

Éditorial: La formation du jugement en évaluation

Lucie Mottier Lopez et Franz Baeriswyl

La genèse de la formation du jugement en évaluation dans l'enseignement est analysée dans les contributions thématiques de ce numéro. D'une façon générale, les objets et les fonctions de l'évaluation dans l'enseignement sont multiples. Ainsi, les performances disciplinaires spécifiques, les savoirs et savoir-faire, mais aussi les compétences transversales telles que l'attitude au travail, l'engagement ou la compétence communicative sont soumises à l'évaluation des enseignants. Lorsque l'évaluation sert à la régulation des processus d'apprentissage (évaluation formative), elle informe l'apprenant de ses progrès, de ses acquis et de ce qu'il doit encore apprendre ou améliorer. Elle fournit aussi des informations à l'enseignant pour qu'il adapte et différencie ses dispositifs de formation par rapport aux besoins des apprenants. L'évaluation formative n'est donc pas une évaluation finale comme c'est le cas de l'évaluation sommative ou certificative qui, quant à elle, atteste de l'aboutissement d'un parcours de formation ou d'une entité temporellement structurée, telle que par exemple une année scolaire. Pour poursuivre la scolarité, le passage d'un degré à l'autre nécessite une recommandation, souvent holistique et de nature pronostique.

Avec le transfert des exigences psychométriques de qualité à l'évaluation en contexte scolaire, une masse de recherches empiriques s'est développée dans les années 1970 et 1980. Elles ont mis en évidence les insuffisances de l'évaluation scolaire. On a, par contre, rarement entendu les voix de ceux qui mettaient l'accent sur l'erreur paradigmatique à la base de ces études, à savoir que les critères de qualité de la psychométrie ne pouvaient être pris en compte pour l'évaluation en classe qui est liée à un contrat pédagogique et didactique. Paradoxalement, les résultats des recherches pointant les «coupables» dans le paysage éducatif ont eu un effet considérable sur le développement d'une recherche propre, pourvue d'une méthodologie adaptée au contexte scolaire. Des aspects de cette recherche propre aux sciences de l'éducation apparaissent dans les contributions thématiques publiées ici. Une nouvelle focalisation s'y observe: la genèse de la formation du jugement en évaluation de l'enseignant à propos d'objets d'évaluation différents.

Pour clarifier ce point de vue original, une discussion notionnelle est publiée ci-après à la suite de l'éditorial.

Les contributions de ce numéro portent toutes sur la formation du jugement de l'enseignant dans des situations d'évaluation certificative ou pronostique, autrement dit dans des situations comportant des enjeux importants pour les parcours de formation des élèves et de leur orientation. Non seulement ces situations de transition sont particulièrement sensibles notamment quand les systèmes scolaires ne permettent que difficilement des réorientations de parcours (Baeriswyl, Wandeler & Christ; Maaz, Neumann, Trautwein, Wendt, Lehmann & Baumert, dans ce numéro), mais de plus elles convoquent des jugements pronostiques qui comportent une part d'incertitude que l'on ne peut jamais totalement éliminer. Ce sont des formes de «paris sur l'avenir» dont les enseignants ont conscience notent Marcoux et Crahay dans ce numéro de la Revue.

La recherche présentée par *Lucie Mottier Lopez* et *Linda Allal*, dans le premier article, examine les caractéristiques du jugement professionnel en évaluation d'enseignants genevois à la fin de l'école primaire. Les résultats de la recherche décrivent leurs pratiques de notation et la nature des ajustements qu'ils effectuent dans le cas de situations d'incertitude pour certains élèves en particulier. Le jugement professionnel, dans cet article, est appréhendé comme un processus co-gnitif et une pratique sociale liée à différents contextes, amenant les auteures à décrire les dimensions individuelles et sociales du jugement en évaluation.

Le deuxième article, de *Dany Laveault*, propose une revue de littérature sur la notion de «jugement professionnel d'évaluation», en puisant dans des apports nord-américains, (francophones et anglophones) ainsi qu'euro-péens. L'auteur met en lumière différentes catégories et indicateurs de jugement et les discute au regard des gestes professionnels et des exigences des réformes en éducation. La complexité de l'évaluation certificative et pronostique est soulignée (e.g. forts enjeux, buts multiples, conflits possibles de loyauté), ainsi que les «synergies» qu'il s'agirait de développer pour améliorer la qualité des jugements professionnels des enseignants.

La contribution de *Géry Marcoux* et *Marcel Crahay* a pour but d'analyser, dans une approche psychologique, la complexité de l'action qui conduit l'enseignant à des décisions de redoublement malgré les nombreux résultats de recherche qui montrent l'inefficacité de cette mesure. Par le moyen d'entretiens, les chercheurs tentent de dégager la trame décisionnelle du jugement de l'enseignant lorsque celui-ci décide pour un élève une promotion ou un redoublement pour un autre. Sur la base de leurs constats de recherche, les auteurs débattent notamment de la professionnalité ou non des jugements des enseignants lorsque ceux-ci continuent, malgré tout, à faire redoubler leurs élèves.

Kai Maaz, Marko Neumann, Ulrich Trautwein, Wolfgang Wendt, Rainer Lehmann et Juergen Baumert étudient les appréciations des enseignants à propos des compétences d'apprentissage lors du passage en fin d'école primaire vers les filières du secondaire I à Berlin. Les auteurs discutent les différentes «sources» ou éléments qui doivent être pris en compte pour former une appréciation des compétences d'apprentissage. En l'occurrence, ils ont observé un effet systématique à propos du genre (fille-garçon) et un effet lié au statut socioéconomique. De plus, le niveau de la classe influence le jugement évaluatif de l'enseignant, ce qui peut être caractérisé comme un effet du groupe de référence.

Un même questionnement est au cœur de la contribution de *Franz Baeriswyl, Christian Wandeler et Oliver Christ*: ils examinent la stabilité de la composition des recommandations de l'enseignant au moment du passage au secondaire I. Bien qu'une grande partie des enseignants changent d'une année à l'autre, sur les sept ans étudiés, les auteurs ne constatent aucun changement de recommandations statistiquement interprétables. La comparaison de ces recommandations est d'un grand intérêt: les évaluations des compétences d'apprentissage de Maaz et al. mettent en évidence un effet significatif lié au genre qui n'est pas systématiquement repérable chez Baeriswyl et al. Par ailleurs, la stabilité temporelle des recommandations des enseignants est interprétée comme une caractéristique positive pour l'équité du jugement.

La troisième contribution en allemand de *Anke Treutlein, Jeanette Roos et Hermann Schöler* étudie l'effet du groupe de référence d'une classe à propos de la notation individuelle en orthographe, en lecture et, prise globalement, en allemand. Contrairement à l'hypothèse et aux résultats de Maaz et al., l'effet du groupe de référence est moindre, voire absent. Le résultat contraire aux attentes initiales est discuté de façon détaillée. Les auteurs mettent en évidence plusieurs explications possibles. Nous y renvoyons le lecteur.

Les six contributions thématiques sur la formation du jugement en évaluation focalisent différents modes et formes d'évaluation. Un puzzle de la genèse de la formation du jugement en évaluation s'amorce ainsi, qui, sur la base de ces seules contributions, n'esquisse toutefois pas encore une image bien différenciée. Les contributions de ce numéro permettent la poursuite de la conceptualisation et de la compréhension des processus en jeu dans l'évaluation pédagogique, en posant notamment la question difficile du rapport entre jugement et évaluation. Un des intérêts est d'étudier les processus de jugement dans des *pratiques ordinaires* d'évaluation, plus particulièrement dans des situations qui, comme on l'a dit, se caractérisent par de forts enjeux au niveau individuel et collectif. La problématique de la formation du jugement en évaluation apparaît importante compte tenu aussi de certaines politiques éducatives actuelles qui tendent à réduire les marges de souplesse et de professionnalité des enseignants en instituant des procédures

mécaniques de prise de décision. Nous pensons notamment aux procédures qui se fondent essentiellement sur des calculs de moyennes de notes, soulevant le problème de la prise en compte des informations autres que les résultats chiffrés et dont les recherches dans ce numéro montrent qu'elles sont capitales pour des décisions d'évaluation certificative et pronostique bien fondées.