

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

Domingos Fernandes

Universidade de Lisboa
dfernandes@ie.ulisboa.pt

Isabel Fialho

Universidade de Évora
ifialho@uevora.pt

Carlos Barreira

Universidade de Coimbra
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 (Giglione & 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 (n=990)	Enseignants (n=4568)	<i>p value</i>	<i>d-Cohen</i>
	\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-Rebelo, 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-Rebelo, 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.