação formativa: concepções psicopedagógicas e modalidades de aplicação?. In Jean Cardinet e Philippe Perrenoud (Org.) *A Avaliação Formativa num Ensino Diferenciado*, (Actas do Colóquio realizado na Universidade de Genebra, Março de 1978), pp. 175-210. Coimbra : Almedina.
- Allal, L. (2010). Assessment and the Regulation of Learning. In: P. Peterson, Eva Baker, Barry Mc Gaw (Eds.), *International Encyclopedia of Education (3)* (pp. 348-352). Oxford: Elsevier.
- Alves, J.M. (2012). Melhorar os processos e os resultados educativos. O que nos ensina a investigação. Vila Nova de Gaia: Fundação Manuel Leão.
- Alves, J.M. & Moreira, L. (2012). O projeto Fénix e a progressiva transformação da gramática escolar. In JM. Alves, *Projeto Fénix – As Artes do Voo e as Ciências da Navegação* (pp. 83-120), 2012. Porto: Fundação Manuel Leão
- Apple, M (1997). *Os Professores e o Currículo: Abordagens Sociológicas*. Lisboa : Educa
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Bloom, B., Hastings, J Th. & Madaus, G. (1971). *Handbook of Formative and Summative Evaluation of Student Learning*. New York: Mac Graw Hill.
- Bobbitt, J. (2004). *O Currículo*. Lisboa : Didáctica Editora.
- Carrilho Ribeiro, A. (1990). *Desenvolvimento Curricular*. Lisboa : Texto Editora.
- Canário, R. (2005). O que é a Escola? *Um “olhar” sociológico*. Porto: Porto Editora. 2005.
- Gaspar, I. & Roldão, M.C. (2007). *Elementos do Desenvolvimento Curricular*. Lisboa : Universidade Aberta.
- Fernandes, D. (2005). *Avaliação das Aprendizagens: desafios às práticas e políticas*. Lisboa : Texto Editora.
- Goodson, I. (1988). *The Making of Curriculum*. Londres: Falmer Press
- Hattie, J & Timperley, H. (2007). The Power of Feedbak. *Review of Educational Research*, 77(1), 81-112.
- Kliebard, H. (1975). Reappraisal: the Tyler's rationale. In: W. Pinar (Ed.), *Curriculum Theorizing: The reconceptualists* (pp. 70-83), Berkeley, CA: McCutchan Publishing Company.
- Marcelo, C. (2009). Desenvolvimento profissional docente: passado e futuro. *Sísifo / revista de ciências da educação* (8).
- Ministério da Educação. (1972). *Documentos preparatórios da Reforma Curricular do Ministério Veiga Simão*.
- Ministério da Educação. (1988). *Documentos preparatórios da Reforma Curricular do Ministério Roberto Carneiro*.
- Ministério da Educação. (1989). *Documentos preparatórios da Reforma Curricular do Ministério Roberto Carneiro*.
- Montero, L. (2005). *A construção do conhecimento profissional*. Lisboa : Piaget
- OECD. (2005). *Teachers Matter*. Paris: OECD documents.
- OECD (2006) *Pisa 2 – Programme for International Students Assessment 2*.
- OECD (2009) *Pisa 3 – Programme for International Students Assessment 3*.
- OECD (2012) *Pisa 4 – Programme for International Students Assessment 4*.
- OECD (2015) *Pisa 5 – Programme for International Students Assessment 5*.

- Perrenoud, Ph.(2003). Dez princípios para tornar o sistema educativo mais eficaz. In J. Azevedo (Coord.). *Avaliação dos Resultados Escolares* (pp. 103–126). Porto : Edições ASA.
- Roldão, M.C. (1983). Pedagogia por objectivos. *O Professor*, 48, 8-15.
- Roldão, M.C. (1999). *Gestão curricular. Fundamentos e práticas*. Lisboa : DEB.
- Roldão, M.C. (2000). A escola como instância de decisão curricular. In I. Alarcão (ed.), *Escola Reflexiva. Uma Escola em Desenvolvimento e Aprendizagem* (pp. 67-77). Porto : Porto Editora.
- Roldão, M.C.(2003a). *Diferenciação Curricular Revisitada. Conceito, Discurso e Praxis*. Porto : Porto Editora.
- Roldão, M.C (2003b). Quem supervisiona o quê? Liderança e colaboração no trabalho curricular da escola. In Ó. de Sousa e M. Manuel Ricardo (org.), *Uma Escola Com Sentido: o currículo em análise e debate* (pp.135-144), Lisboa : Edições Universitárias Lusófonas.
- Roldão, M.C. (2009). *Estratégias de Ensino. O saber e o agir do professor*. Villa Nova de Gaia: Fundação Manuel Leão.
- Roldão, M.C. (2010). Ensinar e aprender: o saber e o agir distintivos do profissional docente. In R. Teodora Ens e M. Aparecida Behrens (orgs), *Formação do Professor – profissionalidade, pesquisa e cultura escolar* (pp 25-42), Prado Velho, Curitiba, PR : Champagnat Editora –PUCPR.
- Roldão, M.C. (2014). Desenvolvimento do currículo e a melhoria de processos e resultados. In J. Machado & J. M. Alves (Orgs.), *Melhorar a Escola: Sucesso escolar, disciplina, motivação, direção de escolas e políticas educativas* (pp. 136-146). Porto: Universidade Católica Editora. ISBN 978-989-96186-4-0. Disponible sur : [http://www.uceditora.ucp.pt/resources/Documentos/UCEditora/PDF%20Livros/Melhorar-a-escola\\_%20ebook.pdf](http://www.uceditora.ucp.pt/resources/Documentos/UCEditora/PDF%20Livros/Melhorar-a-escola_%20ebook.pdf)
- Roldão, M. C. (2016). Currículo enquanto conhecimento necessário – eixos de uma discussão. In Conselho Nacional de Educação (org), *Pareceres e Documentos*. No prelo
- Roldão, M.C.; Neto-Mendes, A.; Costa, J.A. & Alonso, L.(2006). Organização do trabalho docente: uma década em análise (1996-2005). *Investigar em educação – Revista da Sociedade Portuguesa de Ciências da Educação*, 5, 17-148.
- Roldão, MC e Ferro, N. (2015) O que é avaliar? Reconstrução de práticas e concepções de avaliação. *Estudos de Avaliação Educacional*, 26(63), 570-594.
- Shulman, L. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review*, 57, 4-14.
- Shulman, L. & Shulman, J. (2004). How and what teachers learn: a shifting perspective. *Journal of Curriculum Studies*, 36(2), 257-271.
- Sousa, F. (2010). *Diferenciação curricular e Deliberação Docente*. Porto: Porto Editora.
- Taba, H. (1962). *Curriculum Development – theory and practice*. New York: Harcourt, Brace & World.
- Tyler, R. (1949). *Basic Principles of Curriculum and Instruction*. Chicago: The University of Chicago Press.
- UNESCO. (2015). *Rethinking Education - Towards a global common good?* Unesco. Unesco documents.
- Zabalza, M. (1992) *Planificação e Desenvolvimento Curricular*. Porto : Edições ASA

---

<sup>i</sup> Ce texte reprend et adapte des extraits de textes de l'auteure sur le même thème, correctement intégrés dans les références bibliographiques. L'auteure utilise le portugais écrit conformément à la réforme de l'orthographe.

# Changements curriculaires : un levier pour les pratiques évaluatives des enseignants ?<sup>1</sup>

**Walther Tessaro**

Université de Genève  
[walther.tessaro@unige.ch](mailto:walther.tessaro@unige.ch)

**François-Marie Gerard**

Bureau d'Ingénierie en Éducation et en Formation (BIEF)  
[fmg@fmgerard.be](mailto:fmg@fmgerard.be)

**Marcelo Giglio**

Haute Ecole Pédagogique BEJUNE  
[Marcelo.Giglio@hep-bejune.ch](mailto:Marcelo.Giglio@hep-bejune.ch)

**Pour citer cet article :** Tessaro, W., Gerard, F.-M. & Giglio, M. (2017). Changements curriculaires : un levier pour les pratiques évaluatives des enseignants ? *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 51-60.

Ces dernières années, les réformes curriculaires se sont succédé à un rythme soutenu dans de nombreux pays. Dans les nouveaux plans d'études, plus ou moins innovateurs, différents aspects du processus d'enseignement-apprentissage sont concernés par les changements : les intentions, les contenus, les organisations, les méthodes, l'évaluation notamment (Audigier, Crahay & Dolz, 2006).

En considérant le curriculum comme l'ensemble du processus de formation (Lenoir, 2006), les modifications apportées aux apprentissages visés et aux processus didactiques mis en œuvre pour les atteindre ont inmanquablement un effet sur les situations qui permettent d'évaluer leur degré de maîtrise (Demeuse & Strauven, 2013). Il faut par ailleurs admettre que les textes officiels n'ont qu'une influence partielle sur la réalité de la classe : le chemin est long, en effet, avant d'aboutir à la construction de connaissances par les élèves (Jonnaert, 2011). Enfin, certains ont montré qu'il peut exister un écart entre l'adhésion apparente des enseignants aux orientations proposées et leurs pratiques effectives (Lenoir, Larose et Lessard, 2005) ou les modifications qu'ils pensent apporter à leur enseignement (Giglio, Matthey & Melfi, 2014).

Comme l'ont relevé Gerard et Roegiers (2011), réfléchir à l'évaluation durant la conception même du curriculum est sans doute le meilleur moyen de s'assurer qu'il sera mis en œuvre. Cette réflexion peut être envisagée à plusieurs niveaux, dont celui des pratiques d'évaluation en classe. C'est ce dernier aspect que nous traiterons dans ce texte. Plus précisément, nous nous pencherons sur les pratiques d'évaluation certificative des enseignants face aux changements curriculaires en discernant, d'une part, dans quelle mesure elles en sont

---

<sup>1</sup> Ce texte est issu d'une table ronde du 28<sup>e</sup> colloque de l'ADMEE-Europe tenu à Lisbonne.

tributaires et, d'autre part, les adaptations qu'elles nécessitent et les dérives qu'elles peuvent manifester. La question de l'évaluation des apprentissages scolaires constitue en effet un enjeu majeur de toute réforme (Alliata, Dionnet, Ducrey, Guignard, Jaeggi, Nidegger, Osiek & Saada, 2003).

Nous tenterons dès lors d'identifier de quelle façon le système scolaire s'assure d'une cohésion interne dans la transposition curriculaire (Laveault, Dionne, Lafontaine, Tessaro & Allal, 2014). Plus précisément, nous développerons dans ce texte trois axes que nous discuterons à partir de notre insertion professionnelle respective et des travaux que nous avons menés dans le domaine.

## 1. Plan d'études et évaluation des apprentissages

La composante évaluative d'un plan d'études étant rarement spécifiée, il revient aux enseignants de concevoir des outils qui sont en cohérence avec les objectifs d'apprentissage définis et plus généralement avec les intentions éducatives du système scolaire. Le choix de ne pas traiter simultanément l'enseignement et l'évaluation peut paraître surprenant : comment se justifie-t-il ? Peut-on l'interpréter comme un obstacle ou une opportunité pour les enseignants ?

Le fait de préciser les situations qui permettront d'évaluer les compétences des élèves devrait être une garantie que le curriculum sera réellement mis en œuvre, car les enseignants peuvent mieux percevoir ce que les élèves devront être à même de faire. Cependant, si les outils fournis pour traduire un nouveau curriculum en actions concrètes n'abordent le processus d'enseignement-apprentissage qu'à travers des situations d'évaluation, le risque existe que les enseignants continuent à se référer à leurs anciennes pratiques d'enseignement. Si les concepteurs d'un curriculum souhaitent, par exemple, que les élèves soient à même de résoudre des situations complexes, il est utile qu'ils précisent de manière concrète les types de situations que les élèves devraient pouvoir résoudre. Le meilleur moyen d'y arriver est de proposer aux enseignants, dans le curriculum et/ou dans des outils d'accompagnement, des exemples de situations qui correspondent aux compétences visées. Le précédent curriculum de l'école primaire genevoise (*Les objectifs d'apprentissage de l'école primaire genevoise*, 2000) proposait, à titre d'exemple, des attentes de fin de cycle d'apprentissage composées d'activités d'évaluation « ...illustrant, à l'aide de critères, d'indicateurs, de seuil de réussite et d'exemples de réalisations d'élèves, le bagage scolaire indispensable pour tous les élèves » (Introduction, p. 12). Ces situations peuvent ainsi servir à la fois de support pour l'apprentissage des compétences et pour leur évaluation. L'enjeu est donc double : il importe de ne pas dissocier le plan d'études de l'évaluation, car il risquerait de n'être jamais mis en œuvre, ou de manière très parcellaire. À l'inverse, il convient aussi d'éviter de ne présenter un curriculum qu'en termes d'évaluation, car il ne serait mis en œuvre qu'à travers des activités d'évaluation servant à contrôler l'atteinte des niveaux attendus en faisant l'impasse sur les indispensables activités d'enseignement-apprentissage.

Si l'on suit les principes qui la régissent, l'évaluation devrait logiquement être élaborée dans le prolongement du plan d'études dont elle constitue d'ailleurs un volet. On constate cependant qu'elle est le plus souvent élaborée après la mise en œuvre, intégrant ainsi les transformations contextuelles auxquelles l'ont soumise les pratiques enseignantes. Même si celles-ci visent à respecter le plan d'études, elles comportent une marge irréductible d'imprévisibilité qui échappe à la planification ; elles se trouvent ainsi *in fine* positionnées au centre du processus évaluatif. L'évaluation porte donc davantage sur ce qui est enseigné effectivement que sur ce qui était visé lors de la conception de la formation. En étant attentifs

à ce que ce soit les activités didactiques réalisées en classe qui constituent le référent principal lorsqu'ils élaborent un contrôle écrit certificatif, les enseignants sont garants d'une certaine cohésion interne, le plan d'études orientant dès lors la pratique de façon *indirecte* (Mottier Lopez, Tessaro, Dechamboux & Morales Villabona, 2012).

Les composantes enseignement-apprentissage et évaluation peuvent être travaillées de manière conjointe en élaborant des exemples d'évaluation lors de la conception des plans d'études. Les expériences montrent que les pays qui ont développé un plan d'études selon l'approche par les compétences, sans penser en même temps et de manière systématique la conception d'exemples d'évaluation, n'ont pas réussi à implanter la réforme (Cros, De Ketele, Dembélé, Develay, Gauthier, Ghriss, Lenoir, Murayi, Suchaut & Tehio, 2010). En Communauté française de Belgique (CFB)<sup>2</sup>, ce n'est qu'avec les épreuves externes (venues bien après les référentiels et les programmes) que l'approche par les compétences s'est instaurée dans certaines disciplines. Le paradoxe de la CFB est d'ailleurs exemplaire : on a vu cohabiter un référentiel résolument « compétences » à côté de certaines épreuves externes ignorant celles-ci et d'outils de formation à l'évaluation engagés dans cette voie mais parfois selon des approches différentes, voire opposées (Carette & Dupriez, 2009).

Le choix des responsables scolaires de ne pas inclure une composante évaluative au curriculum peut cependant être interprété comme une opportunité pour les enseignants, car il offre davantage de marge de manœuvre et d'initiative : les pratiques ne sont « orientées » qu'en amont par des objectifs d'apprentissage et pas en aval par l'évaluation. Apporter des éléments spécifiques d'évaluation dans la construction d'un curriculum pourrait être un choix pertinent si ces éléments permettent d'offrir une palette d'outils guidant les enseignants à utiliser une évaluation pour améliorer les apprentissages des élèves, et notamment pour soutenir des démarches d'autoévaluation au sens large (Allal, 1999).

Le fait de considérer de nouvelles compétences dans un curriculum peut aussi être une occasion de transformer les pratiques évaluatives et de concevoir ou ajuster des outils. Prendre en compte les capacités transversales<sup>3</sup> du Plan d'études romand (PER) en Suisse permet notamment de ne plus évaluer les élèves uniquement en lecture et dans la compréhension et la résolution d'un problème mais de réfléchir dès lors à d'autres types d'évaluation formative permettant à l'enseignant d'aider ses élèves à progresser dans leurs apprentissages (Giglio *et al.*, 2014).

Ne pas intégrer de dimension évaluative dans un nouveau curriculum peut cependant représenter un obstacle, car cette décision alourdit la responsabilité pédagogique des enseignants. Ces derniers ont non seulement la tâche de l'opérationnaliser, mais doivent, en outre, préserver une certaine équité dans le traitement de leurs élèves et de ceux des collègues, dans la mesure où ce sont alors les enseignants qui décident de la manière et sur quels objectifs sont évalués leurs élèves, avec la possibilité de différences parfois importantes entre collègues. Cela peut représenter une réelle difficulté sur le plan institutionnel, notamment lorsque des systèmes éducatifs, tels que celui de la CFB, développent de nombreuses évaluations externes identiques pour tous les élèves d'un même niveau afin d'uniformiser l'évaluation (et donc le niveau des élèves), alors même que celle-ci n'était pas travaillée en tant que telle dans les curriculums. Par contre, d'un point de vue pédagogique, le fait d'attribuer à l'enseignant la responsabilité de l'évaluation peut certainement s'avérer lourde à porter, mais elle correspond mieux aux attentes qu'on peut avoir aujourd'hui par rapport à

---

<sup>2</sup> Appelée aussi Fédération Wallonie-Bruxelles (FWB).

<sup>3</sup> Le PER propose cinq capacités transversales : la collaboration, la communication, les stratégies d'apprentissage, la pensée créatrice et la démarche réflexive.

l'évaluation au sein des classes. L'avenir de l'évaluation pédagogique, surtout dans une perspective d'équité, réside plus que vraisemblablement dans l'*évaluation différenciée* : selon les besoins de chaque élève, l'évaluer sur des objectifs appropriés, donner plus de temps à un élève pour une évaluation, lui accorder des supports supplémentaires ou différents selon son style d'apprentissage... Dans cette perspective, l'évaluation n'a plus de visée normative ni discriminatoire. Elle cherche à reconnaître l'élève à travers sa situation et son potentiel.

Au niveau organisationnel, une piste pourrait être d'éviter la dissociation du processus entre ceux qui élaborent le plan d'études, ceux qui le mettent en œuvre et, éventuellement, ceux qui l'évaluent. Associer ces trois catégories d'acteurs aux trois phases du processus permettrait sans doute d'en préserver la cohérence et les exigences ainsi que d'introduire un espace de régulation et d'ajustement aux contingences contextuelles.

Au niveau individuel, les savoirs de l'enseignant peuvent se confondre avec des représentations, des préjugés, des réactions non fondés, et cela tant à propos des pratiques d'enseignement que d'évaluation. Une analyse de ce que les enseignants stagiaires prédisent du déroulement d'une auto-évaluation à proposer à leurs élèves et ce qui se passe réellement tant du côté des tentatives fructueuses que des difficultés a, par exemple, mis en évidence l'existence d'un décalage considérable entre la prédiction et l'observation (Giglio *et al.*, 2014). Une piste pourrait être de proposer à l'attention des enseignants des démarches d'autoévaluation de l'élève qui permettraient de promouvoir la création d'un espace de régulation et d'autorégulation.

## 2. Différents types d'apprentissage à évaluer

Un curriculum considère généralement des enseignements développés à partir de champs disciplinaires institutionnalisés. Il comprend aussi des apports qui ne reposent pas sur une forme scolaire habituelle (Rey, 2010) et qui contribuent à la formation générale (p. ex : éducation au développement durable, santé et bien-être). Il vise enfin à développer ce que le PER nomme des *capacités transversales*, liées au fonctionnement individuel et aux interactions (ex. : démarche réflexive, collaboration, stratégies d'apprentissage). Quels sont les effets de cette diversité d'objectifs sur les pratiques évaluatives et sur l'articulation enseignement-apprentissage-évaluation ?

Les systèmes scolaires n'évaluent la plupart du temps que des objectifs d'enseignement facilement évaluables. Or, les capacités transversales sont difficiles à évaluer, ce qui les condamne à n'être évaluées, le plus souvent, que de manière intuitive. Elles ne correspondent pas à une perspective de « transmission », mais s'inscrivent dans une vue de « développement » qui implique une durée et une progression et pour laquelle les pratiques évaluatives habituelles (épreuves ponctuelles) sont inadéquates. Il faut faire appel à des dispositifs temporels de recueil d'indices auxquels on donne sens et où une « indispensable subjectivité est requise » (Gerard, 2002, 2017).

La plupart des enseignants évaluent en réalité ces objectifs qui relèvent des attitudes ou des savoir-être, de manière plus ou moins consciente et souvent avec un sentiment d'infailibilité de leur jugement (Crahay, 2006). Pourtant, les difficultés pour l'évaluation de ces objectifs ne manquent pas, tant sur le plan méthodologique qu'éthique (Gerard, 2011). D'un point de vue méthodologique, les attitudes ou les savoir-être étant par nature non observables (Rosenberg & Hovland, 1960), leur évaluation ne peut se faire que par inférence sur la base d'indicateurs ou de comportements. Par ailleurs, le savoir-être n'existant que quand il se manifeste de manière spontanée, il ne peut donc être évalué qu'en situation naturelle ou en « situation non contrainte » (De Ketele, 2010). Dans le cadre scolaire, il est difficile d'observer les élèves dans

ce type de situations, d'autant plus qu'il conviendrait de disposer de plusieurs observations indépendantes pour pouvoir éventuellement effectuer une inférence. Sur le plan éthique, la difficulté réside dans le fait que l'élève ne doit pas savoir qu'il est évalué, sinon il manifesterait vraisemblablement des comportements « attendus » et non spontanés. Or, une évaluation devrait toujours être transparente, la personne évaluée sachant qu'elle est évaluée et sur quoi, d'autant plus lorsque l'objet d'évaluation porte sur des éléments aussi personnels que ses attitudes ou savoir-être.

Ainsi, s'il y a accord sur la nécessité de développer des capacités transversales à l'école, elles ne sont pas forcément exigibles. Il est en effet légitime de se demander s'il convient de les évaluer et s'il ne faut pas se limiter à créer les conditions de leur émergence afin de prendre en compte l'élève dans sa globalité et de donner davantage de sens à certains enseignements. Mais, si les programmes scolaires prescrivent la créativité, la collaboration, la réflexion et la communication parmi les capacités à développer à travers les disciplines scolaires à l'école, elles nous invitent à repenser les formes d'une « évaluation pour apprendre » en créant les conditions favorables pour que ces capacités puissent être développées chez l'élève.

De plus, les apports à la formation générale courent le risque de se trouver positionnés hiérarchiquement au-dessous des disciplines institutionnalisées (déjà elles-mêmes hiérarchisées entre elles). Cette déqualification potentielle a forcément des conséquences sur l'investissement des acteurs, surtout si ces apports sont pour partie pris en charge par d'autres personnes que des enseignants, par exemple des psychologues ou des éducateurs, et ne revêtent que de faibles enjeux scolaires. En effet, les nouveaux curriculums de l'enseignement obligatoire proposent d'aborder des types de savoirs longtemps abordés seulement dans l'enseignement supérieur et décrits par Bruffee (1995) comme plus réflexifs, voire plus productifs et créatifs. Le travail de mémorisation ou de déduction cède alors place au raisonnement, à l'argumentation. Ces capacités supposent une mise en retrait de l'enseignant pour qu'il puisse observer, apprécier et soutenir un apprentissage autorégulé de ses élèves.

Les nouveaux curriculums mettent en évidence la mission de l'école de préparer les élèves à interagir de manière créative et réflexive, dans une société dans laquelle les professionnels ont à agir avec d'autres partenaires. Pour développer ces capacités, il est important que l'élève puisse se situer dans sa progression, c'est-à-dire s'autoévaluer avec le soutien de l'enseignant. Des discussions en classe entre des enseignants et leurs élèves ont été observées, en lien avec deux capacités transversales indiquées dans certains plans d'études comme le PER : la pensée créatrice et la collaboration. Il apparaît que ce type d'autoévaluation en classe est complexe, mais possible à certaines conditions et en dépassant certains *a priori* de la part des enseignants. Par exemple, lors d'un travail de coévaluation conduit durant un stage, pour plusieurs enseignants en formation il semble difficile de se dégager de certaines représentations, de certaines opinions *a priori* sur la capacité réflexive des élèves. Certains étudiants peuvent penser que l'autoévaluation pourrait devenir une perte de temps. D'autres semblent se préoccuper de la réussite de la tâche d'autoévaluation au détriment de l'apprentissage de l'élève à s'autoévaluer. (Giglio, 2013).

De plus, une activité de créativité, de collaboration et de réflexion nécessite une pratique évaluative qui invite l'élève à verbaliser, voire à dessiner, à écrire ou schématiser ses manières de penser, d'agir et d'interagir en classe. En effet, il est nécessaire que les élèves puissent apprendre à s'autoévaluer et aussi à distinguer et comparer leur point de vue avec celui de l'enseignant (Giglio & Rothenbühler, 2014). De cette manière, l'élève peut prendre conscience de sa progression.

Si les plans d'études reconnaissent de façon unanime l'importance des capacités transversales

dans la progression des apprentissages et dans l'insertion sociale et professionnelle, il apparaît qu'elles ne peuvent être évaluées pour elles-mêmes. L'évaluation de ces capacités transversales, intégrées ainsi dans les différentes disciplines scolaires, questionne la forme de celle-ci, mais aussi les liens implicites ou explicites avec les tâches qui sont proposées. Pourtant, certaines pédagogies tentent de prendre en compte les aspects sociaux, de gestion ou de communication dans l'enseignement et dans l'évaluation. En apprentissage coopératif, par exemple, les habiletés sociales sont aussi importantes que les objectifs disciplinaires abordés. Elles sont évaluées de façon formative au même titre à la fin de l'activité, par l'enseignant mais aussi par les participants.

Il apparaît que si l'on souhaite se pencher sur l'évaluation des capacités transversales, il est nécessaire de clarifier en amont la responsabilité de leur mise en œuvre dans l'établissement scolaire. S'assurer, en l'occurrence, d'un suivi pédagogique qui précise dans quels contextes et par quels enseignants sont abordés les objectifs d'apprentissage visés. Le recours à des techniques telles que le *team teaching*, à travers la mise en commun interactive de ressources complémentaires soit sur un contenu, soit sur le processus pour assurer de meilleures conditions d'apprentissage et d'évaluation, peut se révéler une piste concrète intéressante, avec un temps de concertation hebdomadaire supplémentaire. C'est à cette condition, sans doute, que cette partie d'un plan d'études peut être implémentée de façon progressive et équitable.

### 3. Les enseignants, acteurs de changements ?

La transposition curriculaire est habituellement comprise comme une perspective descendante, qui va des prescriptions et des plans d'études vers les contenus enseignés en classe. Néanmoins, des études ont montré que les pratiques du terrain peuvent aussi participer à une modélisation et à la transformation des documents institutionnels, dans un mouvement ascendant (Mottier Lopez & Tessaro, 2010). Des manières de construire un contrôle écrit certificatif ou un portfolio d'apprentissage peuvent, par exemple, comme pratiques innovantes, être diffusées en formation continue puis progressivement institutionnalisées et mises à disposition des enseignants, voire prescrites. Nous pouvons ainsi nous demander dans quelle mesure l'évaluation du curriculum réellement enseigné peut influencer à son tour le curriculum prescrit, dans une perspective ascendante ou tout du moins dialectique. Plus généralement, il s'agit de comprendre dans quelle mesure et de quelle manière les enseignants peuvent être acteurs des changements curriculaires par leurs pratiques effectives et les constats qui en résultent.

De façon générale, les enseignants ne sont pas réputés pour être des acteurs de changement, car trop souvent « les pratiques innovantes charrient leur lot de déceptions, de conflits, de moments de lassitude, de sentiments d'échec ou d'impuissance, de rancœurs suscitées par l'ingratitude des apprenants ou de l'organisation pour laquelle les professionnels " se défoncent " » (Perrenoud, 2001). Néanmoins, ils sont amenés à créer leurs propres outils d'évaluation, à les partager, d'autant plus que les plateformes d'échange d'outils pédagogiques se multiplient et sont facilement accessibles. Des outils largement partagés pourraient dès lors influencer le curriculum prescrit sur la base de l'évaluation du curriculum enseigné. Il existe par ailleurs un « curriculum tel qu'il est évalué », défini par les praticiens. Brau-Antony et Hubert (2014) montrent comment, notamment dans le domaine de l'éducation physique et sportive, les enseignants se dotent d'une certaine autonomie par rapport aux prescriptions officielles. Ils créent ainsi leurs propres règles, selon les obstacles rencontrés (profils d'élèves à évaluer, niveaux d'exigence attendus), tout en respectant un certain nombre de principes communs. Les auteurs font l'hypothèse que cette prise de distance par rapport aux

instructions officielles reflète un « genre professionnel » (Clot & Faïta, 2000) caractérisé par une expérience professionnelle commune, une sorte de mémoire collective qui permet de faire évoluer les pratiques d'évaluation en fonction de leurs propres façons de faire. Ces nouvelles pratiques, expression de ce genre professionnel, peuvent dès lors à leur tour influencer les prescriptions curriculaires.

Pour qu'un enseignant puisse devenir acteur d'un changement, il devrait disposer de temps et d'espaces de réflexion sur les enjeux du curriculum. Mais cela ne va pas de soi. Afin que le changement soit une amélioration, il nécessite des recherches collaboratives qui attribuent une place importante à la pratique. En Suisse romande, des chercheurs ont travaillé avec des enseignants pendant cinq ans dans le cadre d'une recherche participative afin de co-construire un cadre légal sur l'évaluation des degrés inférieurs (Petignat & Droz Giglio, 2016). Cela a permis aux enseignants d'échanger sur leurs pratiques évaluatives et de proposer certaines innovations au Service de l'enseignement obligatoire du canton de Neuchâtel. Par cette démarche participative, les enseignants ont pu influencer les décisions prises dans le curriculum prescrit en ce qui concerne l'évaluation, et cela dans une perspective dialectique renouvelée, associant des mouvements ascendants et descendants. Outre la modification du cadre légal, de nouveaux outils d'évaluation ont été proposés aux enseignants leur permettant notamment d'évaluer la progression des apprentissages grâce aux traces recueillies et réunies dans un portfolio. Ce processus a été poursuivi en y associant les instances syndicales et politiques ainsi que les établissements scolaires, dans un but de forger tant le cadre légal de l'évaluation que le cadre pratique de l'évaluation en classe.

L'enseignant, en tant que praticien chercheur, peut aussi être acteur d'un changement en développant des séquences didactiques à partir d'un nouveau curriculum. Ce type de démarche permet un élargissement des apprentissages [*Expansive Learning*] (Engeström, 1999) : c'est en analysant les nouvelles situations que les enseignants, avec un chercheur, peuvent créer des outils, les essayer, les consolider et arriver à une nouvelle situation pédagogique améliorée « pas à pas ». Ce travail peut contribuer à des usages nouveaux, à des pratiques d'enseignement innovantes dans le cadre d'un collectif de travail. Par cet agir innovationnel (Cros, 2007), qui alterne de manière réflexive des observations, des réalisations et des résultats de recherche, l'enseignant et le chercheur peuvent devenir des acteurs d'un changement éducatif.

#### **4. D'autres questions en conclusion**

Les trois axes abordés dans ce texte ne représentent qu'une partie des questions qui ont émergé dans les échanges. Le choix de ces axes a été fait en fonction de l'intérêt manifesté par les auteurs et il n'épuise assurément pas le questionnement général de la place de l'évaluation dans le curriculum.

D'autres aspects importants seraient à traiter concernant l'influence d'un changement curriculaire sur les pratiques évaluatives en classe. Par exemple, le recours de plus en plus généralisé aux évaluations standardisées et leurs effets sur le curriculum enseigné et évalué. En effet, en illustrant par des tâches les attentes institutionnelles, ce type d'évaluations est intéressant pour les enseignants et peut avoir plus d'influence que les prescriptions (Baluteau, 2013). Bien que leur usage puisse déboucher sur un rétrécissement du curriculum (Mons, 2009), ces épreuves constituent un moyen intéressant de régulation des pratiques lors de changements curriculaires (Ntamakiliro & Tessaro, 2010). Un autre aspect auquel il nous semble important de réfléchir concerne l'accompagnement des enseignants lors des changements, qui peuvent être parfois sensiblement innovants et qui demandent ainsi une

adaptation conséquente des pratiques. Les outils d'information institutionnels étant insuffisants (Gilliéron Giroud & Tessaro, 2009), ce sont des démarches de formation continue qui peuvent répondre aux craintes suscitées par les changements. Si les enseignants ne sont pas convaincus, ils risquent en effet d'opposer des résistances (Maulini & Wandfluh, 2007).

Si des changements d'importance ont des conséquences sur les pratiques d'évaluation des apprentissages, ils peuvent fragiliser la communication des résultats aux familles qui, pour certaines d'entre elles, doivent reconsidérer une connaissance parfois déjà incertaine du système éducatif. Les prétentions des parents constituent une pression externe qui oblige à une nécessaire coopération : un accent porté sur la communication des nouveaux contenus d'enseignement et des conséquences sur leurs pratiques évaluatives s'inscrit dans une perspective de « rendre compte » (*accountability*) qui caractérise notre société actuelle (Montandon, 1987). Sans doute continueront-ils de cette façon à soutenir les parents dans des choix adaptés aux caractéristiques de leur enfant (van Zanten, 2001).

Parmi les modèles de formation qui contribuent à une harmonisation des pratiques d'évaluation, des dispositifs de modération sociale (Maxwell, 2002), visant à développer les compétences professionnelles des enseignants et à formuler des jugements d'évaluation équitables, constituent à notre sens une perspective intéressante. Les enseignants y engagent des démarches de réflexion collective critique sur les référents et sur leur rôle dans les pratiques d'évaluation interne à la classe ainsi que de confrontation de leurs différentes façons d'interpréter et d'évaluer les productions des élèves. Ces dispositifs contribuent à l'élaboration progressive d'une culture commune en visant la construction de consensus à propos de la correction des productions des élèves et la notation de leurs copies (p. ex. Wyatt-Smith, Klenowski & Gunn, 2010), auxquelles peuvent s'ajouter la construction de l'instrument d'évaluation (Mottier Lopez *et al.*, 2012).

## 5. Références

- Allal, L. (1999). Impliquer l'apprenant dans le processus d'évaluation : promesses et pièges de l'autoévaluation. In C. Depover & B. Noël (Éd.), *L'évaluation des compétences et des processus cognitifs* (pp. 35-56). Bruxelles : De Boeck.
- Alliata, R., Dionnet, S., Ducrey, F., Guignard, N., Jaeggi, J.-M., Nidegger, C., Osiek, F. & Saada, E. (2003). *À la fin de la 6<sup>e</sup> primaire. Enquête auprès de la volée 2000 des élèves de 6P. Acquis et compétences des élèves, représentations des parents et des enseignants*. Genève : Service de la recherche en éducation.
- Audigier, F., Crahay, M. & Dolz, J. (Éd.) (2006). *Curriculum, enseignement et pilotage*. Bruxelles : De Boeck.
- Baluteau, F. (2013). Les dispositifs d'évaluation nationale et les réorganisations locales autour du curriculum. In V. Dupriez et R. Malet (Éd.), *L'évaluation dans les systèmes scolaires* (pp. 64-75). Bruxelles : De Boeck.
- Brau-Antony, S. & Hubert, B. (2014). Curriculum en EPS et évaluation certificative au baccalauréat. *Questions vives*, 22, 67-82.
- Bruffee, K. (1995). Sharing our toys: cooperative learning versus collaborative learning. *Change*, 27, 12-19.
- Carette, V. & Dupriez, V. (2009). La lente émergence d'une politique scolaire en matière d'évaluation des élèves. Quinze ans de transformations en Belgique francophone, *Mesure et évaluation en éducation*, 32(3), 21-45.
- Clot, Y. & Faïta, D. (2000). Genres et styles en analyse du travail. Concepts et méthodes. *Travailler*, 4, 7-42.
- Clot, Y., Faïta, D., Fernandez, G. & Scheller, L. (2001). Entretiens en autoconfrontation croisée : une méthode en clinique de l'activité. *Éducation permanente*, 146, 17-27.
- Crahay, M. (2006). L'évaluation des élèves : entre mesure et jugement. In G. Figari & L. Mottier Lopez (Éd.). *Recherche sur l'évaluation en éducation - Problématiques, méthodologies et épistémologie* (pp. 132-138). Paris : L'Harmattan.

- Cros, F. (2007). *L'agir innovatif. Entre créativité et formation*. Bruxelles : De Boeck.
- Cros, F., De Ketele, J.-M., Dembélé, M., Develay, M., Gauthier, R.-F., Ghriess, N., Lenoir, Y., Murayi, A., Suchaut, B. & Tehio, V. (2010). *Les réformes curriculaires par l'approche par compétences en Afrique*. Document de travail 97. Paris : Agence française de Développement.
- De Ketele, J.-M. (2010). Faces visibles et cachées des classements internationaux. Une tentative de modélisation des tensions dans l'enseignement obligatoire. *Revue Internationale d'Éducation de Sèvres*, 54, 39-49.
- Demeuse, M. & Strauven, C. (2013). *Développer un curriculum d'enseignement ou de formation. Des options politiques au pilotage*. Bruxelles : De Boeck.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Riettinger & R.-L. Punamäki (Éd.), *Perspectives on activity theory* (pp. 19-38). Cambridge: Cambridge University Press.
- Gerard, F.-M. (2002). L'indispensable subjectivité de l'évaluation. *Antipodes*, 156, 26-34.
- Gerard, F.-M. (2011). L'apprentissage et l'évaluation des attitudes et des savoir-être. In X. Roegiers (Éd.), *Curricula et apprentissages au primaire et au secondaire – La Pédagogie de l'Intégration comme cadre de réflexion et d'action* (pp. 146-150). Bruxelles : De Boeck.
- Gerard, F.-M. (2016). Objectiver la subjectivité. In D. Leduc & S. Béland (dir.). *Regards sur l'évaluation des apprentissages en arts à l'enseignement supérieur* (pp. 29-47) Montréal : Presses de l'Université de Montréal.
- Gerard, F.-M. & Roegiers, X. (2011). Currículo e Avaliação: ligações que nunca serão suficientemente fortes. In M.P. Alves, & J.-M. De Ketele. (Éd.), *Do Currículo à avaliação, da avaliação ao currículo* (pp. 143-158). Porto : Porto Editora,. Version française accessible sur [http://www.bief.be/index.php?enseignement/publications/curriculum\\_evaluation\\_des\\_liens\\_qui\\_seront\\_jamais&s=3&r=17&uid=479&lg=fr](http://www.bief.be/index.php?enseignement/publications/curriculum_evaluation_des_liens_qui_seront_jamais&s=3&r=17&uid=479&lg=fr)
- Giglio, M. (2013). L'autoévaluation de l'élève pour apprendre à collaborer créativement : des apriori à découvrir et dépasser. *Actes du colloque de l'ADMEE-Europe*. Fribourg.
- Giglio, M. & Rothenbühler, P. (2014). Analyse d'un continuum entre tâches créatives et collaboratives, métaréflexion et autoévaluation chez l'élève. *Actes du colloque de l'ADMEE-Europe*. Marrakech.
- Giglio, M., Matthey, M.-P. & Melfi, G. (2014). *Réactions des formateurs d'enseignants à un nouveau curriculum scolaire*. Bienne : HEP-BEJUNE.
- Gilliéron Giroud, P. & Tessaro, W. (2009). L'évaluation en classe en Suisse romande : des politiques aux pratiques enseignantes. *Mesure et évaluation en éducation*, 32(3), 47-76.
- Jonnaert, Ph. (2011). Curriculum, entre modèle rationnel et irrationalité des sociétés. *Revue internationale d'éducation de Sèvres*, 56, 134-145.
- Laveault, D., Dionne, E., Lafontaine, D., Tessaro, W. & Allal, L. (2014). L'évaluation des compétences en salle de classe : régulations entre politiques et pratiques. In C. Dierendonck, E. Loarer & B. Rey (Éd.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (pp. 85-96). Bruxelles: De Boeck.
- Lenoir, Y. (2006). Du curriculum formel au curriculum enseigné : comment des enseignants québécois du primaire comprennent et mettent en œuvre le nouveau curriculum de l'enseignement primaire. In F. Audigier, M. Crahay & J. Dolz (Éd.), *Curriculum, enseignement et pilotage* (pp. 119-141). Bruxelles : De Boeck.
- Lenoir, Y., Larose, F. & Lessard, C. (2005). *Le curriculum de l'enseignement primaire : regards critiques sur ses fondements et ses lignes directrices*. Sherbrooke : Éditions du CRP.
- Les objectifs d'apprentissage de l'école primaire genevoise* (2000). Genève : Département de l'instruction publique.
- Maulini, O. & Wandfluh, F. (2007). Organiser et planifier le travail scolaire : des modèles théoriques aux pratiques des enseignants. In M. Gather Thurler et O. Maulini (Éd.), *L'organisation du travail scolaire. Enjeu caché des réformes ?* (pp. 121-149). Québec : Presses de l'Université de Québec.
- Maxwell, G. (2002). *Moderation of Teacher Judgments in Student Assessment*. Brisbane: Queensland School Curriculum Council.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie. Recherches en Éducation*, 169, 99-140.

- Montandon, C. (1987). L'essor des relations famille-école : problèmes et perspectives. In C. Montandon & P. Perrenoud (Éd.), *Entre parents et enseignants: un dialogue impossible ?* (pp. 23-47). Berne: Peter Lang.
- Mottier Lopez, L. & Tessaro, W. (2010). Évaluation des compétences à l'école primaire genevoise : entre prescriptions et pratiques. *Mesure et évaluation en éducation*, 33 (3), 29-53.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L. & Morales Villabona, F. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives*, 18, 159-175.
- Ntamakiro, L. & Tessaro, W. (2010). Vers plus d'épreuves externes: une standardisation de l'évaluation scolaire ? In P. Gilliéron Giroud & L. Ntamakiro (Éd.), *Réformer l'évaluation scolaire: mission impossible ?* Berne : Lang.
- Perrenoud, P. (2001), Préparer les enseignants au changement. Journées "Éducation & Prospective", Forum de l'innovation pédagogique et éducative, Ollioules. [http://www.unige.ch/fapse/SSE/teachers/perrenoud/php\\_main/php\\_2001/2001\\_11.html](http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2001/2001_11.html).
- Petignat, P. & Droz Giglio, C. (2016). La co-construction d'une démarche évaluative au 1<sup>er</sup> cycle primaire. In L. Mottier Lopez et W. Tessaro (Éd.), *Le jugement professionnel, au cœur de l'évaluation et de la régulation des apprentissages* (pp. 313-335). Berne : Peter Lang.
- Rosenberg, M.J. & Hovland, C.I. (1960). Cognitive, affective and behavioral components of attitudes. In C.I. Hovland & M.J. Rosenberg (Éd.), *Attitude Organization and Change*. New Haven : Yale University Press.
- Van Zanten, A. (2001). *L'école de la périphérie. Scolarité et ségrégation en banlieue*. Paris : PUF.
- Wyatt-Smith, C., Klenowski, V. & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59-75.

# Perceptions et pratiques d'évaluation des apprentissages dans l'enseignement universitaire portugais

**Domingos Fernandes**

Universidade de Lisboa

[dfernandes@ie.ulisboa.pt](mailto:dfernandes@ie.ulisboa.pt)

**Isabel Fialho**

Universidade de Évora

[ifialho@uevora.pt](mailto:ifialho@uevora.pt)

**Carlos Barreira**

Universidade de Coimbra

[cabarreira@fpce.uc.pt](mailto:cabarreira@fpce.uc.pt)

**Pour citer cet article :** Fernandes, D., Fialho, I. & Barreira, C. (2017). Perceptions et pratiques d'évaluation des apprentissages dans l'enseignement universitaire portugais. *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 61-75.

## 1. Introduction

Cet article est basé sur les résultats d'une recherche menée dans le cadre du projet AVENA - *Évaluation, enseignement et apprentissage au Portugal et au Brésil : réalités et perspectives* – financé, au Portugal, par la *Fondation pour la science et la technologie* (Projet PTDC/CPE-CED/114318/2009). La recherche a eu lieu entre avril 2011 et novembre 2014, impliquant quatre universités portugaises et trois universités brésiliennes, 36 chercheurs seniors et 20 jeunes chercheurs.

La recherche ciblait, avant tout, deux objectifs essentiels, à savoir : connaître et comprendre les pratiques d'enseignement et d'évaluation des enseignants de l'enseignement supérieur, dans une variété de dimensions, et percevoir les perceptions des enseignants et des étudiants au sujet de ces pratiques. Pour une description détaillée de l'arrière-plan du problème et des questions et de la méthodologie du projet AVENA voir, par exemple, Fernandes, Rodrigues et Nunes (2012).

Dans cet article, nous ne considérons que les données recueillies dans les quatre universités portugaises et l'analyse qui a porté sur les pratiques d'enseignement et les perceptions des étudiants et des enseignants dans le domaine de l'évaluation.

Dans ce contexte, les pratiques d'évaluation des enseignants ont été décrites, analysées et interprétées selon un ensemble de dimensions, ainsi que les perceptions des étudiants et des enseignants à l'égard de ces pratiques.

## 2. Cadre théorique et conceptuel

Le développement de l'Espace européen de l'enseignement supérieur (EEES) a été fondé sur la *Déclaration de Bologne*, de 1999, aussi connue comme le *processus de Bologne*, qui remet en question les routines installées des établissements d'enseignement supérieur européens, grâce à un *agenda pédagogique* qui, au-delà des controverses, a réussi à rassembler la sympathie de larges secteurs du monde académique et de la société en général. D'une certaine manière, les dirigeants politiques ont décidé de confirmer les idées prônées depuis longtemps dans la littérature et les milieux académiques, pour que l'amélioration de l'apprentissage, de l'évaluation pédagogique et de l'enseignement devienne une préoccupation centrale des institutions d'éducation. Ainsi, la *pédagogie* est, à présent, à l'ordre du jour de la vie universitaire. Par conséquent, toutes sortes d'initiatives ont proliféré, donnant lieu à un nombre important de publications qui mentionnent particulièrement la nécessité de : a) investir dans l'articulation et/ou l'intégration des processus d'apprentissage, d'évaluation et d'enseignement; b) créer des conditions pour améliorer les connaissances et les compétences des enseignants dans des domaines tels que l'apprentissage, l'évaluation et l'enseignement; c) promouvoir l'autonomie et la participation des étudiants à des activités liées à l'élaboration du programme; d) encourager l'interaction sociale entre les étudiants et entre ceux-ci et les enseignants; e) reconnaître la possibilité d'améliorer les pratiques curriculaires des enseignants (par exemple, Bryan & Clegg, 2006, Falchicov, 2005).

Cette *dynamique de réforme* orientée vers la transformation et l'amélioration de la dite *pédagogie de/ dans l'enseignement supérieur* n'a pas été développée, et ne l'est toujours pas, indépendamment des contradictions ayant une incidence sur les politiques publiques relatives à l'enseignement supérieur, marquées par certaines tendances, à savoir : a) l'approche adaptée aux logiques du marché à travers des stratégies de nature mercantile; b) les fusions entre institutions, en vue de leur internationalisation et de leur capacité concurrentielle, souvent mesurée par les *classements* internationaux; c) l'utilisation des processus d'évaluation inducteurs de changements importants dans le travail académique. Ces tendances affectent la vie académique et professionnelle des enseignants universitaires, de plus en plus visés par les processus de régulation qui affectent leur autonomie et la liberté académique (Escudero, 2012). Néanmoins, pour cet auteur, la controverse autour du dit *processus de Bologne* est bien plus suscitée par les perspectives technico-bureaucratiques et commerçantes qui l'influencent que par l'*agenda pédagogique* proposé dans son champ d'application. En fait, cet agenda propose des changements depuis longtemps demandés et nécessaires pour améliorer la qualité de la formation des étudiants.

Une analyse de la littérature publiée dans des revues internationales de référence, pendant la première décennie de ce siècle, a permis d'identifier des pratiques d'évaluation généralement associées à l'amélioration de l'apprentissage des étudiants. Dans un ensemble de recherches (par exemple, Klenowski, Askew & Carnell, 2006 ; Miller, 2009), les enseignants utilisaient une variété de tâches d'évaluation, en général entre deux et trois, sur lesquelles les étudiants travaillaient, dans un environnement où l'interaction avec les pairs et les enseignants, la distribution du *feed-back*, le partage des réflexions et l'utilisation de diverses dynamiques d'enseignement et d'apprentissage étaient des caractéristiques frappantes.

Dans d'autres études (par exemple, Dancer & Kamvounias, 2005; Nicol, 2009; Poon *et al.*, 2009), il y avait une participation systématique des étudiants à des actions directement liées à l'évaluation des apprentissages en fonction de certaines dynamiques et telles que : a) le travail collaboratif en petits groupes, avec l'analyse des tâches, la planification et la gestion du temps, la distribution et le partage du *feed-back*; b) la participation à la discussion, la définition et la

clarification des critères d'évaluation relatifs aux tâches qui leur étaient proposées; c) la participation à l'auto-évaluation et à l'évaluation par les pairs.

En ce qui concerne la nature et l'utilisation du *feed-back*, la littérature a montré que, en général, celui-ci était critique, écrit ou oral, et qu'il était distribué par les enseignants afin d'informer, de corriger, de motiver, de guider et d'attribuer une moyenne de l'ensemble des notes de l'évaluation (par exemple, Macmillan & Mclean, 2005, Trotter, 2006). Les étudiants étaient également impliqués dans la distribution du *feed-back* entre leurs pairs. Pour ce faire, ils utilisaient des outils tels que des listes de vérification, des éléments d'évaluation ou des tableaux décrivant différents niveaux de performance pour une tâche donnée (par exemple, Walker & Warhurst, 2000; Walker & Warhurst, 2000; Wanous, Procter & Murshidb, 2009).

La distribution du *feed-back* est particulièrement mentionnée dans la littérature, puisque son rôle est reconnu dans le développement de plusieurs processus de pensée des étudiants plus complexes, ainsi que dans la réglementation et l'autorégulation de leurs apprentissages (par exemple, Albertino & De Sousa, 2004; Andrade & Du, 2007; Klenowski *et al.*, 2006; Macmillan & Mclean, 2005). La participation des étudiants à la définition et à la discussion des critères utilisés pour évaluer leur travail a également été considérée comme étant liée à l'amélioration de leurs apprentissages. En fait, en devenant plus conscients et en intériorisant ces critères, les étudiants étaient mieux préparés à évaluer leur travail et celui de leurs collègues et à se concentrer sur ce qu'ils devaient apprendre ; les étudiants développaient ainsi leurs capacités métacognitives, d'auto-évaluation et d'évaluation par les pairs et de régulation et d'autorégulation de leurs apprentissages (par exemple, Klenowski *et al.*, 2006; Mok *et al.*, 2006; Price & O'Donovan, 2006).

En bref, la littérature abonde d'études avec des bases empiriques solides qui nous indiquent la possibilité réelle d'améliorer considérablement ce que les étudiants de l'enseignement supérieur apprennent et comment ils doivent le faire. À cette fin, les pratiques curriculaires, en général, et les pratiques d'enseignement et d'évaluation des enseignants, en particulier, sont des éléments essentiels pour que les étudiants apprennent davantage, avec une meilleure compréhension et profondeur.

### 3. Méthode

La recherche a été réalisée en cinq phases : a) la phase théorique et conceptuelle; b) la phase de l'analyse documentaire; c) la phase de l'étude extensive; d) la phase de l'étude intensive; e) la phase de l'interaction sociale et de la réflexion. Cet article a pris en compte la *phase de l'étude extensive* et la *phase de l'étude intensive*, pendant lesquelles la stratégie générale de l'étude, son *design*, la méthodologie et toutes les méthodes et procédures de collecte de données ont été matérialisées (Fernandes, Rodrigues & Nunes, 2012).

Pendant la *phase de l'étude extensive*, un questionnaire a été conçu et mené *en ligne* afin d'étudier les perceptions des enseignants et des étudiants à l'égard de chacun des objets de la recherche (enseignement, évaluation et apprentissage). Le questionnaire, identique pour les enseignants et les étudiants (à l'exception des questions sociodémographiques), comprenait 45 items du type Likert, 15 items pour chacun de ces objets de la recherche. Par exemple, chacun des 15 items de l'objet de l'évaluation décrivait, à travers des affirmations, les pratiques d'évaluation par rapport auxquelles les répondants devaient exprimer leur degré d'accord ou de désaccord, sur une échelle à quatre niveaux : *Pas du tout d'accord* (DT = -2); *Plutôt pas d'accord* (D = -1); *D'accord* (C = 1); *Tout à fait d'accord* (CT = 2). Bien que l'option d'un nombre pair de niveaux de réponse ne soit pas consensuelle dans la littérature (Gigliome & Matalon, 2001), dans ce cas il est justifié par une raison pertinente. En fait, les enseignants et les étudiants connaissent

l'essentiel du contenu des questions abordées et, dans de telles circonstances, ils étaient en mesure de prendre une position claire sur l'accord ou le désaccord (Hill & Hill, 2008). Autrement dit, les répondants sont des personnes informées de la matière en question (enseignement, évaluation et apprentissage). D'autre part, l'utilisation du score attribué à chacun des niveaux d'échelle, ci-dessus mentionnée, a permis de compenser l'absence du point intermédiaire "ni d'accord ni en désaccord".

Les réponses obtenues à travers l'enquête par questionnaire ont permis d'analyser les perceptions des enseignants et des étudiants à l'égard des pratiques d'évaluation des apprentissages. Toutefois, le fait que certains éléments ne soient pas unidimensionnels soulève une limitation à l'étude. L'analyse qualitative des données (entretiens et observations) et, également, les contacts informels établis avec les participants ont permis de recueillir des informations supplémentaires sur un large éventail de dimensions des pratiques d'évaluation, réduisant, en quelque sorte, cette limitation.

Les réponses aux questionnaires à tous les enseignants et étudiants de chaque université, étaient anonymes et volontaires, mais sans aucune exigence quant au taux de retour. Plusieurs rappels, demandant aux participants potentiels de répondre aux questionnaires, ont été réalisés, ce qui a permis d'augmenter significativement le nombre de répondants par rapport au nombre obtenu lors du premier contact. Les taux de retour ont varié selon les universités, s'élevant à 10% - 20% pour les étudiants et à 15% - 30% pour les enseignants, pouvant ainsi être considérés comme acceptables dans des situations similaires (Jones, 1999). À la fin du processus, 990 questionnaires d'enseignants et 4568 questionnaires d'étudiants, de tous les domaines de connaissance, et dans le cadre des unités d'enseignement théorico-pratiques ou pratiques qui structurent une licence (1er cycle), ont été validés. Une unité d'enseignement était considérée structurante lorsqu'elle incluait des concepts et des fondements essentiels pour l'étude d'autres unités d'enseignement d'un certain cours supérieur.

La cohérence interne des questionnaires aux enseignants et aux étudiants, en considérant un total de 45 items liés aux questions de l'enseignement, de l'évaluation et de l'apprentissage, a été évaluée par l'indice *alpha de Cronbach*, obtenant les valeurs 0.90 et 0.91, respectivement. Lors de l'évaluation de la cohérence du questionnaire, en ne considérant que les 15 items liés aux questions de l'évaluation, les indices obtenus étaient respectivement de 0.75 et 0.82. Étant donné le petit nombre d'items, ces niveaux de fiabilité peuvent être considérés comme acceptables, voire, tout à fait acceptables.

Les données obtenues dans la *phase de l'étude extensive*, de nature quantitative, ont été analysées en utilisant une variété de techniques statistiques. Par exemple, Barreira, Bidarra, Vaz-Rebello, Monteiro & Alferes (2015a) ont utilisé la différence standardisée entre les moyennes (*d* de Cohen) pour comparer les réponses données par les enseignants et les étudiants à chacun des items du questionnaire. En fait, pour qu'un nombre de degrés de liberté existe, toute différence entre les moyennes obtenues pour les enseignants et les étudiants, même réduite, serait toujours considérée comme statistiquement significative. Les différences entre les moyennes, calculées par le *d* de Cohen, sont considérées comme de petite taille (effet faible) lorsque  $|d| < .20$ ; comprises entre l'effet de petite taille et l'effet de taille moyenne quand  $.20 \leq |d| < .50$ ; comprises entre l'effet de taille moyenne et l'effet de grande taille quand  $.50 \leq |d| < .80$ ; de très grande taille quand  $|d| > .80$ . Pour une consultation approfondie au sujet de la recherche pendant cette phase de l'étude voir, par exemple, Barreira *et al.* (2015a) et Fernandes (2015).

L'enquête des pratiques d'évaluation des enseignants a eu lieu lors de la *phase intensive de l'étude*, basée sur des observations en classe, des entretiens approfondis avec des enseignants et des

entretiens avec des groupes de discussion d'étudiants. À ce stade, 35 enseignants et 168 étudiants ont participé volontairement. Des cours de nature pratique ou théorico-pratique de huit enseignants ont été observés dans chaque université. Ces enseignants étaient responsables des matières structurantes des cours de licence dans chacun des domaines de connaissance suivants : *sciences sociales, arts et lettres, sciences de la santé et sciences et technologies*. Ainsi, des cours de deux enseignants ont été observés par domaine de connaissance et dans chaque université. Étant donné que chaque enseignant a été observé pendant environ 20 heures, environ 160 heures d'observation des cours des unités d'enseignement, par domaine de connaissance (40 heures par université), ont ainsi été réalisées. De ce fait, au total, et dans tous les domaines de connaissance, environ 640 heures de cours ont été observées. De plus, des entretiens approfondis avec chacun des enseignants participants ont été réalisés, ainsi que des entretiens avec les groupes de discussion de trois à cinq étudiants chacun.

Le processus d'intégration et de transformation des données, s'inspirant des recommandations de Wolcott (1994), avait comme base la production successive de récits, en prenant en compte les tailles indiquées dans une *matrice de recherche* de l'étude, qui ont guidé l'analyse des données qualitatives concernant l'*évaluation* et les objets restants. La *matrice* et l'analyse des récits produits, faite par différents groupes de chercheurs, se sont avérées essentielles pour assurer leur cohérence et crédibilité.

Dans un premier temps, trois récits ont été produits par enseignant : un sur la base des observations des cours, un autre sur l'entretien approfondi et l'autre sur l'entretien avec chaque groupe de discussion d'étudiants. Ces trois récits ont été intégrés, donnant lieu à un récit lié aux pratiques d'enseignement et d'évaluation de chaque enseignant. Ainsi, dans chaque université, et par domaine de connaissance, deux récits ont été produits sur les pratiques des deux enseignants qui enseignaient des matières dans ce domaine. En poursuivant l'intégration des données, ces deux récits ont donné lieu à un seul *récit intégré des enseignants* de chaque domaine de connaissance. Quatre *récits intégrés* ont ainsi été produits dans chaque université, un par domaine de connaissance. Enfin, à partir de chaque groupe de quatre récits, un *métarécit* a été construit, en intégrant les pratiques des huit enseignants de chaque domaine de connaissance. Chaque *métarécit* a été construit de manière interactive entre les différents chercheurs, basé sur la *matrice d'évaluation* et les *récits intégrés*. Afin de garantir des niveaux acceptables et crédibles de cohérence et de crédibilité, tous les récits ont été retournés aux participants pour leur analyse et validation.

Le processus d'analyse, d'intégration et de transformation des données n'a pas été algorithmique et n'a pas obéi à tout *système classique* d'analyse de contenu. Une approche trop analytique a été évitée, en développant un travail plus artisanal dans l'identification des modèles, des régularités, des écarts notables, des situations singulières et suivant une perspective inductive de l'organisation et de l'analyse des données.

En bref, la production des quatre *métarécits* a donné lieu à des conditions permettant de décrire, d'analyser et d'interpréter les pratiques pédagogiques d'évaluation et d'enseignement des enseignants participants.

## 4. Résultats

La présentation des résultats a été organisée en trois sections. La première section fait référence à l'analyse des données quantitatives obtenues à partir des questionnaires, tandis que la deuxième est dédiée à l'analyse des données de nature qualitative, résultant de l'observation des cours et des entretiens. Enfin, la troisième comprend la discussion des données recueillies dans les sections précédentes.

### 4.1 Analyse des données quantitatives

Selon les valeurs indexées par le  $d$  de Cohen (cf. tableau 1), dans le domaine de l'évaluation, cinq différences (items 1, 2, 9, 10 et 13) sont inférieures au seuil conventionnel *effet faible* ( $|d| < .20$ ), neuf différences (items 3, 4, 5, 6, 7, 11, 12, 14 et 15) sont comprises entre les seuils *effet faible* et *effet moyen* ( $.20 \leq |d| < .50$ ) et une différence (Item 8) est située entre les seuils *effet moyen* et *effet fort* ( $.50 \leq |d| < .80$ ). Aucune des différences ne dépasse le seuil conventionnel *effet fort* ( $|d| \geq .80$ ).

**Tableau 1.** Comparaison enseignants - étudiants dans les items liés à l'évaluation.

Items du domaine d'évaluation	Étudiants	Enseignants	<i>p value</i>	<i>d-Cohen</i>
	(n=990)	(n=4568)		
	$\bar{x}$	$\bar{x}$		
1. Dans la plupart des matières / unités d'enseignement, l'évaluation se réalise de façon continue tout au long du semestre.	.34	.36	.872	.01
2. Les enseignants discutent assez souvent, et tout au long du semestre, de l'organisation et du développement du processus d'évaluation avec les étudiants.	-.29	-.21	.007	.06
3. Les critères d'évaluation à utiliser dans chaque matière / unité d'enseignement sont connus de la plupart des étudiants.	.91	1.33	.000	.47
4. Dans le processus d'évaluation, une variété de tâches et / ou d'instruments (par exemple, des épreuves, des tests, des devoirs, des rapports, des présentations, des recherches) est souvent utilisée.	.53	.86	.000	.31
5. Dans la plupart des matières / unités d'enseignement, l'évaluation est principalement utilisée pour attribuer des notes / une moyenne de l'ensemble des notes de l'évaluation aux étudiants.	1.26	.99	.000	.33
6. L'évaluation est souvent utilisée pour guider les étudiants à mieux apprendre.	-.13	.10	.000	.20
7. Les enseignants utilisent systématiquement l'évaluation pour que les étudiants prennent conscience de leurs difficultés et / ou de leur progrès d'apprentissage.	.00	.30	.000	.26
8. En général, l'évaluation réalisée est conforme à l'enseignement et aux apprentissages.	.41	1.01	.000	.64
9. Les étudiants sont souvent impliqués dans des situations où ils doivent évaluer leur propre travail et / ou celui de leurs collègues.	-.55	-.49	.003	.05
10. Les enseignants, en général, offrent aux étudiants la possibilité de clarifier et / ou remettre en question les jugements formulés au sujet de leurs apprentissages.	.01	.22	.000	.18

Items du domaine d'évaluation	Étudiants	Enseignants	p value	d-Cohen
	(n=990)	(n=4568)		
	$\bar{x}$	$\bar{x}$		
11. Dans la plupart des matières / unités d'enseignement, les résultats obtenus par les étudiants dans les tests et / ou les examens finaux sont les éléments les plus décisifs pour leur moyenne de l'ensemble des notes de l'évaluation.	1.37	1.08	.000	.35
12. Dans la plupart des matières / unités d'enseignement, les tests et / ou les examens finaux sont la meilleure façon de prendre connaissance de ce que les étudiants savent et sont capables de faire.	-.11	.20	.000	.25
13. Les enseignants encouragent souvent les étudiants à analyser et / ou à réfléchir sur les apprentissages réalisés.	.01	.23	.000	.19
14. En général, les informations issues des évaluations sont utilisées par les étudiants pour guider et / ou réorienter leur façon d'étudier.	.32	-.04	.000	.32
15. Pour la plupart des étudiants, l'évaluation est essentiellement un processus de réalisation des tests et / ou des examens visant à attribuer une moyenne de l'ensemble des notes de l'évaluation.	1.18	.93	.000	.28

Ainsi, des positions convergentes entre les enseignants et les étudiants sont enregistrées. En effet, ils acceptent *que l'évaluation se développe de façon continue tout au long du semestre* (Item 1), que les enseignants *offrent aux étudiants la possibilité de clarifier et / ou remettre en question les jugements formulés au sujet de leurs apprentissages* (Item 10) et *qu'ils les encouragent à réfléchir ou à analyser les apprentissages réalisés* (Item 13).

À leur tour, aussi bien les enseignants que les étudiants sont en désaccord sur le fait *que les enseignants discutent assez souvent, et tout au long du semestre, l'organisation et le développement du processus d'évaluation* (Item 2) et que *les étudiants sont souvent impliqués dans des situations où ils doivent évaluer leur propre travail et / ou celui de leurs collègues* (Item 9).

Toutefois, des positions moins convergentes se manifestent aussi, les enseignants étant plus d'accord que les étudiants sur le fait que *l'évaluation effectuée est conforme à l'enseignement et aux apprentissages* (Item 8), que *les critères d'évaluation sont connus par la plupart des étudiants* (Item 3), qu'*une variété de tâches et/ou d'instruments est utilisée* (Item 4) et que *l'évaluation est utilisée par les enseignants pour que les étudiants prennent conscience de leurs difficultés et/ou de leurs progrès d'apprentissage* (Item 7). À leur tour, les étudiants sont plus d'accord que les enseignants sur le fait que *l'évaluation est principalement utilisée pour l'attribution des notes et des notations* (Item 5), que *les résultats obtenus dans les tests sont les éléments les plus décisifs dans l'attribution d'une moyenne de l'ensemble des notes de l'évaluation* (Item 11) et que *pour la plupart des étudiants, l'évaluation est essentiellement un processus de réalisation des tests et / ou des examens visant à attribuer une moyenne de l'ensemble des notes de l'évaluation* (Item 15).

Des réponses polarisées sont encore enregistrées. Ainsi les enseignants sont d'accord et les étudiants sont en désaccord sur le fait que *l'évaluation est utilisée pour guider les étudiants à mieux apprendre* (Item 6) et que, dans la plupart des matières, *les tests et / ou les examens finaux sont la meilleure façon de prendre connaissance de ce que les étudiants savent et de ce qu'ils sont capables de faire* (Item 12). À leur tour, les étudiants sont d'accord et les enseignants en désaccord sur le fait que *les informations résultant des évaluations sont utilisées par les étudiants pour guider et / ou de réorienter leur façon d'étudier* (Item 14).

Sur la base des données obtenues, il devient ainsi possible de regrouper les réponses des enseignants et des étudiants en des positions moins convergentes, celles-ci présentant le plus grand nombre d'items, suivies des positions convergentes et des positions polarisées. L'interprétation de ces positions sera reprise par la suite, lors de la discussion des données de nature qualitative et quantitative.

## 4.2 Analyse des données qualitatives

Cette section comprend la description et l'analyse des pratiques et des perceptions des enseignants et des étudiants à l'égard des dimensions de l'évaluation constantes dans la *matrice de la recherche*, à savoir : a) les fonctions et l'utilité de l'évaluation; b) la nature de l'évaluation; c) les tâches d'évaluation les plus fréquemment utilisées; d) la nature, la fréquence et la distribution du *feed-back*; e) l'utilisation des stratégies d'auto et d'hétéro-évaluation; f) le rôle des enseignants et des étudiants.

### 4.2.1 Fonctions et utilité de l'évaluation

Dans la plupart des cas, l'évaluation avait une fonction de validation et d'attribution d'une moyenne de l'ensemble des notes de l'évaluation, basée sur des moments spécifiques qui, comme les enseignants le mentionnent, servent à « vérifier si l'étudiant a acquis les connaissances » et à « valider ». Les étudiants ont identifié ces moments spécifiques comme la base de leur moyenne de l'ensemble des notes de l'évaluation : « l'évaluation est toujours présente pendant les cours, mais (...) ce que tu fais là sera ta note », « l'étudiant ne connaît que la moyenne des deux sessions d'évaluation ou de l'examen ».

Pourtant, certains étudiants ont admis que cette évaluation peut les encourager à se motiver pour étudier : « ce jour-là nous fait étudier davantage, ce jour-là nous donne la note », « si la note ne correspond pas à ce que j'attendais, c'est un facteur de plus pour que je travaille davantage. Si la note est celle que j'avais prévue, je garde la même stratégie d'étude ». Mais, lorsque les résultats des tests n'étaient pas positifs, ils généraient une démotivation car, pour certains étudiants, « celui qui ne réussit pas la première session d'évaluation finira par faire encore pire lors de la seconde ».

D'autres étudiants relèvent la moyenne de l'ensemble des notes de l'évaluation, la considérant comme un référentiel normatif de la performance, qui permet de connaître la « place » qu'ils occupent au *ranking* des notations, de voir leurs efforts reconnus ou pas reconnus : « il est important de savoir quelle est ma place au classement ». (...) « savoir où j'en étais par rapport aux autres étudiants ».

Dans la même ligne de pensée, les enseignants reconnaissent que les résultats exercent de l'influence sur l'attitude des étudiants, car « si les choses se passent bien, ils les prennent comme une incitation à maintenir le niveau », « si les choses se passent mal, ils les prennent comme une incitation à améliorer et, donc, ils aiment vraiment savoir quelle est leur note ».

Dans certaines unités d'enseignement, l'évaluation a également une fonction de régulation. Les étudiants affirment que l'évaluation sert à s'autoréguler, à « voir quelles sont nos difficultés et si nous comprenons ou ne comprenons pas la matière », à « identifier points forts et points faibles du travail », mais aussi à réguler l'action enseignante : « si tous les étudiants ressentent la même difficulté, c'est l'enseignant qui est à l'origine du problème, pas l'étudiant et, à travers cela, les enseignants peuvent améliorer leur propre travail ». Pour certains enseignants, l'évaluation est l'occasion de juger de leur propre enseignement : « si les notes sont toutes basses ou toutes élevées, quelque chose ne va pas et finalement cela se révèle être un miroir de la façon dont ils enseignent », « l'évaluation nous permet, à nous les

enseignants, de comprendre jusqu'à quel point nous passons l'information, au-delà de cette notion que nous avons des cours“.

#### 4.2.2 Nature de l'évaluation

En général, les données montrent que l'évaluation a comme objectifs principaux de noter et certifier les étudiants, une perspective qui, dans certaines unités d'enseignement, coexiste avec des pratiques informelles, éventuellement non délibérées, d'évaluation formative, orientées vers l'amélioration des apprentissages.

L'évaluation de nature formative était plus présente dans les cours pratiques, où les enseignants fournissaient un *feed-back*, reconnu par les étudiants comme étant de valeur, considérant qu'il serait nécessaire d'avoir plus de moments qui les guident pour améliorer leurs apprentissages car, comme ils l'ont dit, certains « enseignants ne font pas l'évaluation des devoirs présentés; si on avait un *feed-back*, entre le premier et le deuxième travail, on saurait quoi améliorer ». En revanche, les enseignants ont déclaré que « c'est difficile avec certains étudiants car leur capacité de reconnaître l'erreur est très faible; (...) nous avons un grand pourcentage d'étudiants qui réagit mal à la critique ».

En général, les étudiants essayaient de guider leur étude et leur travail selon le *feed-back* qu'ils recevaient. Les enseignants reconnaissent que les étudiants, « en grande majorité, ont utilisé cette information pour améliorer (...), nous remarquons qu'ils ont appris », un enseignant a souligné que l'évaluation / le *feed-back* des présentations était très utile car « à ce moment, ils font vraiment attention, ils prennent des notes et il nous est possible de voir qu'ils intègrent toutes ces suggestions dans la phase suivante ».

L'existence de différentes perceptions à l'égard de l'évaluation, a pu être constatée. Pour certains enseignants, c'est la collecte systématique d'informations « qui se tiendra à tous les instants, quand nous sommes avec les étudiants, (...) pour évaluer la performance de l'étudiant à chaque instant » ; elle était « faite lors de l'observation des cours, de l'assiduité des étudiants et de leur réaction par rapport aux tâches proposées ». Les étudiants reconnaissent cette collecte systématique d'informations, mais disent que l'évaluation « est très dépendante du travail final et que les enseignants finalement ne s'aperçoivent pas trop de ce que les étudiants ont appris et de ce qu'ils ont fait pendant le temps des cours ». En fait, nous avons pu constater que, en général, l'évaluation dite continue est ponctuelle, basée sur des tests et se mêlant à l'évaluation de nature sommative, destinée à attribuer une moyenne de l'ensemble des notes de l'évaluation. Selon les déclarations de certains enseignants et étudiants, ce sont des « évaluations périodiques, pour faire le point de la situation et puis il y a l'évaluation finale, où tout cela se confirme, le parcours », « elle consiste en deux sessions d'évaluation, une session au milieu du semestre et l'autre à la fin », « il y a un examen final qui comprend tous les modules, c'est-à-dire, toutes les évaluations ponctuelles faites jusqu'à présent ».

Pour certains étudiants « c'est très négatif qu'il n'y ait pas d'évaluation continue car les étudiants, en faisant des sessions d'évaluation tout au long de l'année, seraient en mesure de mieux détecter leurs difficultés », « il devrait y avoir plus d'évaluations pour une meilleure distribution des notes et un meilleur suivi », « l'évaluation [par examen] est injuste, car on peut connaître la matière, s'être vraiment efforcé, puis l'examen ne se passe pas bien », « une évaluation plus continue serait l'idéal ».

#### 4.2.3 Tâches d'évaluation les plus couramment utilisées

Les tâches d'évaluation sont alignées globalement avec la nature, plus théorique ou pratique, des unités d'enseignement. Lorsque celles-ci étaient de nature théorique, l'évaluation relevait de la réalisation des épreuves écrites (examens, tests / sessions d'évaluation), des devoirs écrits et / ou des exposés oraux. Diverses tâches ont été identifiées dans les unités d'enseignement de nature pratique, ou avec un composant pratique, certaines étaient plus conventionnelles, telles que les travaux de recherche et d'enquête, les exposés et les rapports de laboratoire, l'analyse de documents, la présentation orale du travail; d'autres étaient plus innovantes et / ou spécifiques du domaine de connaissance, en particulier les *jeux-questionnaires*, les concours photo, *trivial pursuit*, la construction de maquettes, les dessins, la préparation d'un article de recherche, le portefeuille numérique, les devoirs individuels et les projets informatiques.

Les travaux de groupe ont également été largement utilisés, mais les enseignants ont estimé qu'il était nécessaire d'élaborer des tâches individuelles d'évaluation « qui doit être juste avec le bon étudiant, avec l'étudiant moyen et avec l'étudiant médiocre qui est dans le groupe et qui n'a rien fait. (...) il est important qu'ils soient évalués individuellement ». L'avis de certains étudiants va aussi dans ce sens. Ils affirment que « les évaluations, des devoirs, sont vraiment les plus subjectives, les sessions d'évaluation sont la meilleure façon de tester les connaissances de chacun. C'est le plus juste. Plus rigoureux. ». Cependant, ils estiment que les devoirs favorisent l'apprentissage, « on étudie plus de matières que si on devait étudier pour une session d'évaluation », « sans les devoirs, je ne connaîtrais pas la moitié de la matière ».

Outre les tâches d'évaluation, l'assiduité, la ponctualité, la participation et l'intérêt des étudiants à réaliser les tâches sont aussi des paramètres prévus dans l'évaluation. Cependant, les enseignants ont révélé des difficultés dans leur mise en œuvre, puis, finalement, elles n'ont, en général, pas d'impact sur la moyenne finale. Comme l'ont révélé certains répondants, « ce n'est pas évident, il m'est difficile de faire la différence entre un étudiant 14 et un étudiant 18 rien qu'à partir de la façon dont il répond en classe ou comment il travaille », ce qui mène au choix d' « une note globale évaluant la performance des étudiants en classe ».

#### 4.2.4 Nature, fréquence et distribution du *feed-back*

Dans toutes les unités d'enseignement, nous avons observé que les enseignants avaient distribué un certain *feed-back* (en particulier par voie orale) entre les étudiants « pour qu'ils aient une idée de ce qu'ils peuvent améliorer ». Nous avons constaté que la correction des épreuves effectuée en classe est la plus utilisée et appréciée, par rapport à d'autres formes de *feed-back* (par exemple, des commentaires écrits, des observations, des notes). Comme l'ont souligné certains enseignants, « faire passer un test, puis ne pas faire la correction pour leur montrer ce n'est pas correct, n'a pas de sens ». Les étudiants ont également apprécié cette modalité de *feed-back*, considérant qu' « il est important de savoir ce qui n'était pas correct et ce qu'il faut changer », « cela nous a donné une idée sur nos difficultés ».

Dans les rares cas où le *feed-back* était pratiquement inexistant en classe, les enseignants présentaient comme arguments le nombre élevé d'étudiants ou les changements requis par le processus de Bologne (par exemple, moins d'heures de contact).

#### 4.2.5 Utilisation des stratégies d'autoévaluation et d'hétéro-évaluation

D'une manière générale, les pratiques d'évaluation ne comprenaient pas des stratégies d'autoévaluation et d'hétéro-évaluation. Pour la plupart des enseignants interviewés, le « nombre élevé d'étudiants » et « l'extension des programmes » étaient les arguments présentés pour justifier la non-utilisation de ces stratégies. Cependant, dans certains cas, les étudiants étaient encouragés à réfléchir sur le travail développé, sans comprendre si cette stratégie serait délibérée ou résultant d'une intuition de l'enseignant, combinée à la nature plus pratique de l'unité d'enseignement.

Pourtant, des pratiques d'autoévaluation et d'hétéro-évaluation ont été identifiées, les étudiants ayant l'occasion d'évaluer leur performance et celle des collègues, avec ou sans l'aide de l'enseignant. Des moments de réflexion conjointe étaient ainsi créés, permettant aux étudiants de discuter le travail présenté, de donner des suggestions d'amélioration et de souligner les points forts. Comme un étudiant le disait, « on a vérifié que certains commentaires se faisaient et qu'ils reflétaient l'évaluation réalisée par les étudiants sur leur propre performance et la performance de leurs collègues ».

Certains enseignants ont mentionné que les étudiants avaient du mal à faire de l'hétéro-évaluation, car ils semblaient avoir « une certaine crainte de commenter ou de poser des questions aux collègues », mais ils perçoivent cette procédure comme bénéfique : « il est très important de connaître l'opinion des gens qui sont 'dehors' », « il est toujours gratifiant d'entendre les opinions des collègues ».

#### 4.2.6 Rôle des enseignants et des étudiants

Dans l'ensemble, les enseignants assumaient une position plus importante dans le processus d'évaluation car le rôle de l'organisation, de la sélection et de la définition du processus d'évaluation leur revenait, dans la plupart des cas. Ils procédaient à l'évaluation du travail et de la performance des étudiants et fournissaient le *feed-back* approprié. Dans certaines unités d'enseignement, les étudiants avaient l'occasion d'intervenir, pouvant discuter, voire négocier les aspects liés à l'évaluation (par exemple, les dates des tests et de la présentation des devoirs).

Le rôle essentiel des étudiants était de se préparer pour les tests d'évaluation (par exemple, les épreuves des sessions, les examens, la présentation des devoirs). La correction des épreuves en classe était une pratique suivie par certains enseignants et, parfois, les étudiants avaient l'occasion de discuter des aspects liés à l'évaluation, par rapport auxquels ils n'étaient pas d'accord : « nous avons la possibilité de discuter de l'examen, mais ce n'est pas la note ».

## 5. Discussion des données

Il nous est possible de mentionner, à partir des données obtenues, que les enseignants et les étudiants perçoivent les pratiques d'évaluation comme étant essentiellement sommatives, les enseignants reconnaissant plutôt le potentiel formatif de ces mêmes pratiques. La modalité d'évaluation la plus courante est l'évaluation sommative, orientée vers la moyenne de l'ensemble des notes de l'évaluation et la certification des étudiants, tandis que l'évaluation formative est plutôt résiduelle et presque confinée aux cours de nature plus pratique. Les tâches d'évaluation sont centrées principalement sur l'enseignant et directement liées à la nature des unités d'enseignement, l'utilisation d'une plus grande variété d'activités se vérifiant dans les cours pratiques.

En effet, il y a une convergence de la part des enseignants et des étudiants à l'égard des pratiques d'évaluation présentant une base continue tout au long du semestre. Bien que dans

la plupart des unités d'enseignement l'évaluation continue soit prévue, cette modalité n'est pas effectivement mise en œuvre, les pratiques d'évaluation se trouvant entre une perspective formative (moins évidente) et une perspective clairement sommative, associée à la réalisation de deux ou plusieurs tests ou devoirs, à des moments différents, dont le *feed-back* est fourni, généralement à la fin de l'unité d'enseignement, sous forme de notation quantitative. D'autre part, il y a la perception d'une certaine participation des étudiants à des situations d'auto-évaluation et d'évaluation par les pairs, sans qu'il n'y ait une ouverture aux pratiques d'évaluation formatives alternatives (Fernandes, 2006, 2015).

En ce sens, les enseignants ont exprimé des difficultés à réaliser l'évaluation, en particulier à mettre en œuvre des stratégies d'évaluation favorisant l'amélioration des apprentissages, soit parce qu'ils ne se sentent pas à l'aise avec ce qu'ils considèrent être un processus complexe, soit en raison de la difficulté à individualiser, compte tenu du nombre élevé d'étudiants dans certaines unités d'enseignement (Cid, Fialho, Borralho, Fernandes, Rodrigues & Melo, 2015).

Les données collectées révèlent que le paradigme traditionnel de la méthode expositive dans l'enseignement, avec une évaluation finale par test et examen, est encore très présent. Cependant, lors de l'analyse des perceptions des enseignants et des étudiants, selon le domaine de connaissance (arts et lettres, sciences de la santé, sciences et technologies et sciences sociales), certaines différences sont à souligner. À cet égard, Rodrigues & Soromenho (2015) concluent que, bien que les cours supérieurs en sciences de la santé et en sciences et technologies soient plus orientés vers un paradigme centré sur l'enseignant, les cours supérieurs en arts et lettres ou en sciences sociales adoptent un paradigme plus centré sur l'étudiant, attendu que, dans la pratique, nous semblons être en présence d' "approches multi-paradigmatiques" (p. 412). Ainsi, il est possible, par exemple dans le domaine des sciences et des technologies, de constater que, au cours de la réalisation d'un travail de nature pratique et / ou de laboratoire, l'évaluation assumait des fins de formation claires, dans la mesure où les étudiants recevaient en permanence un *feed-back* sur leur travail, des explications leur étant données et des recommandations formulées, ce qui leur permettait d'améliorer les apprentissages (Barreira, Vaz-Rebelo, Bidarra & Monteiro, 2015b).

En guise de synthèse, dans les matières plus théoriques, les principales fonctions d'évaluation, mises en évidence par les enseignants et les étudiants, étaient les suivantes : a) vérifier si les objectifs et les connaissances prévues dans le programme avaient ou n'avaient pas été atteints; b) noter et valider les apprentissages des étudiants. Par contre, dans les matières de nature plus pratique, outre les fonctions de validation et de notation, d'autres fonctions ont également été mentionnées : a) identifier les étudiants qui avaient besoin d'un plus grand suivi de la part de l'enseignant; b) réguler et améliorer l'enseignement et les apprentissages (Cid *et al.*, 2015).

## 6. Conclusions et réflexions finales

Les principaux objectifs de la recherche sur laquelle cet article repose étaient de décrire, d'analyser et d'interpréter les pratiques d'évaluation des enseignants universitaires et les perceptions des étudiants et des enseignants au sujet de ces pratiques.

L'une des conclusions de cette recherche est conforme à la littérature examinée qui affirme que les pratiques d'évaluation qui se produisent en salle de classe sont liées à plus et à de meilleurs apprentissages de la part des étudiants (par exemple, Albertino & De Sousa, 2004; Klenowski *et al.*, 2006). C'est-à-dire que la nature du travail pédagogique peut générer des environnements où l'enseignement, l'évaluation et l'apprentissage sont étroitement liés et organisés pour soutenir et améliorer la qualité de ce qui est appris.

L'analyse des données a montré que les enseignants utilisent une grande variété de pratiques d'évaluation qui peuvent, en termes généraux, être organisées en deux groupes. Dans un groupe, nous avons des pratiques plus cohérentes avec l'idée de l'enseignant, au centre des processus pédagogiques et, dans l'autre groupe, les pratiques plutôt liées à l'idée des étudiants au centre de ces processus. Dans le premier cas, clairement prédominant, les pratiques d'enseignement plus techniques et normalisées prévalent, avec la parole de l'enseignant, la passivité des étudiants et l'évaluation orientée vers la moyenne de l'ensemble des notes de l'évaluation. Dans le second cas, minoritaire, la participation des étudiants, les pratiques pédagogiques plus réflexives et critiques et un enseignement et une évaluation plus innovants au service de l'apprentissage, étaient présents. Quelles que soient leurs pratiques, les enseignants semblaient agir plus en fonction de leur intuition que de leur connaissance pédagogique, mobilisée délibérément. Cette constatation soulève, naturellement, plusieurs questions dans le domaine de la formation des enseignants et du développement professionnel des enseignants de l'enseignement supérieur.

Les résultats de cette recherche sont cohérents avec ceux obtenus par Berg, Admiraal et Pilot (2006), en ce qui concerne les avantages que les étudiants tirent de leur participation à des processus d'auto-évaluation et d'évaluation par les pairs, car ceux-ci se concentrent mieux sur ce qu'ils doivent apprendre, améliorant leur niveau d'attention et de motivation. De même, et tel que vérifié par Nicol (2009), les étudiants ont apprécié les moments de réflexion et de bilan fournis par ces types d'évaluation.

L'analyse des métarécits a également suggéré que l'articulation de la nature des tâches, de la structure des classes et des dynamiques de travail crée des contextes plus favorables aux apprentissages des étudiants. La relation pédagogique, qui implique la distribution systématique d'un *feed-back* de qualité, a été identifiée comme un processus qui a permis l'encadrement des étudiants, leur induisant des sentiments de sécurité, de confiance et d'estime de soi, essentiels au développement de leur autonomie et à leur implication dans les processus d'apprentissage. La littérature a permis de vérifier ce type de relations qui doit être étudié et compris plus profondément (Fialho & Fernandes, 2012).

La distribution d'un *feed-back* a été très appréciée par les étudiants, en raison des lignes d'orientation qui leur étaient fournies pour clarifier ce qu'ils avaient à apprendre, la situation dans laquelle ils se trouvaient et les efforts qu'ils devaient entreprendre pour réguler et autoréguler leurs apprentissages. Comme le montre la recherche de Trotter (2006), dans l'ensemble, le *feed-back* s'est avéré reposer aussi sur des critères utilisés à des fins de formation mais aussi pour attribuer une moyenne de l'ensemble des notes de l'évaluation.

La recherche a montré que, en général, les pratiques d'enseignement sont encore loin de celles qui sont recommandées dans la littérature et par l'agenda pédagogique découlant du *processus de Bologne*. L'enseignement est un processus qui est toujours associé à l'idée de *dire* et l'évaluation est un processus qui se rapporte aux enseignants et qui, en fin de compte, signifie noter. C'est-à-dire, la *pédagogie centrée sur les enseignants* continue de prévaloir dans les processus universitaires d'éducation et de formation.

Cependant, quelles que soient les autres conclusions qui pourraient être inférées de cette recherche, nous avons identifié un ensemble de caractéristiques qui, selon les données obtenues et la revue de la littérature, est associé à la création d'environnements pédagogiques qui favorisent les apprentissages. Ces environnements sont, par exemple : a) des cours avec des structures bien définies et induisant la participation active des étudiants; b) une sélection rigoureuse d'une variété de tâches pour apprendre, enseigner et évaluer, qui doit être travaillée dans diverses dynamiques de la salle de classe; c) la distribution systématique et délibérée du

*feed-back*, aussitôt que possible, après la réalisation des tâches par les étudiants; d) l'utilisation de l'évaluation formative pour soutenir et améliorer les processus d'enseignement et d'apprentissage; e) l'établissement de bilans des apprentissages réalisés. Ces approches ont été utilisées par les enseignants des quatre domaines de connaissance, rendant ainsi possible le développement de pratiques curriculaires plus cohérentes pour améliorer les processus d'éducation et de formation des jeunes étudiants.

## 7. Références bibliographiques

- Albertino, F., & De Sousa, N. (2004). Avaliação da aprendizagem: O portefólio como auxiliar na construção de um profissional reflexivo. *Estudos em Avaliação Educacional*, 29, 169-189.
- Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education*, 31(3), 341-356.
- Barreira, C., Bidarra, G. M., Vaz-Rebello, P., Monteiro, F. & Alferes, V. (2015a). Percepções de docentes e estudantes de universidades portuguesas sobre ensino, aprendizagem e avaliação. In Fernandes et al (Orgs.), *Avaliação, ensino e aprendizagem no ensino superior em Portugal e no Brasil: Realidades e perspetivas*, Vol.1 (pp. 309-325). Lisboa : EDUCA – Instituto de Educação da Universidade de Lisboa.
- Barreira, C., Vaz-Rebello, P., Bidarra, G. M. & Monteiro, F. (2015b). Práticas pedagógicas no ensino superior: Uma narrativa sobre ensino, aprendizagem e avaliação na área de engenharias e tecnologias. In Fernandes et al (Orgs.), *Avaliação, ensino e aprendizagem no ensino superior em Portugal e no Brasil: realidades e perspetivas*, Vol.1, (pp. 185-226). Lisboa : EDUCA – Instituto de Educação da Universidade de Lisboa.
- Berg, I., Admiraal, W. & Pilot, A. (2006). Design principles and outcomes of peer assessment in higher education. *Studies in Higher Education*, 31(3), 341-356.
- Bryan, C. & Clegg, K. (Org.) (2006). *Innovative assessment in higher education*. New York: Taylor and Francis.
- Cid, M., Fialho, I., Borralho, A., Fernandes, D., Rodrigues, P. & Melo, B. (2015). A avaliação nas práticas curriculares em quatro universidades portuguesas. In Fernandes et al (Orgs.), *Avaliação, ensino e aprendizagem no ensino superior em Portugal e no Brasil: realidades e perspetivas*, Vol.2, (pp. 615-648). Lisboa : EDUCA – Instituto de Educação da Universidade de Lisboa.
- Dancer, D. & Kamvounias, P. (2005). Student involvement in assessment: a project designed to assess class participation fairly and reliably. *Assessment and Evaluation in Higher Education*, 30(4), 445-454.
- Escudero, J. M. (2012). Claros y oscuros del Espacio Europeo de Educación Superior como innovación educativa. In C. Leite & M. Zabalza (Coords.), *Ensino superior: Inovação e qualidade na docência* (pp. 5-34). Porto: Centro de Investigação e Intervenção Educativas da Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto.
- Fernandes, D. (2006). Para uma teoria da avaliação formativa. *Revista Portuguesa de Educação*, 19(2), 21-50.
- Fernandes, D. (2015). Pesquisa de percepções e práticas de avaliação no ensino universitário português. *Estudos em Avaliação Educacional*, 26(63), 596-629.
- Fernandes, D. & Fialho, N. (2012). Dez anos de práticas de avaliação das aprendizagens no Ensino Superior: uma síntese da literatura (2000-2009). In C. Leite & M. Zabalza (Cords.), *Ensino superior: inovação e qualidade na docência* (p. 3693-3707), Porto: Centro de Investigação e Intervenção Educativas da Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto.
- Fernandes, D., Rodrigues, P. & Nunes, C. (2012). Uma investigação em ensino, avaliação e aprendizagens no ensino superior. In C. Leite & M. Zabalza (Coords.), *Ensino superior: Inovação e qualidade na docência* (pp. 932-944). Porto: Centro de Investigação e Intervenção Educativas da Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto.
- Ghiglione, R. & Matalon, B. (2001). *O Inquérito. Teoria e prática* (4ª ed.). Oeiras: Celta Editora.
- Hill, M. & Hill, A. (2008). *Investigação por questionário* (2ª ed.). Lisboa : Edições Sílado, Lda.
- Jones, S. (1999). *Doing internet research: Critical issues and methods for examining the net*. Thousand Oaks: Sage.

- Klenowski, V., Askew, S. & Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment and Evaluation in Higher Education*, 3(3), 267-286.
- Macmillan, J. & Mclean, M. (2005). Making first-year tutorials count: Operationalizing the assessment-learning connection. *Active Learning in Higher Education*, 6(2), 94-105.
- Miller, T. (2009). Formative computer-based assessment in higher education: The effectiveness of feedback in supporting student learning. *Assessment and Evaluation in Higher Education*, 34(2), 181-192.
- Mok, M., Lung, C., Cheng, D., Cheung, R. & NG, L. (2006). Self-assessment in higher education: Experience in using a metacognitive approach in five case studies. *Assessment and Evaluation in Higher Education*, 1(4), 415-433.
- Nicol, D. (2009). Assessment for learner self evaluation: Enhancing achievement in first year using learning technologies. *Assessment and Evaluation in Higher Education*, 34(3), 335-352.
- Poon, W.-Y., Mcnaughtb, C., Lamb, P. & Kwan, H. (2009). Improving assessment methods in university science education with negotiated self and peer-assessment. *Assessment and Evaluation in Higher Education*, 16(3), 331-346.
- Price, M. & O'Donovan, B. (2006). Improving performance through enhancing student understanding of criteria and feedback. In C. Bryan & K. Clegg (Orgs.), *Innovative Assessment in Higher Education* (pp. 100-109). New York: Taylor and Francis Group.
- Rodrigues, P. & Soromenho, G. (2015). Ensino, aprendizagens e avaliação nos diferentes domínios científicos e tecnológicos de quatro universidades portuguesas. In Fernandes et al (Orgs.), *Avaliação, ensino e aprendizagem no ensino superior em Portugal e no Brasil: Realidades e perspectivas*, Vol.1 (pp. 383-423). Lisboa : EDUCA – Instituto de Educação da Universidade de Lisboa.
- Trotter, E. (2006). Student perceptions of continuous summative assessment. *Assessment and Evaluation in Higher Education*, 31(5), 505-521.
- Walker, M. & Warhurst, C. (2000). In most classes you sit around very quietly at a table and get lectured at debates, assessment and student learning. *Teaching in Higher Education*, 1(1), 33-49.
- Wanous, M., Procter, B. & Murshidb, K. (2009). Assessment for learner self evaluation: Enhancing achievement in first year using learning technologies. *European Journal of Engeneering Education*, 34(1), 77-85.



# L'évaluation comme objet d'apprentissage et comme outil de développement professionnel dans le cadre de la formation des enseignants

## **Annick Fagnant**

Université de Liège  
[afagnant@ulg.ac.be](mailto:afagnant@ulg.ac.be)

## **Richard Etienne**

Université Paul Valéry Montpellier  
[Richard.Etienne@univ-montp3.fr](mailto:Richard.Etienne@univ-montp3.fr)

## **Lucie Mottier Lopez**

Université de Genève  
[Lucie.Mottier@unige.ch](mailto:Lucie.Mottier@unige.ch)

## **Marie-Noëlle Hindryckx**

Université de Liège  
[mn.hindryckx@ulg.ac.be](mailto:mn.hindryckx@ulg.ac.be)

## **Résumé**

Ce texte prend appui sur une table-ronde intitulée « *Former les futurs enseignants à l'évaluation et évaluer pour former : deux logiques complémentaires ?* ». En ancrant ses propos dans le contexte spécifique de la France et en s'appuyant sur une analyse des pratiques d'évaluation en salle de classe, la première partie du texte pose un premier cadre de réflexion et permet de dégager plusieurs enjeux importants pour la formation des enseignants à l'agir évaluatif. La suite du texte illustre deux dispositifs qui envisagent « l'évaluation comme objet d'apprentissage et comme outil de développement professionnel » : le premier est mis en place à Genève, dans le contexte d'une formation qui s'adresse à de futurs enseignants du primaire ; le deuxième se situe en Belgique francophone, dans le contexte d'une formation proposée à de futurs enseignants du secondaire supérieur. Le dernier volet du texte reprend quelques questionnements soulevés dans la partie introductive et cherche à les mettre en perspective avec des éléments pointés dans les deux dispositifs présentés. Il se termine en proposant quelques pistes de réflexions en matière de formation à l'évaluation.

## **Mots-clés**

Formation à l'évaluation – Evaluer pour former - Portfolio – Didactique disciplinaire – Stages

**Pour citer cet article :** Fagnant, A., Etienne, R., Mottier Lopez, L. & Hindryckx, M-N. (2017). L'évaluation comme objet d'apprentissage et comme outil de développement professionnel dans le cadre de la formation des enseignants. *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 77-98.

## 1. Introduction

Ce texte prend appui sur une table-ronde intitulée « *Former les futurs enseignants à l'évaluation et évaluer pour former : deux logiques complémentaires ?* ». S'inscrivant dans la thématique générale du colloque, la table-ronde envisageait les liens entre « évaluations et apprentissages » sous deux angles complémentaires : « l'apprentissage de l'évaluation » (ou l'évaluation comme « objet » d'apprentissage) et « l'évaluation pour l'apprentissage » (ou l'évaluation comme « outil » de développement professionnel).

La double originalité de la table ronde était, d'une part, de chercher à relier ce double questionnement et, d'autre part, de croiser les regards de spécialistes issus du champ de l'évaluation en éducation et de la didactique disciplinaire. Organisée par le premier auteur de l'article et avec les trois autres co-auteurs comme intervenants, la table-ronde s'inscrivait dans le cadre des travaux du réseau thématique de l'ADMEE-Europe « EVADIDA » (<http://admee.ulg.ac.be/activites/reseaux-thematiques/evadida/>) dont un des objectifs est de faire dialoguer davantage les spécialistes de l'évaluation (souvent issus du champ des sciences de l'éducation) et les didacticiens disciplinaires (moins historiquement attachés aux questions d'évaluation). Le croisement des deux angles susmentionnés devait conduire *in fine* à interroger « le potentiel de l'évaluation (des pratiques enseignantes) comme *outil de formation* (des enseignants) à l'évaluation ». Le choix des intervenants permettait d'apporter des regards différents sur cette problématique, en fonction des spécificités de leur champ de recherche, du domaine de formation dans lequel ils interviennent et d'une voie d'entrée qui leur avait été demandée de privilégier. S'inscrivant dans le prolongement de cette table-ronde, ce texte ne se présente pas sous la forme d'un article classique. Il fait en effet le pari de rendre compte de ce format de communication qui permet de croiser le regard de différents spécialistes sur une même thématique.

La première partie du texte (rédigée par Richard Étienne) pose un premier cadre de réflexion résumé sous l'intitulé « des pratiques d'évaluation des enseignants à leur formation en matière d'évaluation ». En ancrant ses propos dans le contexte spécifique de la France et en s'appuyant sur une analyse des pratiques d'évaluation en salle de classe, ce premier volet permet de dégager plusieurs enjeux importants pour la formation des enseignants à l'agir évaluatif. Les deux parties qui suivent (rédigées respectivement par Lucie Mottier Lopez et par Marie-Noëlle Hindryckx) envisagent la problématique de « l'évaluation comme objet d'apprentissage et comme outil de développement professionnel » en illustrant des dispositifs mis en place dans des cours destinés à former de futurs enseignants. Le premier exemple se situe en Suisse romande, dans le canton de Genève, dans le contexte d'une formation qui s'adresse notamment à de futurs enseignants du primaire. Le deuxième exemple se situe en Belgique francophone, dans le contexte d'une formation qui s'adresse à de futurs enseignants du secondaire supérieur. Enfin, la dernière partie du texte (rédigée par Annick Fagnant, tout comme la présente introduction) reprend quelques questionnements soulevés dans la partie introductive et cherche à les mettre en perspective avec des éléments pointés dans les deux dispositifs présentés. Avec une partie intitulée « pour terminer sans conclure » (à nouveau proposée par Richard Étienne), le texte se clôture en proposant quelques pistes de réflexion en matière de formation à l'évaluation.

## **2. Des pratiques d'évaluation des enseignants à leur formation en matière d'évaluation**

En France, les futurs enseignants ne savent pas qu'ils vont évaluer, même si, au fil des ans, sont apparues quelques formations à cette pratique, que ce soit dans le cadre des didactiques ou dans celui, plus général, des formations naguère appelées « transversales ». De fait ils croient qu'ils vont enseigner ! Ils découvrent cette dimension du métier dans les premiers jours de travail, parfois dans la solitude d'un dimanche après-midi, avec un paquet de copies ou un tas de cahiers « du jour ». Il n'y a alors d'autre solution que le « courage du commencement », selon l'expression de Jankélévitch (1960). Victime heureuse mais aussi souvent malheureuse de l'évaluation, l'enseignante ou l'enseignant se pose la question d'être juste sans perpétuer l'impression d'arbitraire ressentie parfois pendant sa scolarité et son parcours universitaire. C'est alors que se posent les questions de la nature de l'évaluation, de sa finalité et des modalités de sa communication. Toutes ces questions appellent des réponses avant que ne soit traité le projet de formation à l'évaluation et que ne soit élaborée la problématique de l'évaluation de la formation, qu'elle porte sur les plans et leur mise en œuvre ou qu'elle s'intéresse à la compétence enseignante. Comment traiter les oppositions classiques et traditionnelles entre une évaluation certificative qui a envahi tout le champ scolaire et une formative qui risque de se transformer en jeu de dupes ? Comment enfin faire progresser les pratiques évaluatives pour les inscrire davantage dans une logique de « compétences collectives » (Le Boterf, 1994), gage d'une meilleure acceptation par les parents, les élèves et, finalement, le système éducatif ?

### **2.1 Le complexe de M. Jourdain**

Persuadés que la transmission est une évidence et que l'évaluation sanctionne « justement », enseignantes et enseignants se comportent dans une première phase de leur carrière, et parfois plus longtemps, comme le Bourgeois gentilhomme, ils évaluent, et dévaluent, sans toujours savoir ce qu'ils font et la conséquence de leurs actes (Merle, 2015). Le temps manque et manquera toujours pour former une enseignante ou un enseignant avant sa rencontre avec les élèves. La situation est d'ailleurs sans doute à son paroxysme dans le contexte français actuel. Dans les jours qui suivent son entrée en fonction, le jeune enseignant devra procéder à une première évaluation formelle. C'est alors que, sans préparation, le choc risque d'être rude : l'alternative entre la complaisance qui consiste à fausser le jugement pour ne pas décrocher les élèves et la rudesse d'un diagnostic qui souligne les erreurs, voire l'échec, met le professionnel et la personne dans une impasse dont l'issue semble être la fuite. « Nous sommes là pour enseigner, pas pour juger ». Telle est l'expression recueillie auprès de centaines d'enseignants français lors de dizaines de stages d'évaluation animés de 1985 à 2016.

En France, ils sont bien là pour juger puisque les instructions officielles stipulent que, dans les classes où se décide l'orientation, la notation chiffrée est de règle. Le déni du jugement révèle un malaise lié au fait que les notes traduisent en quelque sorte l'échec de l'enseignement et que la courbe de Gauss plus ou moins répartie autour de la moyenne alimente la « constante macabre » (Antibi, 2003).

### **2.2 L'évaluation et sa communication**

Comment se met en place ce que l'on peut qualifier de « machine infernale » ? Dans un premier temps, si l'on s'appuie sur la synthèse de Barbier sur l'évaluation (1984), on estimera qu'évaluer, c'est prononcer un jugement de valeur en établissant un écart entre un référent

(ce à quoi on se rapporte, un idéal, une norme, un objectif, voire un référentiel que l'on peut qualifier rapidement de catalogue des références disponibles) qui n'est pas forcément stable, stabilisé ou exprimé, déposé sur un support, et un référé (une performance, la copie d'un élève, un oral ou toute production, durable ou non). En accord avec cette perspective, que l'on évalue un repas, une œuvre d'art ou une copie d'examen, l'opération est la même. Elle est en partie mentale et, a priori, non observable.

Dès lors, en matière d'éducation et de formation, se pose la double question du codage et de la transmission de ce jugement si au moins deux personnes sont impliquées dans l'opération. La question n'est devenue cruciale qu'au moment où il fallu attester de la réussite à un examen ou à un concours.

Quelles conséquences de la note chiffrée en termes d'enseignement ? La mesure de l'écart entre le référé et le référent n'est guère utile pour le diminuer. La fréquence des mauvaises notes fait intérioriser l'échec et, en développant les réactions défensives, alimente la « peur d'apprendre » (Boimare, 1999). Pour l'immense majorité des élèves, l'enjeu devient d'obtenir la moyenne, alors qu'elle traduit, si elle est décernée de manière mathématique, une incompétence grave puisque seule la moitié des savoirs est en place ! Comment reconfigurer l'évaluation en cours d'apprentissage pour que la certification atteste de la réelle compétence ? Il s'agit là d'un enjeu important, non seulement pour l'école primaire et secondaire, mais aussi (et sans doute plus encore !) dans le cadre de la formation professionnelle des futurs enseignants.

### **2.3 « Apprendre à évaluer », n'est-ce pas apprendre et savoir ?**

Aujourd'hui, il y a une controverse entre celles et ceux qui considèrent qu'enseigner est un don et les tenants de la formation pour qui c'est « un métier qui s'apprend » (Étienne & Zamaron, 2005). Et c'est le cas aussi pour un « aspect caché » qu'est l'évaluation. Un numéro très récent d'*Éducation permanente* (Tourmen, 2016) développe cet enjeu majeur qu'est le « Apprendre à évaluer ». En effet, qu'on soit élève ou professeur, on ne peut pas apprendre sans apprendre à évaluer. Autrement dit, la peur de l'évaluation gagne à se métamorphoser en levier des apprentissages : ce qui est évalué, ce n'est pas l'apprendre mais les différents savoirs qui en sont l'issue, parfois incertaine. Même les « savoirs savants » (une dissertation ou un contrôle de physique) nécessitent des savoir-faire (produire une copie, des résultats corrects) quand ils sont l'objet d'une évaluation externe. Tout le monde connaît l'importance des savoir-faire et la mise en avant de l'observable comme support du travail d'évaluation (« une activité de l'apprenant identifiable par un comportement observable », Hameline, 1979, p. 62).

Deux autres types de savoir sont aujourd'hui plus recherchés : le savoir agir (« en situation complexe et singulière », définition de la compétence selon Le Boterf, 1994) et le savoir devenir (Boutinet, 1998). Tous deux relèvent du pilotage dans un monde fait de doutes et d'incertitudes. La réaction aux imprévus (Jean, 2009) y a plus d'importance que la mise en œuvre d'un programme bien conçu. Une telle « rupture épistémologique » (Nunziati, 1989) implique de refonder l'évaluation en partant du « réel de l'activité » (Clot, 1999) des élèves et des enseignants pour en modifier les contraintes en vue de mettre en place du « bon travail » (*ibid.*). Il ne s'agit pas de faire le deuil de la normativité, de la normalisation, mais de s'inscrire dans une recherche de la valeur propre à chaque situation singulière. Cette démarche ne peut s'envisager que dans une double contrainte d'analyse du réel et de créativité pour inventer les solutions répondant à la complexité saisie grâce à l'analyse.

#### **2.4 L'élaboration et la communication de l'évaluation dans la communauté de pratique et de discours**

Dès lors, l'évaluation ne peut plus être une simple sentence qui tombe de la personne supposée savoir sur la personne supposée apprendre ou, pour le dire avec encore plus de rigueur et de vigueur, l'élaboration et la communication du référent en amont de la demande de référé deviennent indispensables. Et, comme les recherches dans le domaine de l'information et de la communication ont établi que le *feed-back*, le retour énonciatif, la reformulation sont les plus propices à la transmission d'un message avec le moins de perte en ligne, l'émergence du référent sous une forme co-construite par l'un et l'autre sera l'occasion d'une meilleure appropriation du référent et de ce qui sera demandé pour produire le référé. À cela, on peut ajouter que l'appropriation est facilitée en cas de mise en situation-problème avec une solution et une méthode à trouver, une brève réflexion individuelle puis en travail en groupe afin de provoquer un ou des conflits socio-cognitifs (Bonniol & Vial, 1997). De plus, le travail en boucles itératives permet un cheminement progressif et sûr, tenant compte de la zone de prochain développement. Cette idée et cette pratique sont empruntées à deux sources : d'une part, la communauté de pratique de Lave et Wenger (1991) pour le versant « action » ; d'autre part, la communauté de discours de Bernié (2002) pour le versant « mise en mots partagée ». On pourra se reporter à Doly (2006) qui a travaillé de cette façon sur l'écriture du conte avec de nombreux enseignants et des classes du premier degré.

#### **2.5 Pour en terminer avec l'opposition entre certificative et formative**

Barbier (1984) dénonçait déjà l'opposition factice entre évaluation formative et certificative, terme préférable à celui de « sommative » puisque nous adoptons ce choix fait par Mottier Lopez (2015) en soulignant qu'une évaluation sommative peut avoir une finalité formative si elle est faite au début d'un cycle pour orienter les apprentissages. Ainsi sont articulés les moments de l'évaluation et distingués soigneusement le temps de la formation et le moment de la certification. Le « contrôle continu » a fait oublier cette distinction essentielle qui reconnaît le droit à l'erreur pendant l'apprentissage. Si on peut faire l'hypothèse que 95% du temps scolaire ou universitaire, voire du temps de formation professionnelle, devrait être réservé aux apprentissages, cela vient diminuer le temps strictement réservé à leur contrôle par des évaluations externes. D'où l'impérieuse nécessité d'un pilotage progressivement remis aux intéressés avec une éthique liant autoévaluation et autonomie, la première étant, quand elle est pertinente, cohérente et efficiente, la garante de la seconde.

Rester soumis à des évaluateurs externes ne peut être considéré comme un signe de compétence. Quelques dispositifs de communication de l'évaluation, comme l'élaboration de listes de critères de réalisation et de réussite, puis leur évolution et leur disparition, sont particulièrement propices à une auto- et co-formation à l'évaluation. Remettre les clés de l'évaluation aux intéressés est donc le principe à mettre en œuvre, même s'il n'est pas toujours des plus aisés. Le co-pilotage des apprentissages n'est destiné qu'à la mise en sécurité des personnes qui se forment et sa disparition est la conséquence de sa réussite.

#### **2.6 Analyser les situations d'évaluation pour se former et transformer l'évaluation en moteur de tous les apprentissages**

Dans Étienne et Fumat (2014) et Étienne (2016), il a été souligné l'intérêt d'une homothétie entre ce qu'il est opportun de faire vivre aux élèves, étudiants et personnes en formation professionnelle et la formation des professionnels chargés de cette mission. Ces éléments se retrouvent dans les recherches d'Allal notamment (1999). Qui évaluera l'évaluation ? Telle est la question qui permet de faire passer l'évaluation d'un prescrit dont nous devons

comprendre la nécessité en termes de commande à un remaniement en termes d'action : évaluer induit souvent une marge d'erreur et l'accès à l'autoévaluation permet à l'apprenant de ne plus dépendre d'un évaluateur extérieur.

Berger (1977) posait cette question : « Mais qu'est-ce qui nous prend à évaluer ? » La réponse reste d'actualité : « nous ne savons plus ce que nous faisons » disait-il. C'est dans l'après-coup, dans le débriefing, que nous nous apercevons que nous n'avons pas fait (référé) ce que nous voulions faire (réfèrent). Mais savions-nous vraiment ce que nous voulions faire ? Avons-nous pris conscience de ce qui nous a empêchés, individuellement ou collectivement, de le faire ? Nous avons bien des projets et des désirs et notre action ne se résume pas au déroulement d'un programme. C'est pour cela que nous évaluons plus que nous ne contrôlons.

C'est notamment en se donnant ces temps d'analyse après action, de retour réflexif, que se développe la compétence évaluative indispensable au développement du pouvoir d'agir. Transposée dans l'établissement scolaire en formation continue, cette démarche incite à utiliser ce dernier comme « une nouvelle aire de formation » (Ria, 2015). Pour ce qui est de la formation initiale, un tel constat inciterait à établir un lien fort avec des « établissements formateurs » dont nous continuons à proposer la mise en réseaux et la reconnaissance (Étienne, 1999, 2000).

### **3. L'évaluation comme objet d'apprentissage et comme outil de développement professionnel : deux exemples issus de contextes professionnels différents**

#### **3.1 L'exemple d'un portfolio mis en place dans un cours dédié à l'évaluation destiné à de futurs enseignants du primaire**

Dans le canton de Genève, les futurs enseignants du primaire sont formés à l'Université. Après une première année de tronc commun faisant partie du Baccalauréat en Sciences de l'éducation, ils doivent passer un concours pour être admis à la formation en enseignement primaire qui couvre une deuxième et une troisième années du Baccalauréat, puis une quatrième année de certificat complémentaire à l'enseignement primaire.

Le portfolio qui nous intéresse dans ce chapitre a été conçu pour l'Unité de Formation libre du Baccalauréat, « Evaluation et régulation des apprentissages », donnée par Lucie Mottier Lopez<sup>1</sup> Cette unité se déploie sur un semestre académique de 14 semaines, à raison de deux périodes de 45 minutes par semaine. L'effectif du cours varie entre 60 et 120 étudiants en fonction des années. Si la majorité des étudiants qui choisissent ce cours sont de futurs enseignants du primaire, le cours est ouvert à l'ensemble des étudiants en Sciences de l'éducation. Les objectifs du cours annoncés aux étudiants sont les suivants : « s'approprier les principaux concepts du domaine de l'évaluation des apprentissages ; réfléchir aux enjeux et problématiques éducatives qui sont liés à l'évaluation des apprentissages ; analyser des situations d'évaluation ; concevoir des procédures et des outils d'évaluation des apprentissages ». Au début de chaque cours, ces objectifs généraux sont déclinés en objectifs plus ciblés au regard des thématiques enseignées. Ils sont systématiquement présentés aux étudiants.

---

<sup>1</sup> Une étude étayée de la première expérience de ce portfolio est présentée dans Mottier Lopez (à paraître). Seuls quelques éléments sont repris dans ce chapitre, en ajoutant également quelques ajustements du dispositif qui a été ré-expérimenté en automne 2016.

La conception de ce portfolio visait les buts principaux suivants (Mottier Lopez, à paraître) :

- Proposer aux étudiants un dispositif d'évaluation qui soit cohérent avec les contenus enseignés ;
- Permettre aux étudiants d'expérimenter de nouvelles formes d'évaluation associées à diverses régulations formatives (entre pairs, avec l'enseignante, avec des ressources matérielles et symboliques), dans des temporalités différentes (immédiates pendant le cours, différées dans des formes autonomes d'apprentissage) ;
- Articuler les fonctions formative et certificative par une évaluation continue tout au long du semestre académique ;
- Inciter les étudiants à prendre conscience qu'une évaluation certificative complexe, sollicitant des compétences de haut niveau, peut représenter des opportunités de s'autoréguler à des fins de progression.

Ce portfolio s'inscrit dans la perspective des travaux de recherche sur une « évaluation pour apprendre » (*assessment for learning*) dans le prolongement de l'évaluation formative qui a pour fonction de produire des informations servant à la régulation de l'enseignement et de l'apprentissage : « *Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there* » (Assessment Reform Group, 2002).

Des « stratégies » ont été expérimentées dans les classes montrant l'efficacité de l'évaluation pour apprendre (e.g., Wiliam, 2011). Bien qu'il y ait moins de résultats de recherche disponibles pour l'enseignement supérieur, citons les stratégies encouragées pour ce contexte à partir des travaux de Wiliam<sup>2</sup> : (1) expliciter et partager avec les étudiants les objectifs et critères d'évaluation, (2) concevoir des dispositifs qui intègrent des discussions collectives, des stratégies de questionnement et des tâches qui donnent à voir les apprentissages réalisés, (3) proposer des feedback qui permettent aux étudiants de progresser, (4) inciter les étudiants à être des ressources pour leurs pairs, (5) encourager les étudiants à se sentir responsables de leur propres progressions d'apprentissage. Les sections suivantes illustrent brièvement comment le portfolio a opérationnalisé quelques-uns de ces principes – en les combinant – tout en ayant des attentes académiques de haut niveau.

### *3.1.1 Expérimenter, par soi-même et avec les autres, des formes variées d'évaluation pour apprendre*

Les étudiants devaient insérer trois types de pièces dans leur portfolio, dont certaines étaient à choisir et d'autres qui étaient obligatoires. Le premier type de pièces était associé à des « activités interactives » proposées par l'enseignante tout au long du semestre, souvent brèves, étroitement liées aux contenus enseignés le jour-même. Le but de ces pièces était d'inciter les étudiants à autoévaluer leur compréhension des contenus en jeu, et si besoin de mettre en œuvre des stratégies d'autorégulation pour progresser. Pour chaque activité interactive retenue, les étudiants devaient utiliser un canevas (voir annexe 1) sous forme de fichier informatique à télécharger. Ce canevas, comportant des rubriques ouvertes à renseigner par l'étudiant, visait à structurer la démarche réflexive à entreprendre. Le deuxième type de pièces était lié à deux travaux pratiques (TP) réalisés pendant le semestre. Suite à chacun des TP, des feedbacks systématiques étaient proposé à chaque duo ou trio d'étudiants ayant travaillé ensemble. Enfin, la dernière pièce était un texte à rédiger, visant à

---

<sup>2</sup> <http://www.itlearning.net/assessment-for-learning-5-key-strategies-for-higher-student-achievement>, consulté le 5 décembre 2016.

mettre en relation les différentes pièces du portfolio et à proposer une réflexion conceptuelle critique. L'annexe 2 présente les caractéristiques de chaque type de pièces et donne quelques exemples.

### 3.1.2 *S'autoévaluer pour définir ses objectifs personnels d'apprentissage*

Pour chaque type de pièces, l'enseignante présentait explicitement aux étudiants les objectifs d'apprentissage et critères d'évaluation. Toutefois, il ne suffit pas de communiquer les objectifs pour que les étudiants les comprennent et se les approprient à des fins d'autorégulation. Ce n'est pas assuré non plus que les objectifs définis par une instance externe (l'enseignante) soient pleinement alignés avec les besoins de l'étudiant. Une spécification, voire réorientation des objectifs peut être nécessaire pour orienter les stratégies d'autorégulation à des fins de différenciation. Concrètement, pour chacune des activités interactives, il était ainsi demandé à l'étudiant :

- par rapport aux objectifs annoncés par l'enseignante, de commencer par autoévaluer sa compréhension (au sens large du terme) dans le cadre de tâches précises impliquant des activités d'analyse, de déduction / induction, de mise en relation, de jugement, notamment ;
- d'identifier les contenus (partiellement) incompris ou pouvant faire l'objet d'un approfondissement ;
- puis sur cette base, de rédiger son (ses) objectif(s) *personnel(s)* d'apprentissage (Cartier, 2007), pointant des éléments à retravailler dans une perspective de régulation rétroactive (reprise de contenus non maîtrisés) ou à approfondir dans une perspective de régulation proactive (Allal & Mottier Lopez, 2005).<sup>3</sup>

### 3.1.3 *Encourager la mise en œuvre de stratégies de régulation et d'autorégulation à des fins de progression*

Une fois l'objectif personnel d'apprentissage formulé, la consigne était de mettre en œuvre des stratégies de régulation et de les décrire pour son portfolio en prenant appui sur le champ conceptuel de la régulation des apprentissages travaillés au cours, en lien avec l'évaluation formative. Une mise en relation entre les apports théoriques et l'expérience de l'étudiant était ainsi sollicitée. Les commentaires métacognitifs des étudiants de la première cohorte ayant pratiqué le portfolio ont fait l'objet d'une analyse de contenu (Mottier Lopez, soumis). L'étude a montré que les étudiants ont décrit en majorité des stratégies pendant la réalisation de l'activité interactive et des mises en commun associées (52%) ; puis des stratégies de régulation mises en œuvre hors du temps d'enseignement (41%), et enfin quelques stratégies exploitant d'autres contenus du cours ou en établissant des liens entre les différentes pièces du portfolio (7%). Il est apparu toutefois que la consigne n'était peut-être pas suffisamment contraignante pour véritablement encourager des stratégies d'autorégulation en profondeur. C'est pourquoi, dans la cohorte suivante d'étudiants (portfolio 2016), il a été demandé d'insérer également dans le portfolio des traces d'activité concrètes, y compris métacognitives, qui donnent à voir les stratégies d'autorégulation mises en œuvre.

---

<sup>3</sup> Cette dernière consigne n'existait pas dans la première expérience de portfolio analysée dans Mottier Lopez (à paraître). Elle a été ajoutée pour le portfolio de 2016, suite à cette analyse, afin de davantage soutenir l'apprentissage autorégulé des étudiants.

### 3.1.4. Un texte de mise en lien et d'intégration des concepts

Une des caractéristiques du portfolio est non seulement de consigner des documents dans un même support et de donner à voir des démarches d'autoévaluation (au sens large, Allal, 1999<sup>4</sup>) et d'autorégulation, mais aussi de permettre à l'apprenant de porter un regard sur l'ensemble des traces et de les mettre en relation pour réfléchir sur son cheminement et apprécier sa progression. Le texte final d'intégration et de mobilisation des concepts avait cette fonction. Lors de la première expérience, il avait été demandé aux étudiants d'effectuer un lien avec au moins un exemple concret présenté dans son portfolio et de le conceptualiser (Mottier Lopez, à paraître). Dans le portfolio de 2016, ce texte a pris davantage d'importance. Les deux objectifs, communiqués aux étudiants, étaient de rédiger (1) une réflexion critique sur l'ensemble des stratégies d'autorégulation objectivées dans le portfolio, (2) puis une réflexion conceptuelle sur la problématique de la relation entre les activités d'enseignement, apprentissage et évaluation en contexte scolaire (ou de formation), en s'appuyant également sur une lecture obligatoire.

Cette forme innovante d'évaluation représente une contrainte forte tant pour les étudiants qui doivent s'engager dans une évaluation continue et impliquante, que pour l'enseignante qui doit fortement structurer son cours et les démarches d'évaluation formative qui s'y rapportent<sup>5</sup>. Du point de vue des étudiants, l'obstacle principal observé à cette évaluation « pour se former à l'évaluation » porte sur la compréhension et l'acceptation du nouveau contrat didactique qu'elle sollicite, en rupture avec la représentation de certains étudiants de ce que devrait être un enseignement universitaire et ses modalités d'évaluation. Ainsi, certains étudiants ont-ils considéré<sup>6</sup> que l'enseignement était « trop interactif », « comme s'ils étaient à l'école primaire », et que l'évaluation « exigeait trop la présence des étudiants au cours ». Ces représentations font obstacle à l'apprentissage soutenu par l'évaluation. D'autres étudiants, au contraire, ont manifesté une réelle adhésion face à cette expérience qui était nouvelle pour eux et qui, à leurs dires, leur a notamment permis de mieux « comprendre les notions » enseignées et d'être motivés à « travailler pour le cours ».

### 3.2 L'exemple d'une formation à l'évaluation dans le cadre d'un cours de didactique de la biologie (et des stages associés) destiné à de futurs enseignants du secondaire supérieur

En Belgique francophone, la formation des enseignants qui s'adresseront à des élèves de 15 à 18 ans a lieu à l'Université, sous la forme de 30 crédits insérés dans la formation des Masters ou à leur suite. Cette formation ne comprend donc pas de formation disciplinaire à proprement parler, mais de la didactique des disciplines, de la pédagogie, de la psychologie et des sciences humaines et sociales. L'axe pratique comporte 40 périodes de stage à prester dans une école d'enseignement secondaire, sous la supervision d'au moins un enseignant titulaire de la classe (maître de stage) et des didacticiens disciplinaires (formateurs). L'axe théorique de la didactique disciplinaire comprend des cours théoriques (préparations de leçons, approche des référentiels, techniques de gestion de classe, épistémologie, rencontres

---

<sup>4</sup> Allal (1999) parle d'autoévaluation « au sens large » quand celle-ci inclut des modalités d'évaluation mutuelle entre élèves et de co-évaluation entre élève(s) et enseignant (autoévaluation de l'élève confrontée à l'évaluation de l'enseignant). L'autoévaluation « au sens strict » désigne de façon plus ciblée l'autoévaluation faite par un élève de son propre travail.

<sup>5</sup> Cette contrainte pour l'enseignante a été viable grâce à l'aide d'un assistant, Fernando Morales Villabona, qui en est remercié.

<sup>6</sup> Commentaires anonymes issus du questionnaire de l'évaluation des enseignements par les étudiants (décembre 2016) ou exprimés à l'occasion du bilan final faite lors du dernier cours avec l'enseignante.

d'acteurs de l'enseignement et des institutions...), des observations de contextes réels d'enseignement, des séances de pratiques réflexives...

C'est principalement lors des enseignements de didactique générale et de didactique disciplinaire de la biologie que l'évaluation est envisagée dans différentes dimensions. Ce sont les aspects abordés pendant le cours de didactique disciplinaire de la biologie qui font l'objet des propos ci-dessous. L'annexe 3 reprend les différents moments d'évaluation proposés aux étudiants lors de la formation ainsi que leurs visées pour le métier.

### *3.2.1. Une préoccupation des étudiants : évaluer les apprentissages des élèves en classe*

Contrairement aux étudiants français, une des premières préoccupations des étudiants en Fédération Wallonie-Bruxelles est de construire des questions d'évaluation pour jauger des productions - écrites la plupart du temps - de leurs élèves. Leur réflexe est de reproduire ce qu'ils ont vécu : des questions fermées ou des questions à choix multiples (QCM), quelques questions ouvertes à production courte, des « vrais ou faux »... Il faut, selon eux, vérifier les acquis des élèves après l'apprentissage pour quantifier cet apprentissage et dans une moindre mesure, l'enseignement.

Les étudiants pensent qu'une évaluation est nécessairement teintée d'objectivité et de rigueur ; représentation courante, sans doute encore accentuée par le fait qu'ils ont reçu une formation scientifique. Pourtant, à l'entame de leur formation, peu d'entre eux pensent à définir des attentes précises en termes d'acquis chez les élèves : ils rédigent des questions, parfois les réponses attendues, mais rarement les critères et les indicateurs qui seront nécessaires pour objectiver leurs évaluations. Ces éléments sont alors travaillés explicitement durant la formation, tant du côté des cours de didactique générale que de didactique disciplinaire, au travers de diverses mises en situation (corrections croisées, création de grilles de critères et d'indicateurs, évaluation par compétences et tâches complexes...) visant à déstabiliser leurs conceptions initiales, à leur faire prendre conscience de la complexité du processus évaluatif et à leur fournir des premiers outils techniques qui seront approfondis lors des préparations aux stages d'enseignement notamment.

### *3.2.2 L'évaluation des apprentissages des étudiants par les formateurs*

Les apprentissages des étudiants sont évalués par les formateurs dans le cadre des enseignements théoriques, mais aussi dans le cadre de leurs stages dans des classes du secondaire supérieur. Si l'évaluation dans le cadre des cours théoriques ne semble pas leur poser de problème, tant elle ressemble à leur vécu antérieur, l'évaluation de leurs pratiques en stage leur semble presque déplacée, totalement subjective, voire incongrue...

Pour l'évaluation de la pratique, une grille est utilisée par les formateurs lorsqu'ils assistent aux prestations en stage. Pour appréhender cette grille, les étudiants visionnent un extrait vidéo d'un enseignant lors d'un cours de sciences et notent « tout ce qui va » et « tout ce qui ne va pas », tant au niveau de l'enseignant que des élèves ou du contexte dans lequel la leçon est donnée. Ensuite, partant des observations des étudiants, les formateurs recomposent avec eux cette grille d'évaluation de la pratique. Chaque élément est décrit et discuté afin que les étudiants comprennent bien les attentes en matière de pratiques de classe et que la peur de l'évaluation se métamorphose en levier d'apprentissage. Cette grille contient, outre les éléments contextuels : les éléments liés à la préparation de la leçon (prescrits légaux, scénario prévu, anticipation des difficultés, outils bibliographiques...); les attitudes de l'enseignant stagiaire (expression relationnelle, verbale et non verbale, maîtrise de la langue, gestion du temps, de l'espace et des supports multimédias...); la gestion des apprentissages (prise en compte des représentations, consignes, matériel, bibliographie, méthodes

d'enseignement...); l'interaction avec la classe (interventions, feedbacks, remédiations, types d'évaluation, synthèses des apprentissages...); les attitudes des élèves (prise de note, participation, motivation...); les dispositifs spécifiques (expérimentations, projections multimédias, tableau noir, tableau blanc interactif, notes de cours...) leur pertinence, leur mise en œuvre; les concepts et contenus abordés ainsi que les illustrations utilisées (exactitude, quantité, niveau, ancrage historique, interdisciplinaire ou médiatique ...) et enfin, les éléments de prise de recul, d'analyse et de réflexivité.

### *3.2.3 L'évaluation des pratiques de classe, le recul sur ses pratiques et l'acquisition d'une posture réflexive*

Tout au long de leur formation, les étudiants sont entraînés à pratiquer une auto(hétéro) évaluation de leurs prestations d'enseignement. Au début de l'année, lors de séances appelées « micro-enseignements », chacun prépare et présente une leçon de sciences devant ses pairs, à l'Université. Il s'agit, pour la plupart, d'une première expérience d'enseignement, en contexte artificiel, mais bienveillant. Quand la prestation est finie, ceux qui y ont assisté (étudiants et formateurs) donnent leur avis argumenté sur celle-ci. L'étudiant doit alors rendre un rapport comprenant la préparation formelle de sa leçon, les modifications qu'il y apporterait, ainsi que l'ensemble des remarques qui ont été formulées. Il indique également s'il juge ces remarques pertinentes ou non et pourquoi. Toute la difficulté est de faire comprendre aux étudiants que ce n'est pas tant leur prestation (souvent première et donc peu concluante) qui est ici l'objet principal de l'évaluation, mais bien l'analyse qui en est faite, avec l'aide des pairs et des formateurs. Ce moment de réflexion est crucial dans la formation de l'étudiant pour que, au-delà de sa performance - concluante ou non<sup>7</sup> -, il identifie les enjeux de son action, les analyse et les régule de manière à s'engager dans « une transformation réflexive de l'expérience » (Schön, 1988, p. 25) qui agira sur leur action en situation lors de leurs stages à venir.

Lors des stages ultérieurs, une prise de recul réflexif du même ordre est effectuée après chaque leçon, avec l'aide du maître de stage, et doit figurer parmi les documents relatifs au stage. Lors de sa visite en classe de stage, le formateur peut donc consulter, en plus des notes ou appréciations du maître de stage relatives aux performances de l'étudiant en situation de faire la classe, les écrits portés sur les leçons précédentes.

Par la suite, une leçon de stage, au choix de l'étudiant, fait encore l'objet d'une autoscopie. Concrètement, le futur enseignant dispose au fond de la classe une caméra sur pied et filme sa prestation face aux élèves. Dans un premier temps, il analyse, à chaud, sa leçon en fonction de ce qu'il avait prévu dans sa préparation écrite et selon ses impressions. Dans un deuxième temps, à distance de sa prestation et seul, il visionne son document vidéo et complète le jugement qu'il porte sur sa prestation, à la lumière de la vidéo. Cette expérience est ensuite relatée dans un rapport écrit qui ne fait pas l'objet d'une évaluation chiffrée. À la lecture des rapports, il apparaît que cet exercice semble vraiment significatif pour la plupart des étudiants futurs enseignants, tant sur le plan professionnel que personnel. Une étude systématique des rapports (en cours) permettra certainement d'objectiver davantage ces impressions.

À la fin de chaque période de stage, le formateur fait le point avec chaque étudiant sur les compétences acquises ou non encore acquises. Ils fixent alors ensemble les objectifs à atteindre pour le stage suivant ou pour l'entrée dans la carrière.

---

<sup>7</sup> Si, à ce stade de la formation, la prestation en classe n'est pas évaluée; il en sera autrement lors des stages en situation de classe menant à la certification.

La posture réflexive visée par l'ensemble de ce dispositif<sup>8</sup> est réellement difficile à mettre en place chez les étudiants. En effet, ils ne sont pas habitués à livrer leurs points faibles, mais plutôt à vanter leurs mérites, acquis ou à acquérir... La formation professionnelle, ici à l'enseignement, amène très souvent à révéler des aspects de la personnalité, des comportements sociaux et de savoir-être jusque-là rarement pris en compte : certains, pourtant réputés excellents chercheurs, peuvent s'avérer être de piètres enseignants lors de premières prestations devant des élèves... Il s'avère donc très important de préciser au plus près les attentes des formateurs en termes de compétences, de savoirs, mais aussi d'attitudes et de savoir-être.

### *3.2.4. Evaluer leurs pratiques (lors des stages notamment) peut-il œuvrer à leur formation en matière d'évaluation et sous quelles conditions ?*

Certes, les évaluations de stages attirent l'attention des étudiants sur tout ce qu'il est possible d'évaluer au sein d'une classe, tant du côté de l'enseignant que des élèves. De même, il est tenté de pratiquer avec eux une évaluation dite authentique (au sens de Wiggins, repris par Mottier Lopez, 2015) qui articule apprentissage et évaluation en se basant sur toute une série de traces issues des situations d'enseignement et d'apprentissage (voir annexe 3). Par ailleurs, la conception d'évaluations certificatives critériées pour les élèves est abordée, notamment, en cours de didactique disciplinaire. Cependant, ce qui manque à leur formation est sans doute d'aborder davantage de facettes de l'évaluation des élèves, surtout *pour* l'apprentissage et *en tant* qu'apprentissage.

Si l'évaluation en stage devrait contribuer à former des professionnels armés pour entrer dans leur fonction d'enseignants et pour construire petit à petit leur identité professionnelle, on ne peut pas prétendre, en une si courte période de formation initiale, à les préparer finement à toutes les facettes du métier. Les prémisses d'une démarche réflexive, telle qu'elle peut être initiée en formation initiale, permettent déjà certaines prises de conscience et des formes d'ajustement visant à gérer les tensions entre représentations, valeurs, savoirs et pratiques effectives (Dubar, 1992, cité par Perez-Roux, 2010).

Cependant, faire appréhender aux étudiants une posture réflexive reste complexe, même si les modalités précédemment explicitées ont pour but de les amener à revenir sur l'expérience passée, de l'analyser au sein d'un collectif avec des outils particuliers, et de l'éclairer différemment. Par ailleurs, revenir sur le sens et la portée de son travail, accepter l'épreuve de la découverte de soi, travailler sur le réel du métier et pas sur sa forme idéalisée et entrer dans un processus de détachement, peuvent contribuer à favoriser l'émergence d'une identité professionnelle (Perez-Roux, 2010).

## **4. Apprendre à évaluer et (être) évalu(é)er pour apprendre**

Plusieurs questionnements ont été soulevés tout au long de ce texte : esquissés dans le cadrage proposé au point 2, ils trouvent de nombreux échos dans les exemples développés au point 3. Ces deux exemples, très différents tant par le contexte dans lequel ils s'exercent que par l'angle abordé, donnent néanmoins des éclairages complémentaires. Partant d'un objet précis (le portfolio), le premier exemple illustre la complémentarité entre les versants formatif et certificatif de l'évaluation (deux fonctions dont l'on s'accordera à reconnaître qu'il convient de ne plus les opposer). Relatant différents dispositifs mis en place pour former les futurs enseignants à l'évaluation, le deuxième exemple montre bien toute la complexité de cette entreprise. Former des enseignants nécessite de développer des compétences

---

<sup>8</sup> décrit dans Hindryckx & Poffé (2013).

spécifiques et de construire une identité professionnelle forte (Beckers, 2007). Il s'agit non seulement de former de bons praticiens, mais aussi et surtout de former des « praticiens réfléchis » voire « réflexifs », qui parviennent à établir une distance avec la situation vécue et à étendre le savoir à d'autres situations (Donnay & Charlier, 2006). Une telle posture ne va pas de soi et ne se met pas en place naturellement (Perrenoud, 2001) ; elle est pourtant essentielle car c'est en se donnant ces temps d'analyse « que se développe la compétence évaluative indispensable au développement du pouvoir d'agir ». Le portfolio développé dans le premier exemple et les pratiques d'autoscopies mentionnées dont le second sont deux outils de formation permettant le développement d'une telle posture.

#### 4.1 Apprendre à évaluer ? Evaluer quoi et pour quoi faire ?

Avant même de se poser la question de savoir comment former à l'évaluation, encore faut-il se montrer convaincu de la pertinence de cet objet d'apprentissage. L'évaluation fait-elle partie du rôle de l'enseignant ou ces derniers sont-ils là (en classe) pour enseigner et non pas pour juger ? Mais peut-on enseigner sans évaluer et peut-on même apprendre sans apprendre à évaluer ? S'il faut évaluer, il s'agira alors de se montrer « juste » en se refusant à « perpétuer l'impression d'arbitraire ressentie parfois pendant sa scolarité et son parcours universitaire », mais est-ce suffisant ? Est-ce d'ailleurs « juste » de vouloir tout évaluer ? Et au final, que signifier « évaluer » ?

Si l'ensemble du texte se positionne clairement face à ces interrogations, les étudiants qui entrent en formation ne sont, quant à eux, pas nécessairement convaincus d'emblée du rôle central que joue l'évaluation dans l'enseignement. Que cachent-ils derrière ce terme et sont-ils conscients qu'il recouvre une variété de pratiques ? Plusieurs études menées en France (Issaieva & Crahay, 2010), en Suisse romande (Boraita & Issaieva, 2013 ; Luisoni & Monnard, 2015) et en Belgique francophone (Fagnant & Goffin, à paraître) pointent toute l'importance de prendre en compte les conceptions et le vécu des étudiants en matière d'évaluation. Si on veut viser un changement de posture nécessaire à une évolution des pratiques, il convient non seulement de partir du « déjà-là », mais aussi de faire vivre aux étudiants des expériences menées en accord avec les pratiques qu'on souhaite qu'ils développent, ce que les deux exemples relatés nous ont joliment illustré.

Si les futurs enseignants s'accordent assez rapidement sur la nécessité d'évaluer les apprentissages des élèves (et ceci, même s'ils préféreraient être là pour enseigner et non pour juger), ce n'est pas pour autant qu'ils acceptent aussi facilement que l'on évalue leurs pratiques d'enseignement en stage, la réflexivité sur leurs pratiques ou encore les traces de leurs processus d'autorégulation. Pourtant, les deux exemples développés montrent bien en quoi ces évaluations sont nécessaires pour les aider à développer ces compétences complexes. Dans les deux cas, il s'agit tout d'abord de rendre l'évaluation visible, ce qui se traduit notamment par une explicitation des objectifs et des critères d'évaluation, ainsi que par les divers écrits sollicités. Toutefois, cette explicitation ne suffit pas pour que les étudiants les comprennent et, surtout, pour qu'ils y adhèrent et se les approprient. Dans le premier exemple, c'est au travers de la mise en œuvre d'un processus d'auto-évaluation en trois étapes que les futurs enseignants sont amenés à s'approprier les objectifs du cours et du portfolio ; dans le second exemple, c'est via l'analyse d'une séquence de cours filmée qu'ils sont amenés à intégrer la grille d'évaluation qui sera utilisée en stage. Dans les deux cas, ces expériences les conduisent à commencer à pratiquer ce que nous avons qualifié plus haut d'agir évaluatif ou, autrement dit, à pratiquer les prémices d'un « agir professionnel en situation évaluative » (Jorro & Mercier Brunel, 2011).

Les deux dispositifs décrits montrent tout aussi clairement que l'implication des étudiants dans le processus d'évaluation est primordiale pour atteindre les objectifs visés. Un changement de posture est nécessaire : il faut oser se mettre à nu pour pointer ses faiblesses alors que la société demande souvent de « se vendre sous leur meilleur jour » ; il faut oser faire état des stratégies d'autorégulation que l'on a développées, éventuellement hors du temps d'enseignement, alors que notre système scolaire valorise trop souvent la réussite sans effort comme étant un signe de compétence.

Enfin, même s'ils ne développent pas vraiment cette thématique, les deux exemples éclairent également toute l'importance de développer un « jugement professionnel » (Mottier Lopez, 2008) qui permet de donner plus de corps à cette notion de justice et peut aider les étudiants à dépasser la dichotomie qu'ils imaginent entre une évaluation (qu'ils voudraient objective) des apprentissages des élèves et l'évaluation (qu'ils imaginent subjective) de compétences complexes comme celles de leur propre pratique d'enseignement notamment.

#### **4.2. Prendre ses propres pratiques comme objet d'évaluation pour apprendre à enseigner**

À partir du moment où l'on dépasse la seule vision d'une évaluation de type notation-sanction et que l'on s'accorde à reconnaître la nécessaire complémentarité entre les fonctions formative et certificative de l'évaluation, on peut aussi admettre « qu'apprendre à enseigner », c'est aussi « apprendre à évaluer » puisqu'on ne peut pas apprendre sans apprendre à évaluer et puisque l'évaluation fait partie intégrante du processus d'enseignement apprentissage (*evaluation for learning*).

En formation d'enseignants, d'aucuns considèrent comme essentiel de faire vivre aux étudiants des situations d'évaluation qui s'accordent pleinement à ce que l'on veut leur faire apprendre et à ce que l'on souhaite qu'ils transfèrent dans leurs classes selon un principe d'homothétie, d'isomorphisme ou d'homologie (Campanale, 2007 ; Étienne & Fumat, 2014 ; Goffin & Fagnant, 2017). Les deux expériences relatées au point 3 s'accordent avec cette logique : dans le premier exemple, le dispositif de formation mis en place autour du portfolio est en cohérence avec les contenus enseignés ; dans le second exemple, des activités telles que les micro-enseignements permettent aux étudiants de vivre des activités proches de celles qu'ils rencontreraient lors de leurs stages d'enseignement.

Dans les deux exemples décrits dans ce texte, les futurs enseignants sont invités à analyser différentes situations d'évaluation dans lesquelles ils sont directement impliqués. Par exemple, lorsque les étudiants doivent rédiger un texte faisant preuve d'une mise en lien et d'une intégration des concepts du cours dans le cadre d'un portfolio, ils doivent aussi donner à voir les démarches d'autoévaluation et d'autorégulation qu'ils ont mises en œuvre pour progresser dans leurs apprentissages. Lorsqu'ils doivent analyser leurs pratiques de stage pour démontrer leur acquisition d'une posture réflexive, c'est là aussi une volonté de « transformer l'évaluation en moteur d'apprentissage » qui est au cœur du processus. Apprendre à évaluer et (être) évalu(é)er pour former semblent donc bien constituer les deux faces d'une même médaille.

#### **4.3 Pour terminer sans conclure...**

Même préparée en formation initiale, la pratique de l'évaluation provoque de l'émotion chez les enseignants, les élèves et les parents. La tentation de la fuite, du refuge dans un enseignement débarrassé de tout contrôle peut apparaître comme illusoire pour instaurer la paix scolaire et améliorer le climat dans la classe. Pourtant, il est une autre voie qui consiste

à se former à l'évaluation et à sa communication en pratiquant une co-évaluation formative et en signifiant aux élèves et aux personnes en formation que le but est de leur remettre, à terme, le pilotage de leurs apprentissages pour qu'ils puissent contrôler l'acquisition de leurs savoirs et le développement de leurs compétences pour agir dans des situations singulières. Ces démarches peuvent s'envisager dans un cadre individuel, mais elles peuvent aussi gagner en puissance en devenant l'objet de projets d'établissement (Étienne, Connac, Marquié-Dubié, 2017). Elles pourront alors être mises en œuvre, puis collectivement analysées dans le cadre de boucles évaluatives et correctrices. Une formation pédagogique et didactique à l'évaluation s'inspirant de ces principes peut être mise en place et accompagner la transition d'une évaluation-sanction unanimement repoussée vers une évaluation moteur et motivante.

Enfin, si les recherches menées sur la formation à l'évaluation en formation initiale doivent être développées, il convient aussi de poursuivre les études menées au moment de l'entrée dans le métier (De Stercke *et al.*, 2010 ; Luisoni & Monnard, 2015) et tout au long du parcours professionnel, au travers de recherches collaboratives notamment. A cet égard, nous renvoyons le lecteur à un numéro thématique publié dans cette même revue (voir Morissette & Tessaro, 2016 pour le texte introductif) et coordonné par le réseau thématique RCPE de l'ADMEE-Europe (Recherches collaboratives sur les pratiques évaluatives : <http://admee.ulg.ac.be/activites/reseaux-thematiques/rcpe/>).

## 5. Références

- Allal, L. (1999). Impliquer l'apprenant dans le processus d'évaluation : promesses et pièges de l'autoévaluation. In C. Depover & B. Noël (Ed.), *L'évaluation des processus cognitifs dans l'apprentissage* (pp. 33-52). Bruxelles : De Boeck.
- Allal, L. & Mottier Lopez, L. (2005). Formative Assessment of Learning : A Review of Publications in French. In *Formative Assessment - Improving Learning in Secondary Classrooms* (pp. 241-264). Paris : OECD Publication.
- Antibi, A. (2003). *La constante macabre*. Toulouse : édition Math'Adore.
- Assessment Reform Group 2002, *Assessment for Learning: 10 principles research-based principles to guide classroom practice*. London, UK : Assessment Reform Group.
- Barbier, J.-M. (1984). *L'évaluation en formation*. Paris : PUF.
- Beckers, J. (2007). *Compétences et identité professionnelles. L'enseignement et autres métiers de l'interaction humaine*. Bruxelles : De Boeck.
- Berger, G. (1977). Mais qu'est-ce qui nous prend à évaluer ? *Pour*, 55.
- Bernié, J.-P. (2002). L'approche des pratiques langagières scolaires à travers la notion de « communauté discursive » : un apport à la didactique comparée ? *Revue Française de pédagogie*, 141, 77-88.
- Boimare, S. (1999). *L'enfant et la peur d'apprendre*. Paris : Dunod.
- Bonniol, J.-J. & Vial, M. (1997). *Les modèles de l'évaluation*. Bruxelles : de Boeck .
- Boraita, F. & Issaieva, E. (2013). *Les conceptions de l'évaluation chez les futurs enseignants et enseignants en exercice : une étude dans le canton de Genève*. Actes du 25e colloque de l'ADMEE-Europe: Evaluation et autoévaluation, quels espaces de formation. Du 9 au 11 janvier 2013, Fribourg (Suisse).
- Boutinet, J.-P. (1998). *L'immaturité de la vie adulte*. Paris : PUF.
- Campanale, F. (2007). Evaluation réflexive en formation professionnelle et évaluation interactive dans les classes. In L. Allal & L. Mottier Lopez. (Eds.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 191-206). Bruxelles : De Boeck.
- Cartier, S. C. (2007). *Apprendre en lisant au primaire et au secondaire : Mieux comprendre et mieux intervenir*. Anjou : Editions CEC Enseignants Collégial et universitaire.
- Clot, Y. (1999). *La fonction psychologique du travail*. Paris : PUF.

- De Stercke, J., Renson, J.-M., Cambier, J.-B., Leemans, M., Maréchal, C., Radermaecker, G., Temperman, G., Beckers, J. & De Lièvre, B. (2010). *Dynamiser l'insertion professionnelle des enseignants débutants du secondaire*. Récupéré [en mars 2017] du site enseignement.be : [http://www.enseignement.be/index.php?page=23827etdo\\_id=8042etdo\\_check=](http://www.enseignement.be/index.php?page=23827etdo_id=8042etdo_check=)
- Doly, A.-M. (2006). La métacognition : de sa définition par la psychologie à sa mise en œuvre à l'école. In G. Toupiol. *Apprendre et Comprendre. Place et rôle de la métacognition dans l'aide spécialisée* (pp. 84-124). Paris : Retz.
- Donnay, J. & Charlier, E. (2006). *Apprendre par l'analyse des pratiques. Initiation au compagnonnage réflexif*. Namur : Presses universitaires de Namur.
- Étienne, R. (1999). L'établissement formateur. Quelle contribution du "terrain" à la formation générale des professeurs de l'enseignement secondaire ? *Recherche et formation, INRP, 31*, 137-151.
- Étienne, R. (2000). *Les réseaux d'établissements, enjeux à venir*. Paris : ESF.
- Étienne, R. (2016). Former à évaluer : enjeux, tensions, solutions. *Éducation permanente, 208*, 101-113.
- Étienne, R., Connac, S., Marquié-Dubié, H. (2017). *Effets de l'accompagnement dans l'action et la recherche au sein d'un collège avec des classes coopératives*. Communication à paraître dans les Actes du colloque international « Invention d'espaces de travail. Entre chemins individuels et pistes collectives » les 17, 18 et 19 mai 2017, Université de Lille 3.
- Étienne, R. & Fumat, Y. (2014). *Comment analyser les situations éducatives pour se former et agir*. Bruxelles : De Boeck, collection Guides pratiques, Former & Se former.
- Étienne, R. & Zamaron, A. (2005). Enseigner, un métier qui s'apprend. *Les Cahiers Pédagogiques, 435*, 7-54.
- Fagnant, A. & Goffin, C. (2017). Les conceptions des futurs enseignants du secondaire en matière d'évaluation : entre un accord de principe et une vision limitée de l'évaluation formative. *Mesure et Evaluation en Education, 40(1)*, 1-32.
- Goffin, C. & Fagnant, A. (2017). Faire vivre une expérience d'évaluation mutuelle en formation initiale : quelle transposition vers les pratiques d'enseignement projetées ? *Contexte et Didactique, 9*, 46-58.
- Hameline, D. (1979). *Les objectifs pédagogiques en formation initiale et en formation continue*. Paris : ESF.
- Hindryckx, M.-N. & Poffé, C. (2013). Pratiques réflexives en formation initiale des enseignants. *Education et Formation, 298(2)*, 87-104.
- Issaieva, E. & Crahay, M. (2010). Conception de l'évaluation scolaire des élèves et des enseignants : validation d'échelles et études de leurs relations. *Mesure et Evaluation en Education, 33(1)*, 31-62.
- Jankélévitch, V. (1960). Hommage à Bergson. Extrait du *Bulletin de la Société Française de Philosophie, 54(1)*, 55-62.
- Jean, A. (2009). *Le traitement des imprévus par les professeurs stagiaires de technologie en formation initiale à l'UFRM. Quels gestes professionnels d'ajustement ? Quelle utilisation pour leur professionnalité ?* Thèse de doctorat (sous la direction de R. Étienne, soutenue à Montpellier le 18 juin 2008). Lille : Atelier National de Reproduction des thèses.
- Jorro, A. & Mercier-Brunel, Y. (2011). Les gestes évaluatifs de l'enseignant dans une tâche de correction collective. *Mesure et Evaluation en Education, 34(3)*, 27-50.
- Lave, J. & Wenger, E. (1991). *Situated learning. Legitimate peripheral participation*. Cambridge : Cambridge University Press.
- Le Boterf, G. (1994). *De la compétence : essai sur un attracteur étrange*. Paris : Les éditions d'organisation.
- Luisoni, M. & Monnard, I. (2015). L'évolution des conceptions de l'évaluation scolaire chez les futurs enseignants en formation initiale. In P.-F. Coen & L. M. Bélaïr (Eds.), *Évaluation et autoévaluation. Quels espaces de formation ?* (pp. 175-196). Louvain-la-Neuve : De Boeck Supérieur.
- Merle, P. (2015). L'école française et l'invention de la note. Un éclairage historique sur les polémiques contemporaines. *Revue française de pédagogie, 193*, 77-88.
- Mottier Lopez, L. (2008). Discussion notionnelle : les relations entre jugements et évaluation. *Revue Suisse des Sciences de l'Education, 30(3)*, 455-463.
- Mottier Lopez, L. (2015). *Evaluations formative et certificative des apprentissages. Enjeux pour l'enseignement*. Bruxelles : De Boeck.

- Mottier Lopez, L. (2017). L'étude d'un dispositif d'évaluation formative et certificative visant à soutenir l'autorégulation des apprentissages des étudiants en contexte universitaire. In S. C. Cartier & L. Mottier Lopez (Ed.), *Soutien à l'apprentissage autorégulé en contexte scolaire* (pp. 55-83). Québec : Presses Universitaires du Québec.
- Nunziati, G. (1989). Pour construire un dispositif d'évaluation formatrice. *Cahiers pédagogiques*, 280, 47-64.
- Paquay, L., Perrenoud, Ph., Ajtet, M., Étienne, R. & Desjardins, J. (2014). *Travail réel des enseignants et formation : quelle référence au travail des enseignants dans les objectifs, dispositifs et pratiques ?* Bruxelles : De Boeck Université.
- Perrenoud, P. (2001). Mettre la pratique réflexive au centre du projet de formation. *Cahiers pédagogiques*, 390, 42-45.
- Perez-Roux, T. (2010). Quelle prise en compte du Sujet dans la formation des enseignants ? Enjeux et limites d'une approche clinique. In G. Baillat, D. Niclot & D. Ulma (Eds.). *La formation des enseignants en Europe* (pp. 128-138). Bruxelles : De Boeck.
- Ria, L. (2015, dir.). *Former les enseignants au XXI<sup>e</sup> siècle. 1. Établissement formateur et vidéo-formation*. Bruxelles : De Boeck.
- Saint-Arnaud, Y. (2001). La réflexion dans l'action : un changement de paradigme. *Recherche et formation*, 36, 17-27.
- Schön, D.A. (1988). Coaching Reflective Teaching. In P.P. Grimmet & G.L. Erickson (Eds.), *Reflection in teacher education* (pp. 19-29). Vancouver : Pacific educational press.
- Tourmen, Cl., (2016). Apprendre à évaluer. Dossier d'*Éducation permanente*, 208, 5-136.
- Wiliam, D. (2011). *Embedded Formative Assessment*. Bloomington IN, USA : Solution Tree Press.

## Annexe 1. Canevas pour les traces 1 du portfolio

**COURS : EVALUATION ET REGULATION DES APPRENTISSAGES – Lucie Mottier Lopez**  
**Semestre d'automne 2016**

### **Fiche d'autoévaluation et d'autorégulation**

*À insérer dans votre portfolio, agrafée à l'activité interactive choisie*

Nom de l'activité interactive faite pendant le cours : *(rubrique ouverte à renseigner par l'étudiant)*

Objectif(s) d'apprentissage visé(s) annoncé(s) par l'enseignant(e) : *(rubrique ouverte à renseigner par l'étudiant)*

### **Rubriques à renseigner après le cours**

#### **Modélisation de mon apprentissage autorégulé :**

- 1) Concrètement, ce que j'ai appris par cette activité par rapport à l'objectif d'apprentissage / aux objectifs associés : *(rubrique ouverte à renseigner par l'étudiant)*
- 2) Ce que je n'ai pas (bien) compris, les interrogations concrètes qui sont demeurées quand j'ai fait l'activité : *(rubrique ouverte à renseigner par l'étudiant)*

→ **Formulation de « mon objectif personnel d'apprentissage » pour progresser** : *(rubrique ouverte à renseigner par l'étudiant)*

- 3) Ce que je mets/ j'ai mis en œuvre pour mieux comprendre et progresser par rapport à mon objectif personnel d'apprentissage : *(rubrique ouverte à renseigner par l'étudiant)*

→ **Insérer dans le portfolio des traces concrètes qui montrent ce que j'ai mis en œuvre.**

- 4) Là où j'en suis actuellement dans ma compréhension par rapport à mon objectif personnel d'apprentissage (bilan en fin de semestre) : *(rubrique ouverte à renseigner par l'étudiant)*

**Annexe 2. Tableau qui reprend les caractéristiques des pièces à insérer dans le portfolio (adapté de Mottier Lopez, soumis)**

	<b>Traces 1 : AI</b>	<b>Traces 2 : TP</b>	<b>Trace 3 : TIR</b>
<b>Description</b>	Activités interactives (AI)	Travaux pratiques (TP)	Texte d'intégration et de réflexion critique (TIR)
<b>Thèmes</b> (Portfolio, automne 2016)	AI 1 : <i>Fonctions de l'évaluation</i> AI 2 : <i>Quelles informations solliciter auprès des élèves dans des démarches d'autoévaluation ?</i> AI 3 : <i>Evaluation en collaboration</i> AI 4 : <i>Régulation de l'enseignement et de l'apprentissage</i> AI 5 : <i>Rédigez vos critères d'évaluation...</i> AI 6 : <i>Identification de biais en cours de correction</i>	TP1 : <i>Fonctions de l'évaluation et implication de l'élève / des élèves</i>  TP2 : <i>Evaluation formative et interventions à des fins de régulation de l'apprentissage</i>	En 2 parties : 1. Réflexion critique sur les stratégies d'autorégulation mises en œuvre en étayant avec les documents insérés dans le portfolio  2. Réflexion conceptuelle critique sur les liens entre enseignement, apprentissage, et évaluation
<b>Types de tâches et exigences</b>	Tâches spécifiques - Appropriation de concepts Tâches complexes - Autoévaluation - Mise en œuvre de stratégies d'autorégulation - Ecriture réflexive	Tâches complexes - Utilisation de concepts comme cadre d'analyse de données et de conception d'interventions formatives	Tâche complexe - Mobilisation des concepts pour produire des mises en relation avec des données empiriques - Production d'une argumentation conceptuelle
<b>Choix &amp; obligations</b>	Choix par l'étudiant de 3 AI sur 6 proposées en cours de semestre	2 TP obligatoires	1 texte obligatoire
<b>Articulation avec l'enseignement</b>	- Insérées dans l'enseignement - Puis décrochées : mises en œuvre par l'étudiant de stratégies d'autorégulation	- Ponctuels : un cours complet consacré à chaque TP	- Décroché : à réaliser hors du cours

<p><b>Conception par le formateur des modalités de feedback et de régulation</b></p>	<p>Diverses modalités pendant le cours :</p> <ul style="list-style-type: none"> <li>- Interaction entre pairs</li> <li>- Mise en commun</li> <li>- Feedback du formateur (à la demande de l'étudiant)</li> </ul> <p>Hors du cours :</p> <ul style="list-style-type: none"> <li>- Ressources sur la plateforme d'enseignement</li> <li>- Permanence (à la demande de l'étudiant)</li> </ul>	<p>Suite à la correction par l'enseignante/assistant du cours :</p> <ul style="list-style-type: none"> <li>- Feedback écrit personnalisé</li> <li>- Retour collectif</li> </ul> <p>Hors du cours :</p> <ul style="list-style-type: none"> <li>- Permanence (à la demande de l'étudiant)</li> </ul>	<p>Pendant le cours :</p> <ul style="list-style-type: none"> <li>- Un temps collectif pour répondre aux questions</li> </ul> <p>Hors du cours :</p> <ul style="list-style-type: none"> <li>- Permanence (à la demande de l'étudiant)</li> </ul>
--	--	--	---

Annexe 3. Tableau qui reprend les moments d'évaluation proposés aux étudiants futurs enseignants lors de la formation ainsi que leurs visées pour le métier (Hindryckx, document non publié)

Visées de la formation	Évaluation des élèves	Évaluation des étudiants futurs enseignants Évaluation <i>de</i> l'apprentissage		Évaluation des pratiques de classe, posture réflexive Évaluation <i>pour</i> l'apprentissage		
Cours théoriques	Évaluation par compétences Critères et indicateurs Corrections croisées	Exercices de microenseignements : construction avec l'aide des formateurs d'une micro leçon de sciences à donner devant des pairs (thème et méthodologie fixés)		Exercice d'évaluation d'un stagiaire sur prestation filmée (re)Construction d'une grille		
Stages pratiques (40 périodes)			Stages (40 leçons) en trois périodes au minimum	Utilisation de la grille par les formateurs lors des visites de stage	Film d'une leçon de stage au choix de l'étudiant	Entrevue de chacun des futurs enseignants avec les formateurs
<i>Evaluation formative</i>	Construction de grilles critériées avec indicateurs	Prestation évaluée par l'ensemble des participants (auto-hétéroscopie)	Rapports de stage (préparations et reculs réflexifs) Évaluations du/des Maître(s) de stage		Rapport d'autoscopie Ecart entre préparation et prestation, recul avant et après visionnage de la prestation	Faire le point sur les compétences en stage et fixer ensemble des objectifs à travailler pour les périodes de stage suivantes ou pour l'entrée dans le métier

<p><i>Evaluation certificative</i></p>	<p><b>Examen</b> : Leçon à produire avec une évaluation certificative par compétences adaptée</p>	<p><b>Rapport de microenseignement</b> reprenant les principales remarques faites ainsi qu'une discussion à propos de celles-ci</p>	<p><b>Evaluations des stages</b> : quatre évaluations par les formateurs sur sites d'enseignement Grilles complétées et résumés fournis aux étudiants futurs enseignants</p>			
--	---	---	--	--	--	--

# Evaluer : pour contrôler ou pour apprendre ?

**Christophe Dierendonck**

Université de Luxembourg  
[christophe.dierendonck@uni.lu](mailto:christophe.dierendonck@uni.lu)

**Carmen Cavaco**

Universidade de Lisboa  
[carmen@ie.ulisboa.pt](mailto:carmen@ie.ulisboa.pt)

**Pour citer cet article :** Dierendonck, Ch. & Cavaco, C. (2017). Evaluer : pour contrôler ou pour apprendre ? *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 99-100.

Les textes rassemblés dans ce numéro spécial offrent à la fois une perspective internationale et une couverture de plusieurs niveaux du système d'éducation et de formation. Quel que soit le contexte envisagé, ils mettent en lumière - si besoin en était encore - la prédominance des pratiques d'évaluation visant le contrôle et la régulation des apprentissages, des individus et des institutions. Mais ils permettent également de mieux comprendre pourquoi certaines pratiques d'évaluation éprouvent tant de difficultés à se développer. Chacune à leur manière, les six contributions montrent l'intérêt de dépasser les logiques de contrôle des acquis scolaires et de régulation de l'hétérogénéité traditionnellement associée à l'acte évaluatif posé en classe par l'enseignant ou par une évaluation externe (évaluation *de* l'apprentissage) pour entrevoir les perspectives offertes par les versants formatif (évaluation *au service de* l'apprentissage) et formateur (évaluation *en tant qu'*apprentissage) de l'évaluation.

## 1. Evaluer pour contrôler

L'évaluation *de* l'apprentissage renvoie à une pratique évaluative classique des enseignants dont la fonction est de certifier si, au terme d'une séquence d'enseignement, les élèves ou les étudiants maîtrisent ou non les apprentissages visés. La démarche adoptée dans ce genre d'évaluation est généralement de nature sommative ; une note globale est attribuée en faisant la somme des réussites observées à une série de questions ou d'items inclus dans un test, un contrôle ou un examen. Les résultats obtenus sont consignés dans un document officiel (comme le bulletin) pour décider *in fine* de la promotion des élèves et des étudiants vers l'année d'étude suivante ou de l'obtention d'un certificat ou diplôme. L'évaluation *de* l'apprentissage n'est cependant plus de la seule responsabilité des enseignants. Depuis plusieurs décennies, l'évaluation externe s'impose avec force dans tous les milieux. Les résultats des élèves et des étudiants ne concernent plus uniquement ces derniers. Une fois agrégés à un niveau spécifique, ils servent à évaluer les systèmes éducatifs, les établissements de formation, voire même les enseignants.

## 2. Evaluer pour apprendre

L'évaluation *au service de* l'apprentissage et l'évaluation *en tant qu'*apprentissage sont considérées comme parties intégrantes du processus d'enseignement/apprentissage. Cette forme d'évaluation est selon Black et William (1998) au cœur de l'enseignement efficace. L'évaluation n'est plus vue comme une évaluation-contrôle ou évaluation-sanction, elle cherche à favoriser la motivation et l'implication des apprenants en les informant de manière continue ou en les amenant à se rendre compte par eux-mêmes des apprentissages réalisés ou en voie de l'être. L'évaluation *au service de* l'apprentissage renvoie à la fonction formative classique de l'évaluation, à savoir une évaluation qui informe l'enseignant et l'élève (diagnostique quant aux apprentissages réalisés mais aussi quant à l'efficacité des pratiques d'enseignement) et qui permet d'améliorer les apprentissages (régulation, en envisageant ensemble les stratégies à adopter). L'évaluation *en tant qu'*apprentissage renvoie aussi à la fonction formative, mais cette fois avec une perspective élargie (Allal et Mottier Lopez, 2005 ; Fagnant et Goffin, 2017) qui intègre l'évaluation au sein du processus d'enseignement/apprentissage et qui peut aussi faire des élèves et des étudiants des acteurs de leur propre évaluation (co-évaluation, auto-évaluation, ...).

## 3. Conclusion

Force est de reconnaître que l'évaluation pour contrôler demeure la pratique prédominante pour des raisons historiques, politiques et économiques. Selon Rey et Feyfant (2014), évaluer pour apprendre, c'est d'ailleurs « *se situer à rebours d'une tradition où l'évaluation est conçue dans un modèle cybernétique comme un contrôle a posteriori, quasiment technique, d'un apprentissage qui serait indépendant des modalités d'évaluation* ». Cela demande une professionnalisation accrue, des outils d'évaluation adaptés et un investissement quotidien majeur pour piloter les apprentissages, au plus près des particularités de chaque élève ou étudiant. A lire les contributions de ce numéro thématique, il semble évident que l'évaluation ne doit pas se limiter à un questionnement sur la conformité ou la cohérence, mais qu'elle doit surtout questionner le sens et les significations des actes posés (Ardoino et Berger, 1989). L'évaluation en tant que problématisation du sens met en évidence l'importance de la réflexion *dans et sur l'action* et ses liens avec l'apprentissage.

Au-delà de l'évaluation *instituée*, davantage étudiée, reconnue et qu'il faudrait sans doute remettre en question, il semble fécond, si pas nécessaire aujourd'hui, de s'intéresser aux dimensions de l'évaluation *instituant* qui valorise la richesse, la singularité et le potentiel d'apprentissage des processus d'évaluation vécus par les élèves, les étudiants et les enseignants.

## 4. Références bibliographiques

- Allal, L. & Mottier Lopez, L. (2005). Formative Assessment of Learning: A Review of Publications in French. In *Formative Assessment - Improving Learning in Secondary Classrooms* (pp. 241-264). Paris: OECD Publication.
- Ardoino, J. & Berger, G. (1989). Fondements de l'évaluation et démarche critique, *AECSE*, 6, 3-11.
- Fagnant, A. & Goffin, C. (2017). Les conceptions des futurs enseignants du secondaire en matière d'évaluation : entre un accord de principe et une vision limitée de l'évaluation formative. *Mesure et évaluation en éducation*, 40 (1), 1-32.
- Rey, O. & Feyfant, A. (2014). Evaluer pour (mieux) faire apprendre. *Dossier de veille de l'IFÉ*, 94, 19-27. Lyon : ENS de Lyon.