

Pour citer cet article : Donzé, T., & Pasche Gossin, F. (2023). Tensions et ambivalences chez les enseignant·es face à l'évaluation externe à visée d'orientation. *La Revue LEeE, varia*. <https://doi.org/10.48325/rleee.varia.001.01>

TENSIONS ET AMBIVALENCES CHEZ LES ENSEIGNANT·ES FACE À L'ÉVALUATION EXTERNE À VISÉE D'ORIENTATION

Tristan DONZÉ, Françoise PASCHE GOSSIN

Version de la publication : juin 2023

Évaluation ouverte et collaborative

Rétroacteurs : Dominique Broussal, Philippe Haerberli

Résumé

Cette recherche porte sur les « épreuves communes », tests pronostiques qui visent l'orientation dans le cycle 3 dans le canton du Jura en Suisse. La recherche a pour but de révéler les contextes et enjeux culturels de l'évaluation externe à visée d'orientation au travers de diverses dimensions de tensions et ambivalences chez les enseignant·es. L'article se situe dans un cadre théorique inspiré de la sociologie critique et compréhensive et présente une perspective critique des pratiques évaluatives externes et standardisées. Un des résultats majeurs pointe les diverses axiologies de faits et de valeurs qui transparaissent du discours des enseignant·es au sujet de l'évaluation des élèves qui mettent tension évaluation externe et évaluation interne au sujet notamment de principes de justices.

Mots-clés : évaluation externe, évaluation à visée d'orientation, croyances des enseignant·es

Abstract

This research focuses on external tests, prognostic tests aimed at curriculum counselling in cycle 3 in the canton of Jura in Switzerland. The research aims to reveal the contexts and issues at stake in external evaluation for guidance through various dimensions of tension and ambivalence among teachers. This article is located in a theoretical body inspired by critical and comprehensive sociology and presents a critical perspective of internal and standardized evaluative practices. One of the major results points to various axiologies of facts and values that emerge from the discourse of teachers about student assessment. This could explain the tension between external evaluation and internal evaluation concerning principles of justice.

Keywords: external tests, curriculum counselling, justice in assessment, teacher's beliefs

« Pour filer une métaphore médicale, notre embrouillement sur l'utilisation des tests standardisés revient à confondre le pouls avec l'effet total d'un régime sain. »
(Wiggins, 1989)

Eu égard à la gouvernance et la gestion du système scolaire par les chiffres (Felouzis & Hanhart, 2011), certaines études mettent au jour les paradoxes entre un désir de « bien faire » étatique et l'appréhension sur le terrain face à des outils de mesure standardisés (e.g., Mons, 2009 ; Yerly, 2022). D'autres posent un regard critique sur les formes plus ou moins réussies d'accointances entre mesures statistiques et éducation, stipulant, notamment, que les concepts et catégories élaborés dans un cadre scientifique de la mesure ont une incidence indirecte sur la gouvernance et la conception générale de l'éducation (Kieffer & Tréhin-Lalanne, 2011). En trame de fonds émerge la question fondamentale du rapport entre l'évaluation standardisée et les « bénéfices potentiels du testing » (Mons, 2009, p. 131).

Fortes sont les opinions concernant les épreuves standardisées et leur garantie d'objectivité dans l'évaluation scolaire. Un argument récurrent de la *doxa* en faveur de ce type de tests pointe l'importance de repères extérieurs à la classe pour guider le personnel enseignant. Ces indicateurs permettent-ils « d'objectiver et d'encadrer la subjectivité en évaluation », ou révèlent-ils de « la chimère » de l'objectivité en évaluation (Romainville, 2011, p. 1) ? Car la confiance en la solidité d'une balise externe repose aussi bien sur la fiabilité de la mesure des faits que sur les valeurs qui donnent appui à cette dernière.

C'est cette large étendue qu'interroge cette recherche au sujet d'une petite pièce d'eau : la procédure d'orientation par « épreuves communes » mise en œuvre en Suisse romande, plus spécifiquement dans le canton du Jura¹. Prescrites par la loi cantonale scolaire, ces épreuves communes s'appliquent à tous les élèves en fin de 2^e cycle de la scolarité obligatoire (élèves de 10-12 ans). Ces tests élaborés dans un cadre externe à la classe, organisés par l'État, ont comme visée l'orientation des élèves par niveaux au cycle 3 (élèves de 12-15 ans). Par quoi les épreuves marquent l'entrée dans l'enseignement segmenté par discipline et par groupes d'enseignement distincts répartis sur trois niveaux d'enseignement distincts².

Cette procédure a plus de trente ans d'existence (entrée en vigueur le 20 décembre 1990). La recherche commanditée³, menée en vue d'exposer les tenants et aboutissants du système d'orientation actuel, a conduit à interroger les principaux partenaires et actants du système en place (corps enseignant, groupe de rédaction, parents, membres du syndicat). Nous présentons ici un focus sur quelques résultats issus de la recherche susmentionnée selon une approche théorique inspirée de la sociologie critique et compréhensive, dans

¹ Cette évaluation est considérée comme une *évaluation étatique externe à fonction de sélection et d'orientation* en raison du statut de l'organisation qui la mandate, un canton suisse. Dans un cadre fédéraliste, la Suisse compte 26 « États enseignants », dotés de larges libertés juridiques et organisationnelles (Hofstetter, 2012), la législation scolaire étant du ressort des cantons.

² Dans la forme actuelle, le processus d'orientation pourrait être qualifié d'évaluation mixte (évaluation interne et évaluation externe), en raison de la prise en compte des notes des bulletins scolaires (75%) associées aux résultats de deux épreuves communes (25%).

³ Cette recherche concerne un mandat attribué par le Département de la formation, de la culture et des sports (DFCS) de la République et Canton du Jura au Département de la recherche de la HEP-BEJUNE. Ce mandat a été confié au domaine de recherche « Apprentissage et évaluation ». La demande s'inscrit dans le cadre de l'évaluation des épreuves communes et se donne pour objectif de dresser un état des lieux de la situation et notamment de déterminer le positionnement du personnel enseignant vis-à-vis des épreuves communes. Dans la foulée, il s'agit également de proposer des pistes d'action et de réflexion quant aux possibles évolutions des épreuves communes jurassiennes.



une perspective critique des pratiques évaluatives externes et standardisées. La problématique repose sur une question centrale : quelles tensions liées au sentiment de justice et de responsabilité animent le corps enseignant face à une évaluation externe standardisée à visée d'orientation ?

1. Des épreuves communes pour orienter

Les tests prescrits par le Département de la formation de la République et Canton du Jura soumettent les élèves à trois séries d'épreuves communes dans les disciplines de base : français (langue 1), mathématiques et allemand (langue 2). Lors du passage du test, les réponses sont reportées par l'élève sur un questionnaire à choix multiples (QCM). La première épreuve commune ne compte pas dans l'évaluation sommative, elle est nommée « préparatoire » et vise à familiariser l'élève à la procédure d'évaluation⁴ (Pasche Gossin & Donzé, 2021). Ces évaluations standardisées se réfèrent, dans leur contenu, au programme d'apprentissage de la Suisse francophone, le Plan d'étude romand (PER), adopté par les cantons membres en 2010 (CIIP, 2011). Suite aux tests, les élèves sont répartis dans trois niveaux en fonction de leur rang (par degré de réussite) jusqu'à concurrence de proportions définies *a priori* (soit avant les tests) : 40% accéderont au niveau A (meilleures performances pronostiquées), 35% au niveau B et le 25% restant au niveau C⁵.

En vue de cette décision d'orientation, au terme de la huitième année scolaire, le canton du Jura établit, dans chacune des trois disciplines, le classement des élèves en appliquant un procédé statistique de moyenne entre les deux notes de bulletin semestriel et les résultats obtenus aux deux épreuves communes (originellement de moitié pour la moyenne des deux épreuves internes et de moitié pour la moyenne des deux épreuves externes). Les moyennes des évaluations internes menées par le corps enseignant peuvent être indexées par le service de l'enseignement.

Notons que durant la pandémie de 2020 (pour des questions organisationnelles), la pondération entre épreuve externe et épreuve interne a été modifiée. Le service de l'enseignement de la république et canton du Jura a choisi de passer d'une pondération de deux $\frac{1}{2}$ (moitié pour épreuves externes et moitié pour les épreuves internes) à une pondération de $\frac{3}{4}$ pour les évaluations internes à la classe et $\frac{1}{4}$ pour les « épreuves communes ». Par l'outil statistique, créé au moment de l'instauration des épreuves communes, les résultats obtenus dans chaque discipline sont indiqués en *stanines* (abréviation de « standard nine-point scale »), situant l'élève par rapport aux résultats cantonaux⁶. L'appareillage statistique y est présent : un bureau administratif a été créé (La CEVES), associé à une logistique (tri, corrections des copies) informatisée. Les épreuves communes ne sont pas utilisées à des fins de contrôle. Ces tests font l'objet d'une certaine sanction (entrée au cycle 3 dans un niveau spécifique) avec un enjeu pondéré par deux facteurs : les évaluations internes entrent aussi dans le calcul d'orientation, le premier semestre à l'école secondaire peut être considéré comme transition puisqu'en fonction des

⁴ Notamment aux réponses par QCM auxquelles les élèves ne sont pas familiers.

⁵ Les parents exercent une liberté de choix pour les enfants se situant dans les franges de 5% à la jointure entre les niveaux A et B, ainsi que B et C.

⁶ Cette méthode statistique appelée « échelle des stanines » comprend neuf classes ayant des intervalles égaux entre les classes 2 à 8, comportant chacune un demi-écart type, la cinquième étant centrée sur la moyenne, les première et neuvième classes étant d'étendue illimitée. Le 1^{er} stanine comprend les résultats les plus faibles et le 9^{ème} stanine les résultats les plus élevés. Un stanine, par discipline, est attribué pour chaque épreuve.



moyennes obtenues au premier semestre, les élèves peuvent monter ou descendre d'un niveau.

2. Cadre théorique

2.1 Évaluation, orientation et justice de l'équité

Quand on parle d'équité dans le champ scolaire dans un contexte démocratique faisant référence à un État social et un principe de justice distributive, il s'avère difficile de ne pas prendre en considération la répartition des biens éducationnels par le biais d'un système éducatif (Brighouse et al., 2016). Ces derniers relèvent de connaissances, de compétences, de dispositions et de comportements ou attitudes, appartenant à l'individu, et susceptibles de contribuer à son propre épanouissement comme à celui des autres. L'évaluation scolaire a pour effet indirect la répartition des biens éducatifs : la fonction de classement par l'évaluation est une des composantes fonctionnelles de l'orientation. Dans une perspective rawlsienne, la structure politique et démocratique de base d'une démocratie sociale suppose un aménagement selon le principe d'équité où « les attentes des plus désavantagés tout au long de leur vie ... soient aussi grandes que possible dans le contexte d'institutions fixes garantissant les libertés de base égales et établissant une égalité équitable des chances » (Rawls, 2009, p. 14). Partant du principe que « le public n'est pas un objet qui s'impose par l'expérience » (Hulak & Girard, 2018, p. 227), comment s'articulent la répartition du bien public et les biens éducatifs ?

Dans sa dimension tangible, l'évaluation scolaire pose, en implicite, un diagnostic, du moins temporaire, de possible accession aux positions sociales : diagnostic parfois contaminé par l'effet de l'inflation scolaire dans une « méritocratie » souvent factice (Duru-Bellat, 2006). Ainsi, comme le résume Coulangeon (2021) à propos des travaux de Dubet (2004) et Duru-Bellat (2009),

L'idéologie du mérite scolaire, consacré par la massification de l'éducation, tend à faire peser sur l'individu une responsabilité démesurée de ses succès comme de ses échecs, du fait d'une forte intériorisation de la légitimité de verdicts vécus sur le mode de la sanction, de l'échec ou de la réussite individuelle. (p. 175)

En cela, les systèmes scolaires produiraient des « hypertrophies » de justice dans les appréhensions des élèves, des parents et du corps enseignant à (Dubet & Duru-Bellat, 2010). D'autant que la massification de l'éducation n'empêche pas la gradation des extrêmes et le déclassement, la courbe des diplômés se dessine souvent en « U », selon Coulangeon (2021) se référant aux travaux de Torche (2011). Cette courbe s'expliquerait notamment par la thèse de Falcon et Bataille (2018), selon laquelle l'effet décroissant de l'origine sociale se limite à l'accès à une égalisation de la reproduction sociale pour des formations moyennes et non l'accès aux hautes études.

Ceci n'est pas sans effet sur les croyances et pratiques des enseignant·es quand survient la « justice » d'orientation. Certaines recherches sur les effets de la croyance en la méritocratie chez les enseignant·es montrent que plus les individus estiment que l'école récompense l'effort et le mérite, moins les personnes s'engagent dans la mise en place d'une pédagogie



égalitaire (e.g., Darnon et al., 2018). La question de ces effets doublement délétères⁷ se pose d'autant plus fortement dans le cadre d'une évaluation externe à la classe où la fonction de sélection semble encore plus marquée. Ce biotope est fécond pour faire croître l'idée que la méritocratie scolaire « not only entails believing that processes in schools are fair but also that the differences in life outcomes based on the educational hierarchy are also fair » (Batruch et al., 2022, p. 12).

2.2 Conceptions de l'évaluation

2.2.1 Évaluation mesure et évaluation soutien d'apprentissage

Ces épreuves s'inscrivent dans la logique de l'évaluation-mesure, évaluation reposant sur des standards de références observables et mesurables, des épreuves ponctuelles standardisées validées par des modèles psychométriques, où l'idée de mesure renvoie à un « quasi-marché » (De Ketele, 2010) : les places dans les niveaux d'enseignements segmentés étant limitées. Vial (2012) définit de manière archétypique l'évaluation-mesure avec, pour fer de lance, la métrique et la docimologie comme science des examens, suivi historiquement de l'évaluation-gestion ayant pour but la rationalisation par les objectifs comme aide à la prise de décision. Ce qui fait référence à la première génération de cette conception de l'évaluation s'appuyant sur des outils de mesure, d'indicateurs de performance individuelle et de tests standardisés renforcée par un outil numérique ne visant pas une évaluation durable (Younès, 2020). La conception « de l'évaluation de la maîtrise des apprentissages » comme la conception de l'évaluation soutien d'apprentissages (ou *Assessment for Learning*) se sont construites sur les critiques que l'on pouvait porter au modèle de l'évaluation mesure (De Ketele, 2010), principalement parce que les résultats de l'évaluation-mesure ne sont en général pas construits pour servir et réguler les apprentissages dans une fonction formative (De Ketele, 2010). Black et William (2018) soulignent aujourd'hui l'importance de ne pas user de l'évaluation formative comme « théorie du n'importe quoi » en faisant abstraction, dans le champ de l'éducation, de la démarche sommative et de la fonction certificative.

Les épreuves critériées ou assistées par l'informatique à des fins d'évaluation diagnostique fournissent souvent une « illusion instrumentale » (e.g., Bain, 1988 cité par Mottier Lopez & Laveault, 2008) et amènent le risque d'élaguer la réflexion théorique sur les processus d'enseignement/apprentissage rattachée à l'évaluation formative (Mottier Lopez & Laveault, 2008).

2.2.2 L'évaluation standardisée

L'évaluation standardisée peut être définie, selon Mons (2009), comme « une évaluation qui a fait l'objet d'un traitement statistique dont les résultats qui en découlent sont relativement fiables en raison des items censés mesurer de façon quantitative des savoirs » (p. 100). Mons la qualifie « d'évaluation quantitative » dont les fonctions et les effets attendus sont à la fois « un outil de pilotage (ou de régulation) des systèmes éducatifs » tout comme un « instrument pédagogique des acquis individuels des élèves » (p. 101). Cette évaluation standardisée, fondée sur une « redevabilité » des écoles publiques, s'établit sous l'influence d'une perspective économiste du *New public management*. Au-delà des résultats scolaires, l'évaluation peut donc renseigner aussi au sujet des performances des

⁷ Échec de l'élève empiré par une absence de soutien pédagogique.



établissements scolaires, voire être utilisée à des fins de contrôle du corps enseignant (Maroy, 2008, 2012) tout en garantissant des standards aux résultats qui décideront de l'orientation. Mons & Dupriez (2010) distinguent deux types de redditions de compte : l'*accountability* dure, fondée sur des tests à forts enjeux (*high-stakes testing*) dont les résultats font l'objet de sanctions et de récompenses pour les écoles, pour le corps enseignant et les élèves (par exemple la voie vers des diplômes ou le redoublement des élèves) ; l'*accountability* douce qui consiste à réaliser des tests à faibles enjeux dont les résultats servent de référence au corps enseignant pour améliorer son professionnalisme et la qualité de son enseignement. Ce qui équivaut, à certains égards, à la régulation des apprentissages par l'évaluation externe à la classe. Les évaluations externes à visée d'orientation devraient participer à la création d'une culture évaluative comme « culture commune en véhiculant auprès des enseignant·es la nécessité d'assurer [des] étayages » (Tessaro, 2015, p. 59).

2.2.3 Corollaires de l'évaluation externe

Les différents travaux de recherche en matière d'évaluation externe (et souvent standardisée) ne permettent pas de révéler d'effets significatifs des bienfaits, tant sur les plans de l'efficacité que de l'égalité scolaire. Et pourtant, de nombreuses enquêtes conduites aux États-Unis (Jones, 2007), dans les pays du Nord (Helgoy & Homme, 2007), en Suisse (Yerly, 2014 ; Yerly & Maroy, 2017), en France (Demailly, 2001, 2006 ; Mons, 2009) montrent qu'une partie du personnel enseignant perçoit parfois de façon positive ce type de tests, notamment parce que ces derniers tendent à donner des repères par rapport à des standards et de mettre en évidence les faiblesses et les forces des élèves, pour autant bien sûr que les tests soient utiles à de futurs apprentissages.

En ce sens, ce type d'évaluation externe et standardisée a un impact sur les pratiques pédagogiques (e.g., Suchaut, 2008 ; Tessaro & Ntamakiliro, 2010). Selon Tessaro (2015), les régulations qui résultent des épreuves externes donnent lieu à des régulations, par exemple, à une adaptation de l'enseignement pour la cohorte d'élèves suivante. L'enseignant·e se soucie dès lors de la correspondance entre les résultats aux épreuves externes et sa propre appréciation (Tessaro, 2015). En proposant des repères susceptibles d'engager une régulation des pratiques, les épreuves externes peuvent être perçues comme complémentaires aux évaluations sommatives internes à la classe. Les épreuves externes contribuent donc à l'alignement et à la cohésion interne des enseignements (Laveault et al., 2014).

D'autres recherches révèlent l'émergence d'entraînements intensifs aux tests, au détriment d'un enseignement soutenant les apprentissages : il s'agit du phénomène bien connu du « teaching to the test » identifié dans la littérature scientifique (Bélair, 2005 ; Jones, 2007). Il est également constaté un rétrécissement des curricula : le corps enseignant oriente son enseignement sur les apprentissages spécifiques qui seront évalués, souvent appuyé de méthodes pédagogiques qui font appel à la seule mémorisation ou l'apprentissage spécifique détaché du développement de compétences réelles⁸. Ces tests peuvent également conduire selon Mons (2009) à « un sentiment de déprofessionnalisation, qui se traduit par des phénomènes de démotivation du corps enseignant » (p. 122). Ces dispositifs sont perçus comme négatifs et mis en cause avec plus ou moins de virulence par le corps enseignant (Mons, 2009) principalement quand il n'est pas tenu compte des

⁸ Entendons ici l'idée de compétence aussi bien « affectives, sociales, sensorimotrices » que cognitive, qui suppose une « actions finalisées », la compétence étant « fondée sur l'appropriation de modes d'interaction et d'outils socioculturels » (Allal, 2002).



caractéristiques sociales des acteurs et actrices concerné·es. Notons que, selon Mons (2009), plus l'évaluation standardisée à forts enjeux oriente les démarches pédagogiques du corps enseignant et s'éloigne d'une évaluation qualitative continue du travail des élèves, plus les résistances du corps enseignant apparaissent nombreuses.

Enfin, les institutions peuvent être amenées à isoler certaines catégories d'élèves qui ne seront pas soumises aux tests pour favoriser un score élevé de réussite (Bélair, 2005). Selon Demailly (2001) les résistances concernent à la fois les intérêts du personnel enseignant, leurs valeurs, comme les ressources cognitives invoquées et perçues comme plus ou moins utiles. Les résultats des épreuves externes et standardisées apparaissent souvent comme incomplets aux yeux des enseignant·es, car ils ne portent que sur une partie de la mission de l'école (Yerly, 2014).

2.3 Une objectivité de faits et de valeurs

Les résultats de l'évaluation standardisée semblent relativement solides en raison d'un traitement statistique de données fiables (Mons, 2009). Survient donc la question réelle de la définition même du degré de « *fiabilité* » estimé par calcul statistique de résultats. Ce relevé légitime est érigé, en quelque sorte, en « standard » normatif par une axiologie construite. Mais selon quelle conception de l'objectivité ?

L'objectivisme considère usuellement que l'objectivité repose sur des faits empiriques identifiables et accessibles à l'expérience de toute personne. Un objectivisme non naturaliste considérerait par ailleurs que ces qualités objectives sont relatives aussi à des *valeurs* (souvent liées à des obligations sociales ou morales).

À l'objectivisme qui repose sur des éléments tangibles⁹ (objets mesurés ou valeurs partagées) s'oppose le subjectivisme empreint de relativisme, d'absence de reconnaissance, voire l'artificialité. À titre d'exemple, le jugement émanant d'un seul sujet-évaluant agissant sans égard pour le sujet-évalué, le traitant comme un objet pour parvenir à des fins autres que pédagogiques. L'artificialité mènerait à évaluer un élément étranger aux apprentissages sans aucun lien avec l'objet ou l'action évaluée (par exemple : enjeu stratégique, intérêt masqué...).

Tableau 1

Conceptions du subjectivisme/objectivisme

Conceptions du continuum objectivisme/subjectivisme	Types de faits et de valeurs associés
Objectivisme de faits et de valeurs	Associée à des faits et des valeurs qui existent indépendamment du sujet évaluant (obligations morales, dialogues, intuitions...)
Relativisme de faits et de valeurs	Qui n'est pas associée à des faits ou des valeurs qui existent indépendamment du sujet-évaluant (obligations morales, dialogues, intuitions...)
Artificialité de faits et de valeurs	Qui construit des faits et valeurs mesurés à seule fin stratégique (intérêt de personne ou de groupe, déclassement, défense de privilèges...)

⁹ Ce qui est « tangible » selon la définition de Chateauraynaud (1991), soit ce qui résiste à l'épreuve de vérité.



À titre d'exemple, la théorie de l'évaluation authentique souligne que les tests de masse, dans leurs standards, évitent la prise en compte du contexte des apprenant·es, par artificialité, car les normes et les tâches authentiques pour juger des capacités intellectuelles demandent beaucoup de travail et de temps (Wiggins, 1989a). Le problème de l'artificialité liée au manque d'éléments tangibles met à mal l'objectivité de l'évaluation ; même appuyée de nombreux critères, quand les problèmes sont artificiels, et les issues ne répondent pas au réel, l'évaluation manque d'ancrage au « réel » (Wiggins, 1989a) et risque un relativisme, vecteur de création ou de perpétuation d'injustices.

Ainsi, la procédure d'évaluation se contente souvent de spécifier une catégorie d'objectifs en combinant les contenus et les qualifications clés (Mottier Lopez & Laveault, 2008), comme l'on combinerait faits et valeurs au cœur d'un seul et même produit chiffré, la note. Ce produit s'apparente au résultat de la démarche sommative : *summa* (sommet) et *sigma* (somme) d'une évaluation. Reste que sommet et somme ne sont jamais qu'un point fixe, une suite finie qui valide une norme particulière et contingente d'une dimension du réel¹⁰, norme qui peut apparaître ailleurs comme artificielle.

Force est de constater que la mesure et ses « produits » suscitent souvent d'importantes résistances auprès des acteurs concernés, car ils apparaissent hors de leurs réalités et de leurs expériences quotidiennes, et peut-être aussi parce qu'ils entretiennent une certaine confusion entre des résultats qui concernent des actions collectives et la perception d'une évaluation qui, de fait, interpelle aussi l'individu. (Mottier Lopez, 2013, p. 944)

Dès lors, la définition même de ce que l'on considère comme objectif, soit les faits et les valeurs mesurés, impacte le modèle sur lequel est construite et/ou estimée l'évaluation externe, modèle articulant par là même le micro de l'évaluation (interne à la classe) au macro (du système scolaire, des attentes sociales ou politiques).

2.4 L'orientation et la justice d'orientation

L'évaluation à visée d'orientation est souvent associée à la « préparation d'une nouvelle action » : soit ce qui vise à organiser une nouvelle année ou une nouvelle filière d'études pour un élève (e.g., De Ketele, 2010), diagnostic qui contribue à orienter la personne évaluée vers un contexte de formation future. Pourtant l'évaluation à visée d'orientation ne peut être imperméable à d'autres fonctions de l'évaluation relevées par Bonami (2005) : par exemple, l'aide à la décision dans le pilotage de l'unité considérée du système scolaire ; la mobilisation sur des objectifs (projet d'établissement, amélioration de la qualité, activités de remédiation, apprentissage par cycles) ; la certification attestant l'atteinte d'objectifs (standardisation des diplômes et homogénéité au sein de l'unité considérée du système scolaire). Enfin, l'évaluation à visée d'orientation renvoie à la reddition de compte : aux politiques, à l'opinion publique, aux parents, aux collectivités locales (Bonami, 2005).

Vue sous l'angle d'une fonction univoque, l'évaluation à visée d'orientation génère possiblement ambivalences et tensions. Ne considérer, de manière monosémique, qu'une seule de ses fonctionnalités offre un point de vue limité sur la justice scolaire et sociale, comme sur la répartition des choix décisionnels d'orientation. Ainsi, deux larges logiques dirigent l'évaluation en éducation, celle du « contrôle et attestation » qui s'oppose à celle du « développement formatif » ; cette dichotomie croise celle qui oppose la valeur

¹⁰ Par réel, nous pouvons évoquer de manière pragmatique « tout ce qui résiste aux variations perceptuelles, instrumentales et argumentatives auxquelles le soumettent des acteurs dotés de représentations et d'intérêts divergents » (Chateauraynaud, 2004).



certifiante ou des évaluations internes et externes (Mottier Lopez, 2009). Bulletins, notes par exemple se confrontent aux auto-évaluations ou évaluation régulative, par exemple, pour l'évaluation interne, et tests pour décision d'orientation à des tests à valeur diagnostique, pour l'évaluation externe (Mottier Lopez, 2009).

2.5 Segmentation et justices

Plusieurs études montrent que la segmentation par niveaux d'enseignement tend à accroître les écarts de résultats entre élèves (e.g., Dupriez, 2010). À cela s'ajoute que le regroupement des élèves les plus faibles contribue à créer des situations difficiles sur le plan socio-pédagogique (Dupriez, 2010). Les cantons suisses qui font le choix d'une « organisation segmentée de leur enseignement secondaire obligatoire sont ceux pour lesquels les acquis scolaires dépendent le plus fortement de l'origine sociale des élèves » (Felouzis & Charmillot, 2017, p. 11), quand bien même « les élèves plus faibles progressent moins bien quand ils sont regroupés dans des filières ou des classes à niveaux, où ils bénéficient d'un enseignement moins exigeant, que lorsqu'ils sont mélangés avec les élèves forts dans des classes hétérogènes où ils suivent un enseignement plus exigeant » (Ntamakiliro & Daepfen, 2016, p. 8).

Malgré cela, le pilotage des résultats des élèves par niveau dans les voies d'accès en formations tertiaires reste une des fonctions de l'école obligatoire en Suisse. Alors que les formations tertiaires sont souvent considérées comme garantie de trouver un emploi résistant à la crise comme à de meilleures perspectives d'avenir (OCDE, 2011), Duru-Bellat et Kieffer (2008) montrent que sur les seuls bacheliers français, les enfants d'ouvriers représentent 38 % des entrants au collège et ne forment que 9 % des entrants en classes préparatoires aux grandes Écoles (Lemaire, 2008). En Suisse, les enfants, dont les parents ont au plus achevé la scolarité obligatoire, ne représentent que 6,7% des étudiant·es entrant dans les hautes écoles (OFS, 2021). Les biens éducationnels constituent des biens positionnels, d'où la prise en considération majeure de la justice comme équité en social-démocratie dans la construction des normes sociales comme dans le pilotage des systèmes éducatifs.

Another reason to consider equality is that some educational goods are, in part, what economists call positional goods. That is, their worth, to those who have them, depends in part on how much they have relative to others. When competing for jobs and other positions that are allocated (partly) on the basis of one's holdings of educational goods, what matters is not one's own absolute level, but one's level relative to that of one's competitors (Brighthouse et al., 2016, p. 13)

Cette pensée politique conséquentialiste, qui lie biens éducatifs et biens positionnels suggère dès lors qu'il pourrait exister des tensions (adhésion ou non à des conséquences de justice pressenties)¹¹ et ambivalences (quant à l'objectivité)¹² liées au sentiment de justice et de responsabilité au sein du corps enseignant. Comment s'animent ces représentations et prises de positions face à une évaluation externe standardisée à visée d'orientation ?

¹¹ Terme de tension issu d'autres recherches qui montrent par exemple l'adhésion ou non à l'évaluation issue d'un pilotage du système, les enjeux de pouvoir (Yerly, 2022).

¹² L'ambivalence, terme hérité également, renvoyant à cette légitimité reconnue aux évaluations externes (Yerli & Maroy, 2017).



3. Méthodologie

Cet article repose sur une recherche menée de 2019 à 2021 dans le canton du Jura sur mandat de son Département de la formation, de la culture et des sports pour dresser un bilan relatif à la procédure des épreuves communes et pour déterminer le positionnement des différentes parties prenantes impliquées dans le dispositif cantonal d'orientation des élèves. Un premier rapport intermédiaire a été rédigé, en septembre 2020, sur la base d'une enquête exploratoire qui a mené à décrire et mieux comprendre le phénomène tel qu'il est vécu par les acteurs et actrices concerné·es. La méthode reposait sur un entretien semi-directif sur la base d'un guide d'entretien construit sur une base thématique. Dix-huit entretiens enregistrés d'une moyenne de soixante minutes ont été réalisés. Les résultats de cette première phase ont permis d'effectuer une première analyse qualitative des entretiens, ainsi que de permettre la construction d'un questionnaire diffusé ensuite auprès du personnel enseignant (7-8P et 9S) durant l'année scolaire 2020-2021. Le taux de réponse pour le questionnaire en ligne a été de 69% pour les enseignant·es du 7^e et 8^e degrés (134 invitations, 87 réponses) et de 58% (191 invitations, 111 réponses) pour les enseignant·es du 9^e degré (premier degré du 3^e cycle). Ces données ont permis la rédaction d'un rapport final (Pasche Gossin & Donzé, 2021).

Les résultats présentés dans cet article constituent un prolongement aux deux rapports de recherche (2020 et 2021) déposés auprès du Département de la formation, de la culture et des sports de la République et canton du Jura. La méthodologie s'appuie sur une approche compréhensive (Kaufmann, 1996/2016) en vue de mettre au jour les unités composites (Mbiatong, 2022) qui articulent situation, objets, discours, sentiments, motif d'action, décrire le processus de construction de sens des enquêté·es, décrire leur activité interprétative. La scène des évaluations externes et standardisées est bien connue du fait d'un grand nombre d'enquêtes sur ce modèle d'évaluation. Les références théoriques pour faire progresser l'argumentation nous ont permis de baliser le degré d'adéquation (Kaufmann, 1996/2016).

L'analyse des questionnaires repose sur une analyse statistique à plat. L'analyse des entretiens, quant à elle, s'appuie sur l'analyse de discours et la construction de catégories sémantiques à partir d'unités de sens, par méthode de comparaison et mise en relation.

4. Résultats et interprétation des résultats

Les résultats présentés ici trouvent leur origine dans une recherche mandatée visant à faire un état des lieux des avantages et inconvénients de certaines modélisations d'épreuves d'orientation dans un territoire spécifique. Cette recherche a permis, à un second niveau, de mettre au jour les tensions et ambivalences chez les enseignant·es eu égard à ce modèle d'orientation comme aux conceptions de l'évaluation à visée d'orientation. L'analyse des données sous le jour des premières catégories de tensions et d'ambivalences nous a amené à une construction conceptuelle d'autres catégories principales (Glaser, 2002) que nous espérons suffisamment solides¹³ pour être réutilisables en d'autres contextes (sociaux, politiques).

Cette conceptualisation, présentée ci-dessous sous forme de schéma, vise à circonscrire la complexité du terrain étudié selon quatre sources de tensions corrélées (figure 1) qui

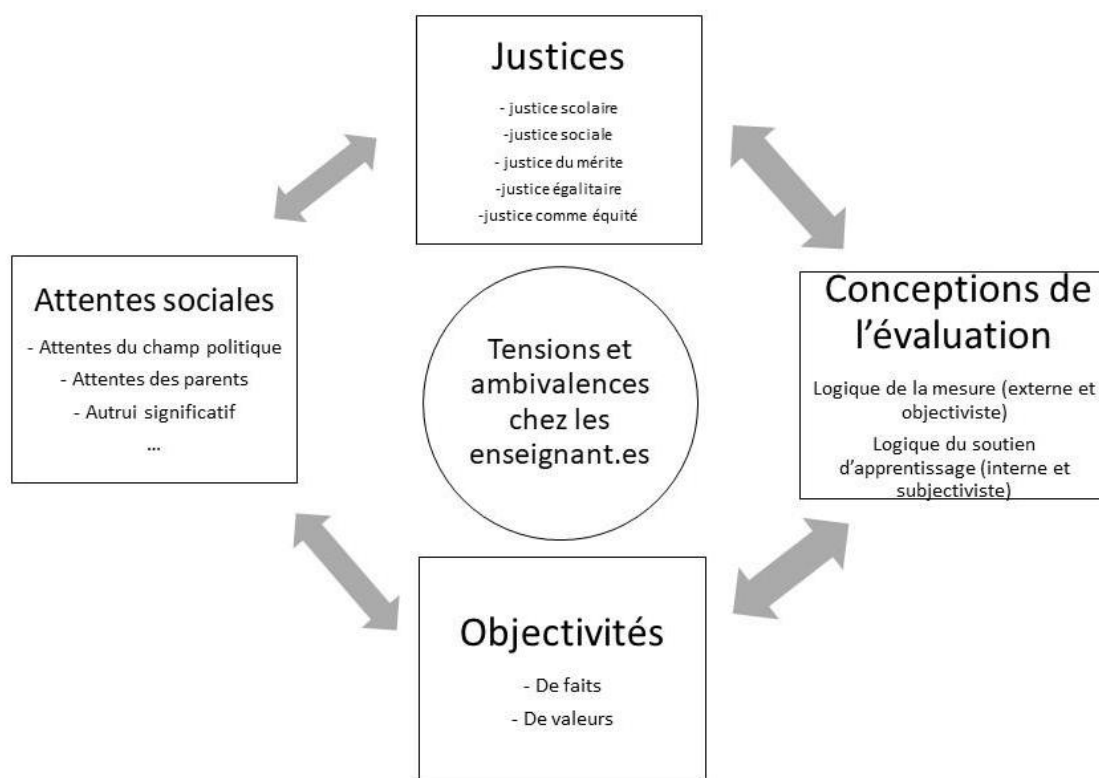
¹³ « Abstract of time, place, and people, and that concepts have enduring grab » (Glaser, 2002, p. 1).



conduisent à des compromis et des justifications divergentes en lien avec le sentiment de responsabilité du corps enseignant. Partant des tensions et ambivalences décelées, l'analyse interroge la réalité de l'objet de recherche, en lien avec les différents contextes qui en émanent, au travers des discours et des positionnements des enseignant·es.

Figure 1

Modélisation de quatre sources de tensions et d'ambivalences chez les enseignant·es



4.1 Ambivalences entre justice sociale et justice scolaire

La justice sociale pourrait être définie sommairement comme une justice en prise avec la réalité sociale large d'individus d'une société donnée (sous l'effet de principes éthico-politiques et/ou de croyances : en la justice du mérite, justice distributive...). La justice scolaire prendrait dès lors la forme de l'effet (réussi ou non) de l'application de ces principes dans le champ scolaire ; notamment par le choix de diverses pratiques évaluatives. Notons toutefois que les notions s'entremêlent quand « la réduction de la justice scolaire à une pure méritocratie chez bien des sociologues, des enseignants et des experts ... trouvent là des modes d'évaluation des systèmes commodes » (Duru-Bellat & Meuret, 2009, p. 55).

Le corps enseignant du cycle 2, chargé de préparer les élèves aux épreuves communes, a un avis partagé sur la question des inégalités : 45% (36, n=81) estiment que la procédure d'orientation amène à créer de l'inégalité scolaire et sociale (ces catégories émergent du cadre théorique confronté à l'enquête qualitative préliminaire de la recherche mandatée). Du côté du corps enseignant du troisième cycle, qui accueille les élèves répartis par niveaux, l'avis est plus tranché : seul un tiers (27%) (29, n=105) estime que cette procédure



standardisée d'orientation amène de l'inégalité scolaire et sociale (Pasche Gossin & Donzé, 2021).

Une première ambivalence oppose ainsi *la présomption d'injustice dans la procédure d'orientation* à un *sentiment de nécessité systémique de classement des élèves*. Cette tension est d'autant plus forte quand l'on fait référence à des particularismes (cognitifs ou sociaux) chez les élèves, la prise en compte des contingences du sujet-évalué. Plus de la moitié (56%) du corps enseignant enquêté est d'avis que ces tests entraînent une forme d'injustice scolaire et sociale pour les élèves allophones (98, n=175) et 59% pour les élèves présentant un trouble spécifique de l'apprentissage (105, n=176) (Pasche Gossin & Donzé, 2021). Les chiffres montrent toutefois divers indices d'opinions controversables : 66% des enseignant·es qui préparent les élèves aux tests (cycle 2) estiment que les épreuves communes entraînent des injustices pour les élèves allophones contre 48% du cycle 3. En ce qui concerne notamment les troubles spécifiques des apprentissages, la proportion est encore plus marquée : 79% des enseignant·es du cycle 2 considère que les épreuves conduisent à des injustices, contre seulement 43% du cycle 3 (Pasche Gossin & Donzé, 2021).

Dès lors, la compréhension du contexte, de la réalité sociale et des spécificités cognitives des élèves semble *a priori* plus prégnante chez les personnes accompagnant les élèves dans le passage des tests d'orientation. Ainsi, 87% de ces enseignant·es soulignent une forme d'injustice pour les élèves sujets au stress (Pasche Gossin & Donzé, 2021).

Les trois verbatims ci-dessous montrent un parfait continuum de tension et d'ambivalence, au sein des représentations des enseignant·es, quant à la congruence entre justice sociale et justice scolaire. Cette congruence influe sur la force du sentiment de responsabilité, cet engagement à produire ou à empêcher certains résultats (Lauermann & Karabenick, 2011). Ce sentiment, associé à la maîtrise de la réussite des apprenant·es (apprentissage, performance et progrès scolaires durant l'année), constitue l'un des enjeux forts du sentiment de responsabilité des enseignant·es (Vaudroz & Berger, 2019).

Tableau 2

Congruence et incongruence entre justice sociale et justice scolaire

Relation entre justice sociale et justice scolaire	Verbatims exemplifiant la relation	Conséquences sur la perception de la subjectivité et le sentiment de maîtrise
Congruence entre justice scolaire et justice sociale	« Je suis favorable parce que je constate que les épreuves communes reflètent chez moi totalement mes notes . Ça me donne une possibilité de pouvoir argumenter avec les parents . Donc pour moi, elles sont assez bien adaptées à ce qu'on doit faire » (Enseignant 2)	Sentiment d'équivalence entre mesure externe et mesure interne : objectivité de faits et de valeurs Sentiment de maîtrise en lien avec les attentes sociales.
Incongruence possible et doute sur la justice scolaire	« Si on était les seuls à détenir ce pouvoir de classement , s'il y avait que nos notes finalement qui comptaient, j'ai l'impression que les parents ne nous feraient pas confiance, donc finalement c'est un peu une assurance : l'épreuve commune va confirmer les notes qu'on a mises à l'élève, donc c'est aussi un peu une protection » (Enseignant 1)	Sentiment de possible effet d'artificialité et de subjectivité dans la mesure interne comme externe. Sentiment de maîtrise fragile par un certain artifice de protection qui manque peut-être d'objectivité



<p>Incongruence entre justice scolaire et justice sociale</p>	<p>« En même temps, chaque année, on repart avec une nouvelle volée puis on se dit : celle-là, ça sera la bonne. Et puis en fait, on se rend quand même compte d'année en année qu'il y a une inadéquation, enfin que c'est dur, c'est trop dur. On cherche les meilleurs. Donc même des élèves qu'on voit là, qui sont fantastiques, qui ont des qualités incroyables on sait que ça ne va pas jouer » (Enseignant 4)</p>	<p>Sentiment de possible effet d'artificialité et d'incomplétude dans la mesure externe : pas d'objectivité de faits et de valeurs Sentiment de non-maîtrise qui a pour cause une attente sociale d'excellence ou de méritocratie factice ; trahison de l'idéal pédagogique</p>
---	--	---

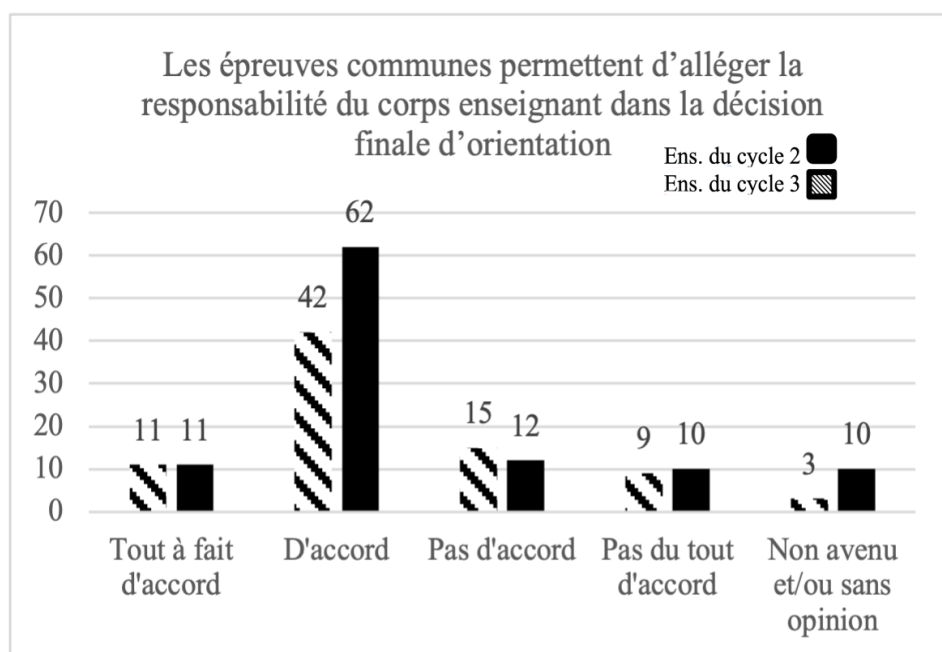
L'engagement et les obligations de l'enseignant·e reposent sur des « repères faisant foi » comme des doutes (Thévenot, 2019). Ainsi, le sens ordinaire du juste se mesure à l'aune d'ordres divers de grandeurs, axiologie de valeurs donnant lieu à moult compromis dans une société plurielle (Boltanski & Thévenot, 1991). Les justifications font référence au contrôle et à la perte de la maîtrise, on y décèle le divorce entre l'idéal de l'enseignant·e et la contingence (exposée comme nécessaire) du système de classement. Les tensions et ambivalences s'animent ainsi par la prégnance des questions de définitions diverses de justices et d'objectivités (voir figure 1).

4.2 Tension entre sentiment de maîtrise et responsabilités multiples

Les données qualitatives comme les résultats quantitatifs (figure 2) révèlent un sentiment d'allègement de responsabilité pour le corps enseignant qui entre en tension avec le sentiment de se faire déposséder d'une responsabilité majeure, notamment eu égard au principe d'éducabilité ou d'un certain idéal pédagogique.

Figure 2

Sentiment d'allègement de responsabilité chez les enseignant·es



Quelques enquêté·es posent l'hypothèse que la confrontation des mesures internes et externes participe d'une évaluation en silos qui a un impact sur la véritable valeur de la décision d'orientation : en priorisant, par exemple, la justification administrative versus la réalité des compétences. Il est à souligner que ces éléments sont mis en doute par les enquêté·es : la concordance entre le bilan communiqué (De Ketele, 2013) aux élèves ou aux parents par le corps enseignant et le classement final entre toutes les élèves du canton ; la correspondance entre les compétences développées en classe et les « standards » testés (De Ketele, 2013). Une chose est certaine, les différents contextes n'agissent pas explicitement de façon dialectique au sein de « l'activité évaluative » (Mottier Lopez et al., 2021).

Selon les chercheurs des espaces de discussion au sujet des modalités d'évaluation permettraient une meilleure mise en lien dialectique ou actionnelle. Aux yeux des enseignant·es enquêté·es, si la responsabilité décisionnelle est allégée, l'engagement pédagogique dans la responsabilité de l'enseignant·e semble mise à rude épreuve quand la « raison administrative » conduit à prioriser et justifier des décisions d'orientation entrant en conflit avec le contexte de la classe et les évaluations internes.

Le sentiment de faible maîtrise des effets désignés (Lauermann & Karabenick, 2011) semble d'autant plus marqué quand il s'agit d'élèves à besoins éducatifs particuliers. « Là j'ai un élève typiquement qui est "cas particulier"¹⁴, on a décidé pendant deux ans que, toutes les notes qu'il aurait, je les noterais, mais j'y mettrais un 4 ». L'enseignante relève ici combien elle situe et différencie son évaluation dans la notation¹⁵ de manière assumée pour soutenir la motivation de cet élève « cas particulier ». « C'est idiot hein, donc il sait que par rapport au groupe classe, il a peut-être un 3,5, mais moi je lui mets 4. Et dans le bulletin, il aura 4. Ça, c'était établi en décision avec le conseil pédagogique ». Puis, suite à une décision de la personne en charge du suivi pédagogique, l'enseignante a le sentiment d'avoir été remise à l'ordre.

[Maintenant] on ne veut plus de ces pratiques qui ont « trop perduré », maintenant c'est comme ça. On lui met les notes qu'il vaut, et s'il ne peut pas obtenir la moyenne, alors on devra le sortir du système, et puis le prioriser.

L'enseignante témoigne d'une souffrance dans cette perte de maîtrise et dans l'impossibilité d'incarner son rôle d'accompagnante dans une évaluation qui soutiendrait les apprentissages et atteindrait la motivation intrinsèque (Pulfrey et al., 2011) :

Très bien donc là, depuis 3 semaines, je lui mets des 2, des 3. Il pleure à chaque test et puis il pleure à chaque résultat. Et maintenant il vient me dire « toute façon je suis nul, toute façon j'arriverai pas ». Quel intérêt ? Quel intérêt ?

L'émotion est vive chez l'enseignante concernée, elle tente de regagner la confiance de l'élève par un discours convoquant une idée de la justice sociale détachée du fait scolaire « tu sais moi j'ai un frère qui était C [niveau d'orientation le plus bas], dyslexique, c'était une cata, puis maintenant il est chef d'entreprise ... ». Il est à noter que le corps enseignant endosse une responsabilité forte qui s'inscrit dans une appartenance à des contextes et une communauté large. L'enseignante présente développe une croyance faible en la méritocratie dite *prescriptive* : la manière dont on estime que l'école devrait octroyer le mérite (Duru-Bellat & Tenret, 2012). La critique de la méritocratie *descriptive*, celle des faits,

¹⁴ Cette expression est couramment usitée sur le terrain : les enquêté·es parlent de « cas particulier » pour désigner les élèves non soumis aux épreuves communes (élèves à besoins éducatifs particuliers, allophones...).

¹⁵ Le 4 étant la note qui situe l'élève dans la suffisance sur une échelle de 6 en Suisse.



celle qui légitime la distribution des mérites dans le système scolaire est forte chez l'enseignante concernée. Ce type de tension n'est pas sans effet sur le sentiment de liberté et de responsabilité face aux pratiques pédagogiques permises ou empêchées, chez les enseignant·es enquêté·es. Beaucoup ressentent une tendance à forcer les pratiques dans la mise en place d'une pédagogie « inégalitaire » (Darnon et al., 2018), une tendance à vouloir soutenir l'égalité dans un système inégalitaire. Ce manque de continuité entre les contextes, cette rupture entre méritocratie de droit et méritocratie de faits n'est pas sans rappeler l'idée de décentration selon Lave et Wenger (1991) :

The master as the locus of authority (in several senses) is, after all, as much a product of the conventional, centered theory of learning as is the individual learner. Similarly, a decentered view of the master as pedagogue moves the focus of analysis away from teaching and onto the intricate structuring of a community's learning resources. (p. 94)

L'enseignante endosse un rôle que justifie un idéal pédagogique qui repose sur un principe égalitaire et une justice commutative : face à l'élève en échec elle remet en cause la valeur même des savoirs scolaires (« *j'ai un frère qui était C [niveau d'orientation le plus bas], dyslexique, c'était une cata, pis maintenant il est chef d'entreprise* »), montrant par-là que, selon elle, l'évaluation à visée d'orientation s'inscrit dans une structure bien plus complexe des ressources d'apprentissage en appelant à une communauté large (community's learning resources), à une vue qui ose la décentration de l'enseignant comme pédagogue (view of the master as pedagogue) ; sa réaction agit en révélateur également de l'importance d'une forme de « justice pédagogique » pour les enseignant·es : justice qui articulerait dans la *praxis*¹⁶ justice sociale et justice scolaire.

4.3 Tensions entre conceptions de justices et attentes sociales

Les tensions de justices (pédagogique, sociale, égalitaire, de l'équité, scolaire...) au cœur des représentations touchent aussi bien les enquêté·es actant que ceux qui ne participent que passivement au processus d'évaluation. Par exemple, la sélection objective (justice scolaire), associée à une certaine artificialité de classement (manque d'équité), est pleinement assumée chez un parent d'élève questionné.

La quarantaine, d'un parcours de formation linéaire jusqu'en formation tertiaire, ce parent rêve d'épreuves standardisées (« *qui sont les mêmes pour tout le monde* ») qui permettraient aussi aux enseignant·es de se mettre à jour et que les élèves passeraient « *comme une épreuve normale* » sans s'en rendre compte.

Si la dimension méritocratique des résultats de classement par l'évaluation est discutée (« *on peut remettre un petit peu en question les niveaux* »), l'idée d'un standard plus objectif que l'évaluation émise par le sujet enseignant est inscrite dans ses représentations. Ce parent évoque une évaluation externe plus objective en faits et valeurs (« *je pense qu'avec seulement les notes du bulletin, on ne va pas réussir à classer les gens correctement.* »). Notons que, dans ce contexte, les propos du parent se transforment, pour les enseignant·es, en attente sociale d'un autrui significatif, le parent qui attend un classement « correct » des élèves.

¹⁶ praxis au sens originel de ce qui articule théorie et pratique: idées et mise en mouvement des idées. Ainsi s'agirait-il d'articuler justifications théorico-politiques (au sujet de la justice sociale) et justice scolaire (soit les possibilités de mise en pratique des théories sociales dans le cadre professionnel du métier d'enseignant·e).



À ce point, surgit la conception de l'évaluation dans sa finalité de classement élaguée théoriquement d'autres dimensions (formative, soutien des apprentissages...). La notion de justice de l'évaluation repose ici, en conséquence, essentiellement sur la fiabilité des résultats : objectivité des faits et valeurs mesurés. Cette représentation normative de l'évaluation justifie le processus de classement pronostique visant la segmentation en divers niveaux d'apprentissages au troisième cycle, selon quelques compromis d'artificialité de classement :

et puis justement, justement, tous ces enfants entre guillemets, qui sont un peu moins scolaires, alors je sais pas ... , c'est pas pour dénigrer hein mais les personnes qui travaillent, voilà les deux parents travaillent dans une usine, ils ont déjà peut-être pas, peut-être les capacités peut-être pas le, oui l'envie de...

Cette justification du mode d'évaluation en place repose sur une certaine « nécessité » du classement, fatalité de la segmentation par niveau d'enseignement qui s'oppose à la « contingence » de l'évaluation scolaire menée par les enseignant·es. Il convient d'en extrapoler la naturelle « hypertrophie » d'injustices que ce type de justification peut engendrer chez les parents, principalement quand cette orientation limite les accès à certaines formations ou diplômes (Dubet & Duru-Bellat, 2010).

Ainsi, 54% du corps enseignant enquêté (Pasche Gossin & Donzé, 2021) considèrent les facteurs familiaux comme déterminants pour la réussite des élèves aux épreuves communes. Selon une enseignante, tout est évident, « on cherche les meilleurs. Donc même des élèves qu'on voit là, qui sont fantastiques, qui ont des, qui ont des qualités incroyables, on sait ... qu'on devra être celle qui donne l'étiquette de t'es moyen ou t'es pas bon ». Ces éléments révèlent l'importance de considérer les tensions et ambivalences, au sein du corps enseignant, entre l'appel d'une normativité construite sur une logique de classement (attentes sociales ou attente d'un système politique) et l'appel d'une normativité construite sur la logique d'accompagnement pédagogique (attentes sociales de l'élève, des parents). Un parent relève ce fait :

dès qu'on est en C [niveau d'orientation le plus bas], je pense que ça, au niveau du marché du travail ça coince. Ouais ça permet d'avoir moins d'opportunité et puis c'est ça qui, qui stresse un petit peu tout le monde. C'est un peu plus les débouchés. C'est pas vraiment les épreuves communes, c'est, c'est vraiment la perspective de l'avenir quoi.

Cette analyse compréhensive peut être mise en perspective avec ce qui est relevé par Mons (2009) : pour les parents, les objectifs cognitifs qu'évaluent les tests ne sont qu'une des missions des écoles ; leurs principales inquiétudes restent la sécurité, la discipline, le respect pour les enseignant·es et l'enseignement des valeurs. Le bas niveau scolaire n'est qu'en toute fin de classement. Derrière la justice scolaire se cache pourtant un enjeu souvent social, la peur du déclassement, quand le miroir scolaire révèle la minorisation.

4.4 Tensions face à diverses objectivités et de multiples conceptions de l'évaluation

Une enseignante souligne une autre ambivalence :

On a affaire à des formulations de questions auxquelles les élèves ne sont pas habitués ... ce n'est pas par volonté un peu sadique des rédacteurs, c'est finalement, une chose à



laquelle ils sont contraints. ... Parce que ça a pour but de sélectionner. Si on veut sélectionner alors il faut faire l'outil qui le permet.

Dans les propos de la personne interrogée survient l'absence de rapport authentique au savoir (au sens de ce qui est *truly representative of performance in the field* (Wiggins, 1989b), une artificialité de l'évaluation à visée d'orientation qui, dans une démarche de sélection, ne peut être soutien des apprentissages (ces derniers étant considérés comme moyen plutôt que comme fin).

Dans ce contexte, la démarche de sélection sociale est d'autant plus présente que l'épreuve fonctionne selon le procédé du concours : le nombre de places dans chaque section étant limité à un pourcentage donné *a priori* (40%, 35%, 25%). Apparaît, dans les propos de l'enseignante (« *si on veut sélectionner alors il faut faire l'outil qui le permet* »), l'arbitraire possible d'une orientation non basée sur la juste égalité des chances (d'acquisition de compétences et d'apprentissages), mais qui repose sur la prédominance d'une sélection sociale non centrée sur une évaluation à visée pronostique en vue de la régulation d'apprentissages, mais centrée sur une orientation reproductive d'inégalité.

Autrement dit, sont possiblement évalués des savoirs informels, selon des registres culturels démarquant, entraînant les injustices du déterminisme social bien identifié par la sociologie de l'éducation depuis une cinquantaine d'années (e.g., Duru-Bellat & Van Zanten, 2009). Un enseignant l'affirme :

Si on disait aux écoles secondaires, au 30 juin, voilà la répartition des élèves, c'est inimaginable, ils n'arriveraient pas à mettre en place les horaires pour la rentrée, donc il y a aussi cette obligation.

Ou à d'autres égards :

Si certaines écoles doivent dédoubler un niveau A et d'autres se retrouvent avec un effectif d'élèves de niveau A insuffisant. Ce 40, 35 et 25% ne se justifie plus tout à fait ? Après reste à savoir comment est-ce qu'on établit les niveaux.

Ainsi, selon une enseignante, s'exprimant au sujet du corps enseignant en charge de la rédaction des épreuves communes :

Sur la forme, il y a quand même des choses où ils exagèrent. On a des espèces de consignes avec des négations, des doubles négations, même moi quand je les passais avec mes élèves en même temps qu'eux, pour justement voir la complexité, je trouvais ça pas facile.

Cette remarque met en relief un certain embrouillement des pistes pressenti par les acteurs et actrices engagé·es dans le champ pédagogique, embrouillement quant aux fonctions prioritaires de l'évaluation standardisée. Qui se cache derrière le pronom « ils » ? La ou le collègue en charge de la rédaction des tests ? L'administration ? Surviennent ici des tensions, pour le corps enseignant, entre un idéal pédagogique qui serait inscrit dans une cité inspirée (faisant fi des utilités marchandes) et un idéal civique (visant le bien de tous) et les justifications pratiques liées à l'administration de l'évaluation à visée d'orientation : justifications marchandes et industrielles (Boltanski & Thévenot, 1991) – quand la division du travail social en appelle à la segmentation et au classement scolaire qui oriente vers le monde professionnel. Le compromis semble fragile, peu signifiable et sujet à controverse. Controverses en reflet des critiques de certaines recherches au sujet du « détachement de



l'école des besoins fonctionnels » ou « des effets pervers de la fuite en avant de la sur-éducation » (Coulangeon, 2021, p. 166).

Enfin, nombre d'enquêté·es dénoncent l'argumentaire désuet de l'innéisme de l'intelligence, craignent d'opérer une sélection faisant fi de la justice sociale et du rôle fondamental de l'école, dans sa vision démocratique, celle de maximiser les chances d'acquérir des compétences pour chaque élève et de distribuer les biens éducationnels de manière équitable. Le risque est d'autant plus fort, si la modélisation du test repose sur le principe de l'outil sélectif. Cette réalité sélective est évoquée par une personne en charge de la rédaction des épreuves communes depuis un certain nombre d'années :

On part d'une thématique commune, niveau A on commence en général, et puis on fait une partie B et C, plus ou moins commune, hein avec des petites subtilités et puis on différencie les niveaux A, où là on fait quand même une épreuve nettement plus difficile. On dira ça comme ça. Et le but c'est que dans la même thématique on ait trois épreuves cohérentes, mais avec des critères qui sont propres à ces niveaux.

Si peuvent être entendu des « choix négociés » (Tessaro, 2015) chez la personne en charge de la rédaction des épreuves, peut être perçu aussi dans son discours, la nécessité de définir *a priori* des niveaux de compétences ayant pour objectif le classement.

Selon Tessaro (2015) la proportion des tâches complexes présentes dans l'évaluation standardisée est inférieure à celle des évaluations faites en classe, car elle reste difficile à évaluer. Apparaît ici le risque d'un rétrécissement curriculaire en des référés limités (et limitants) pour légitimer une future segmentation de l'école par niveaux. Dit autrement, l'évaluation standardisée et la modération statistique dépossèdent les enseignant·es de certaines compétences évaluatives (Laveault & Yerly, 2018), et donc d'une responsabilité.

Tableau 3

Inégalités et injustices selon les enseignant·es enquêté·es

La procédure d'orientation amène à	Cycle 2	Cycle 3
créer de l'inégalité scolaire et sociale	45%	29%
des injustices pour les élèves allophones	66%	48%
des injustices pour les élèves ayant des troubles spécifiques des apprentissages	79%	43%

Les craintes souvent évoquées des enquêté·es au sujet de l'injustice de la sélection opérée (tableau 3) relèvent parfois d'un sentiment de faible maîtrise des résultats de leurs enseignements face aux sources d'arbitraires et de déterminismes :

- La focalisation sur des tâches simples au détriment des tâches complexes en action en vue de développer des compétences qui supposent des « actions finalisées » et non une « mesure statique de l'état des capacités » (Allal, 2002)
(« Si je prends le français, on nous dit tout au long de l'année, il faut faire écrire aux élèves, il faut travailler sur leur production (...) Et puis là, on a un truc qui fonctionne totalement à l'inverse. Les élèves doivent faire des croix. On dit, en allemand, maintenant, il faut inviter des élèves à s'exprimer. Mais ils ne s'expriment pas dans les tests. Ils posent des croix »)



- L'usage du capital culturel comme outil de démarcation sociale (Bourdieu & Passeron, 1977) ou la mise à l'écart de minorités par l'usage de registres d'apprentissages en appelant à des habitus symboliques ou culturels (Rayou, 2020). (« Ces épreuves sont très culturalisées, c'est la culture qui fait foi et certains élèves allophones sont pénalisés (...) C'est un désavantage pour les enfants qui ont des difficultés avérées »)
- L'usage du savoir académique « expert » sacrifiant au principe de démocratie (Lemay, 2000) (« Le français, c'est une machine à piège. Le troisième cahier c'est le plus difficile. C'est toujours un peu le cahier où l'on se casse la figure »).

4.5 Ambivalence entre objectivité annoncée et place du sujet dans la logique de classement

À la question « les pourcentages (40%=A, 35%=B, 25%=C) devraient varier d'une année à l'autre ? », soit à cette possible répartition arbitraire d'une population, la polarisation est peu marquée : le corps enseignant enquêté, dans une toute petite majorité, se prononce toutefois en faveur de la suppression des contingents *a priori*, de l'évaluation-concours (Pasche Gossin & Donzé, 2021). C'est le personnel enseignant au cycle 2, en charge de préparer les élèves aux épreuves communes, qui se prononce dans une plus large majorité pour la fluctuation des pourcentages de répartition selon le niveau de chaque volée d'élèves¹⁷. Étant donné que les préoccupations administratives (horaire, effectifs des classes...) et d'enseignement (homogénéité des classes...) occupent davantage les intérêts du corps enseignant engagé dans la formation secondaire, on peut identifier ici la justification d'intérêt administratif. Cette justification entre en contradiction avec l'objectivisme comme la justice du mérite, d'autant si le choix systémique de la segmentation de l'enseignement secondaire obligatoire repose essentiellement sur l'origine sociale des élèves (Felouzis & Charmillot, 2017, p. 11).

Pour les enquêté·es, il n'y a pas de connaissances des attentes officielles exprimées par la gouvernance et l'on ne peut dès lors construire un consensus au sujet des pratiques d'évaluation comme le propose Wyatt-Smith et al. (2010) cité par (Tessaro, 2015). Ce qui semble être au centre de l'ambivalence pour une partie corps enseignant enquêté, c'est donc la fonction, qui leur est assignée, comme leur rôle, dans les dispositifs de pilotage du système : à travers quoi se joue la nature de leurs rapports à l'État ainsi que le statut de leur profession (Crahay, 2009), dans la prise en compte de chaque individu. Ainsi, comme le souligne une enseignante fort impliquée dans un engagement face à l'élève :

Là, on a 3 filles en français, elles méritent amplement un niveau A. Mais il y en a une qui s'est vraiment plantée aux épreuves communes et ce n'est pas du tout mérité, par rapport aux capacités, à ce qu'elle démontre en classe. Elle est clairement d'un niveau A, à l'école secondaire. Mais avec le stanine qu'elle a fait, cela l'en empêche. Ma moyenne... vont-ils peut-être la descendre ? Enfin, je ne sais pas, du coup... voilà, c'est ça que je ne trouve pas juste. Ce n'est pas du tout mérité.

Cette enseignante perçoit l'élève sous le jour d'autres axiologies d'« évaluation » inscrites dans des contextes plus larges. L'axiologie de l'évaluation interne n'étant pas nécessairement alignée avec l'axiologie de l'évaluation externe. L'indexation opérée par la gouvernance des moyennes de l'enseignant·e (évaluations internes à la classe) met donc fortement en doute, avec une certaine violence, l'objectivité du jugement professionnel de l'enquêté·e.

¹⁷ 47 contre 20 pour le primaire ; 49 contre 35 pour le secondaire, 36 sans opinion, n= 175.



Selon un autre enseignant, assumant plus pleinement la dimension émotionnelle : « on a 35 semaines, fois 5. Voilà, ça fait 175 jours et puis il y en a 6 qui sont pris en compte en comptant pour moitié. » Ces propos permettent un argumentaire empreint d'un autre « réalisme » quant au manque d'objectivité de ce pointage externe au contexte de la classe : l'évaluation pronostique à visée d'orientation proposée à l'élève se fait à proportions qualitatives et quantitatives de traces bien plus faibles que les proportions de la mesure interne. Le nombre et la fréquence des évaluations internes à la classe dépassent de loin ceux de l'évaluation externe : il existe donc une disproportion de mesure d'éléments tangibles entre les évaluations des personnes en charge de la transmission des apprentissages et l'évaluation standardisée. La prise en considération de l'individu semble centrale dans la recherche d'une mise en situation. Ce qui pour d'aucun·es relève du subjectivisme apparaît ici sous le jour d'un objectivisme non naturaliste ou positiviste. Objectivisme en évaluation qui prendrait en compte la valeur de l'individu dans une temporalité et un espace qualitatif et quantitatif le plus large possible. Dans une perspective

d'évaluation située [qui] incite à conceptualiser la relation entre l'individu-en-activité (membres qui évaluent/qui sont évalués) et le contexte de développement de cette activité, tant dans l'idée d'un contexte immédiat (la classe, l'école) que dans une relation à un contexte socioculturel, économique et historique plus large (le système scolaire, ses traditions, son rapport au marché du travail). (Mottier Lopez & Allal, 2008, p. 468)

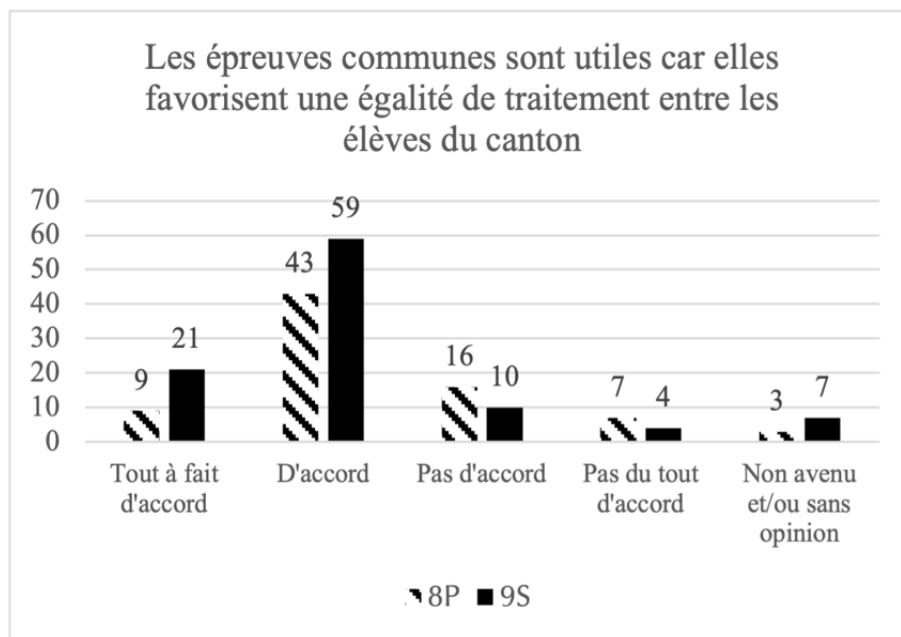
Le nombre et la variété limités des traces entre l'évaluation externe standardisée portent à croire qu'elle se rapporte à un sommet (*summa*) limité par une somme finie (*sigma*). En ce sens l'évaluation externe dans sa démarche sommative manquerait de force scientifique et devrait être sujette à une interprétation limitée. Survient ici, pour le corps enseignant développant un doute, un regard critique au sujet des « repères faisant foi » (Thévenot, 2019).

Ce doute conduit à un dilemme d'objectivité de l'action morale. « An action can be objectively but not subjectively right where, despite the objective rightness of the action, the agent has no beliefs about its rightness or believes falsely that it is not objectively right » (Timmons, 2015). Le système perdure tant que le corps enseignant croit à la possible objectivité morale de l'action décisionnelle issue de l'évaluation standardisée en favorisant une égalité de traitement (figure 3), quand bien même l'action apparaît, dans une perspective individuelle, comme injuste, inégalitaire ou inéquitable¹⁸.

¹⁸ Pour 45% (36, n=81) des enseignant·es qui préparent les élèves au tests interrogés estime que la procédure d'orientation amène à créer de l'inégalité scolaire et sociale.



Figure 3
Épreuves communes et égalité de traitement



Les dilemmes et tensions morales auxquels sont soumis les enseignant·es montrent l'importance d'une vision décentrée qui considérerait les contextes imbriqués (familiaux, sociétaux, culturels...) dans l'évaluation à visée d'orientation. La littérature montre que l'épreuve externe s'appuie souvent sur des situations d'application, de mémorisation, de tâches spécifiques et à des pratiques d'enseignement plus « traditionnelles » (Tessaro, 2015). Le corps enseignant comme la gouvernance entrent dans une structuration complexe.

Il y a quand même ce côté un peu officiel en disant : voilà je quitte le primaire pour aller au secondaire, j'ai quelque chose à montrer. Et on est finalement évalué toute notre vie. Donc peut-être que c'est une première manière, à 11 ans, d'entrer un peu dans ce genre d'évaluation ... On est évalué toute notre vie.

Le glissement sémantique, opéré ici par un enseignant, entre *valeur d'une mesure* et *rituel* n'est pas anodin. La réduction opérée entre ressources et mesures d'apprentissages n'est pas sans incidence sur des considérations éthiques.

5. Conclusion

Selon Chateauraynaud (1991), « l'attitude objectivante » neutralise le jugement moral et interdit à celle ou celui qui l'adopte d'entrer dans une relation interpersonnelle. D'où l'importance de « l'attitude performative », qui octroie le pouvoir au sujet de valider l'action par son énoncé. Accorder le pouvoir aux enseignant·es de tenir un discours performatif sur la décision d'orientation, comme responsabilité morale de l'individu dans les relations interpersonnelles, semble essentiel. La relation interpersonnelle demande à toute personne tierce, ici l'État, le canton, de considérer la légitimité du jugement de faits et de valeurs et la performativité de la parole enseignante, quant à la « juste » orientation compte tenu des risques d'hypertrophie de justice si cette parole est tue. Il semble essentiel de



maintenir vivante l'enquête au sujet de l'objectivité relative des faits tangibles sur lesquels s'appuient les tests externes et standardisés. À titre d'exemple, quand il est souhaité au niveau macro, que les enseignant·es « régulent » leurs pratiques d'évaluation, des espaces institutionnels doivent être accordés pour ajuster les pratiques (Yerly, 2022).

Enfin, les données tendent à montrer en fil rouge de l'analyse des discours que la définition même de la fonction centrale de l'orientation n'est pas rendue explicite. S'agit-il d'un classement en vue d'une division du travail scolaire, du social, d'accès à des diplômes, d'une acquisition juste et équitable des biens éducatifs... sur quelle normativité se construit le jugement scolaire de l'évaluation à visée d'orientation ? Si l'ensemble des acteurs et actrices comme bénéficiaires du système scolaire répondent diversement à cette question, la recherche et les politiques pourraient s'accorder davantage sur des éléments en tensions qui animent les intentions de l'évaluation externe standardisée à visée d'orientation. Ainsi,

Both researchers and politicians have focused heavily on student achievement - understood as performance on standardized tests in mathematics and reading - and its distribution. A focus on these outcomes may come at the cost of other goals such as students' ability to work collaboratively or to appreciate music or art. (Brighouse et al., 2016, p. 4)

Si cette perspective utilitaire peut paraître provocante ou insolite par cette focalisation sur les issues de la scolarité (*outcomes*) dans des habiletés sociales, elle a toutefois le mérite de questionner le non-dit : la diversité des contextes explicites et implicites qui participent de l'évaluation à visée d'orientation et les justices qui l'animent.

Références

- Allal, L. (2002). Acquisition et évaluation des compétences en situation scolaire. In J. Dolz (Ed.), *L'énigme de la compétence en éducation* (pp. 75-94). De Boeck Supérieur. <https://doi.org/10.3917/dbu.dolz.2002.01.0075>
- Batruch, A., Jetten, J., Van de Werfhorst, H., Darnon, C., & Butera, F. (2022). Belief in school meritocracy and the legitimization of social and income inequality. *Social Psychological and Personality Science*, 14(5), 621-635. <https://doi.org/10.1177/19485506221111017>
- Bélaïr, L. (2005). Les dérives de l'obligation de résultats ou l'art de surfer sans planche. In C. Lessard & P. Meirieu (Eds.), *L'obligation de résultats en éducation* (pp. 179-203). De Boeck.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25, 1-25. <https://doi.org/10.1080/0969594X.2018.1441807>
- Boltanski, L., & Thévenot, L. (1991). *De la justification. Les économies de la grandeur*. Gallimard.
- Bourdieu, P., & Passeron, J. C. (1977). *La Reproduction. Éléments pour une théorie du système d'enseignement*. Minuit.
- Bonami, M. (2005). Évaluation interne et évaluation externe : concurrence ou complémentarité ? *Les Cahiers de Recherche en Éducation et Formation*, 38, 1-20. <https://shs.hal.science/halshs-00603486/document>
- Brighouse, H., Ladd, H. F., Loeb, S., & Swift, A. (2016). Educational goods and values: A framework for decision makers. *Theory and Research in Education*, 14(1), 3-25. <https://doi.org/10.1177/1477878515620887>
- Chateauraynaud, F. (1991). L'exercice du jugement moral. In F. Chateauraynaud (Ed.), *La Faute professionnelle* (pp. 436-445). Éditions Métailié. <https://www.cairn.info/la-faute-professionnelle--9782864241034-p-436.htm>
- Chateauraynaud, F. (2004). L'épreuve du tangible : expériences de l'enquête et surgissements de la preuve. In B. Karsanti & L. Quéré (Eds.), *La croyance et l'enquête : aux sources du*



- pragmatisme (pp. 167-194). EHESS, Raisons pratiques. <https://books.openedition.org/editionsehess/11215?lang=fr>
- Coulangeon, P. (2021). *Culture de masse et société de classes : le goût de l'altérité*. Presses universitaires de France.
- Crahay, M. (2009). Articuler l'évaluation en classe et le pilotage des systèmes, est-ce possible ? In L. Mottier Lopez, & M. Crahay (Eds.), *Évaluations en tension. Entre la régulation des apprentissages et le pilotage des systèmes* (pp. 233-251). De Boeck.
- CIIP. (2011). *Plan d'études romand*. PER-MER. <https://portail.ciip.ch/per/>
- Darnon, C., Smeding, A. & Redersdorff, S. (2018). Belief in school meritocracy as an ideological barrier to the promotion of equality. *European Journal of Social Psychology*, 48(4), 523-534. <https://doi.org/10.1002/ejsp.2347>
- De Ketele, J. (2010). Ne pas se tromper d'évaluation. *Revue française de linguistique appliquée*, 15(1), 25-37. <https://doi.org/10.3917/rfla.151.0025>
- De Ketele, J. (2013). Les épreuves externes servent-elles l'action ? *Revue française de linguistique appliquée*, 18(1), 9-27. <https://doi.org/10.3917/rfla.181.0009>
- Demailly, L. (2001). *Évaluer les politiques éducatives : sens, enjeux, pratiques*. De Boeck.
- Demailly, L. (2006). En Europe : l'évaluation contre la crise des systèmes scolaires, l'évaluation en crise. *Éducation et société*, 17, 105-120. <https://doi.org/10.3917/es.017.0105>
- Dupriez, V. (2010). *Séparer pour réussir ? Les modalités de groupement des élèves*. UNESCO, Institut international de planification de l'éducation (IIPÉ).
- Dubet, F., & Duru-Bellat, M. (2010). *Les sociétés et leur école. Emprise du diplôme et cohésion sociale*. Seuil.
- Duru-Bellat, M. (2006). *L'inflation scolaire : les désillusions de la méritocratie*. Seuil.
- Duru-Bellat, M., & Kieffer, A. (2008). Du baccalauréat à l'enseignement supérieur en France : déplacement et recomposition des inégalités. *Population*, 63, 123-157. <https://doi.org/10.3917/popu.801.0123>
- Duru-Bellat, M., & Meuret, D. (2009). *Les sentiments de justice à et sur l'école*. De Boeck.
- Duru-Bellat, M., & Tenret, E. (2009). L'emprise de la méritocratie scolaire : quelle légitimité ? *Revue Française de Sociologie*, 50, 229-258. <https://doi.org/10.3917/rfs.502.0229>
- Duru-Bellat, M., & Van Zanten, A. (2009). *Sociologie du système éducatif : les inégalités scolaires*. PUF.
- Falcon, J., & Bataille, P. (2018). Equalization or reproduction? Long-term trends in the intergenerational transmission of advantages in higher education in France. *European Sociological Review*, 34(4), 335-347. <https://doi.org/10.1093/esr/jcy015>
- Felouzis, G., & Charmillot, S. (2017). Les inégalités scolaires en Suisse. *Social Change in Switzerland*, 8. <https://doi.org/10.22019/SC-2017-00001>
- Felouzis, G., & Hanhart, S. (Eds.) (2011). *Gouverner l'éducation par les nombres ? Usages, débats et controverses*. De Boeck.
- Helgoy, I., & Homme, A. (2007). Towards a new professionalism in school? A comparative study of teacher autonomy in Norway and Sweden. *European educational research journal*, 6(3), 232-249. <https://doi.org/10.2304/eeerj.2007.6.3.232>
- Hulak, F., & Girard, C. (2018). *Philosophie des sciences humaines. Tome 2 : Méthodes et objets* (Vol. 2). Vrin.
- Hofstetter, R. (2012). La Suisse et l'enseignement aux XIXe-XXe siècles. Le prototype d'une « fédération d'États enseignants » ? *Histoire de l'éducation*, 134, 59-80. <https://doi.org/10.4000/histoire-education.2499>



- Jones, B. (2007). The unintended outcomes of high-stakes testing. *Journal of applied school psychology, 23*(2), 65-86. https://doi.org/10.1300/J370v23n02_05
- Kaufmann, J.-C. (2016). *L'entretien compréhensif* (4^e ed.). Armand Colin. (Ouvrage original publié en 1996)
- Kieffer, A., & Tréhin-Lalanne, R. (2011). La fabrication d'un consensus : la révision de la Classification Internationale Type de l'Éducation. *Sociologie et sociétés, 43*(2), 273-299. <https://doi.org/10.7202/1008247ar>
- Lauermann, F., & Karabenick, S. A. (2011). Taking teacher responsibility into account(ability): Explicating its multiple components and theoretical status. *Educational Psychologist, 46*(2), 122-140. <https://doi.org/10.1080/00461520.2011.558818>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- 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 (Eds.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (pp. 85-96). De Boeck. <https://www.cairn.info/l-evaluation-des-competences-en-milieu-scolaire-et--9782804181949-page-85.htm>
- Laveault, D., & Yerly, G. (2018). Modération statistique et modération sociale des résultats scolaires : Approches opposées ou complémentaires ? *Mesure et évaluation en éducation, 40*(2), 91-123. <https://doi.org/10.7202/1043569ar>
- Lemay, V. (2000). *Évaluation scolaire et justice sociale. Droit, éducation et société*. Éditions du Renouveau pédagogique.
- Maroy, C. (2008). Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe ? *Sociologie et société, 40*, 31-55. <https://doi.org/10.7202/019471ar>
- Maroy, C. (2012). Les politiques d'accountability au service de la confiance dans l'institution scolaire et les enseignants ? *Revue suisse des sciences de l'éducation, 34*(1), 57-70. <https://doi.org/10.25656/01:10040>
- Mbiatong, J. (2022). L'entretien compréhensif : conduite, relances, catégorisation des matériaux. In B. Alberto & J. Thievenaz (Eds.), *Traité de méthodologie de la recherche en Sciences de l'Éducation et de la Formation. Tome II* (pp. 204-212). Éditions Raison et Passion.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie, 169*, 99-140. <https://doi.org/10.4000/rfp.1531>
- Mons, N., & Dupriez, V. (2010). Les politiques d'accountability. *Recherche & formation, 65*(3), 45-59. <https://doi.org/10.4000/rechercheformation.131>
- Mottier Lopez, L., & Allal, L. (2008). Le jugement professionnel en évaluation : un acte cognitif et une pratique sociale située. *Revue suisse des sciences de l'éducation, 30*(3), 465-482. <https://doi.org/10.24452/sjer.30.3.4798>
- Mottier Lopez, L., & Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation, 31*(3), 5-34. <https://doi.org/10.7202/1024962ar>
- Mottier Lopez, L. (2009). Introduction. L'évaluation en éducation : des tensions aux controverses. In L. Mottier Lopez (Ed.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (pp. 7-25). De Boeck Supérieur.
- Mottier Lopez, L. (2013). De la mesure à l'évaluation collaborative en éducation. *Revue française d'administration publique, 148*(4), 939-952. <https://doi.org/10.3917/rfap.148.0939>
- Mottier Lopez, L., Girardet, C., & Naji, T. (2021). L'évaluation continue pour apprendre : enjeux de la pluralité des feedbacks entre pairs dans un cours universitaire. *Mesure et évaluation en éducation, 44*(2), 1-34. <https://doi.org/10.7202/1090461ar>



- Ntamakiliro, L., & Daeppen, K. (2016). *Orientation scolaire et satisfaction des élèves avant et depuis la réforme du secondaire I*. URSP, 168. https://www.vd.ch/fileadmin/user_upload/organisation/dfj/ursp/Publications/Publications_99-ajd/168_rapport_ursp_2016.pdf
- OFS. (2021). *Situation sociale et économique des étudiants (SSEE)*. <https://www.bfs.admin.ch/bfs/fr/home/statistiques/education-science/enquetes/ssee.html>
- Pasche Gossin, F., & Donzé, T. (2020). *Évaluation des épreuves communes. Rapport intermédiaire* [Rapport de recherche non publié]. HEP-BEJUNE.
- Pasche Gossin, F., & Donzé, T. (2021). *Évaluation des épreuves communes. État de lieux et tendances générales du positionnement du personnel enseignant de 7-8 P et de 9S dans le canton du Jura* [Rapport de recherche non publié]. HEP-BEJUNE.
- Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology*, 103, 683-700. <https://doi.org/10.1037/a0023911>
- Rayou, P. (2020). Des registres pour apprendre. *Éducation & didactique*, 14(2), 49-64. <https://doi.org/10.4000/educationdidactique.6737>
- Rawls, J. (2009). Unité sociale et biens premiers. *Raisons politiques*, 33(1), 9. <https://doi.org/10.3917/rai.033.0009>
- Romainville, M. (2011). Objectivité versus subjectivité dans l'évaluation des acquis des étudiants. *Revue internationale de pédagogie de l'enseignement supérieur*, 27(2). <https://journals.openedition.org/ripes/499>
- Suchaut, B. (2008). Évaluation externe et recherche empirique en éducation : quels apports pour l'amélioration de l'école primaire? *Recherches & éducations*, 1, 89-104. <https://doi.org/10.4000/rechercheseducations.443>
- Tessaro, W. (2015). Les épreuves externes comme soutien au développement professionnel des enseignants dans leurs pratiques d'évaluation certificative : les apports d'une recherche collaborative. *Évaluer. Journal international de Recherche en Education et Formation*, 1(2), 49-65. <https://journal.admee.org/index.php/ejiref/article/download/52/24/>
- Tessaro, W., & Ntamakiliro, L. (2010). *Les évaluations externes en Suisse romande : enjeux, pratiques et effets*. Actes du congrès de la section suisse de l'ADMEE-Europe, Martigny.
- Thévenot, L. (2019). Ce qui engage : la sociologie des justifications, conventions et engagements, à la rencontre de la norme. *Revue des droits de l'homme*, 16. <https://doi.org/10.4000/revdh.6452>
- Timmons, M. C. (2015). Objective rightness. In R. Audi (Ed.), *The Cambridge Dictionary of Philosophy*. Cambridge University Press.
- Torche, F. (2011). Is a college degree still the great equalizer? Intergenerational mobility across levels of schooling in the United States. *American Journal of Sociology*, 117(3), 763-807. <https://doi.org/10.1086/661904>
- Vaudroz, C., & Berger, J.-L. (2019). Validation de la version française de l'Échelle du sentiment de responsabilité des enseignants (Teacher Responsibility Scale). *Mesure et évaluation en éducation*, 41(3), 87-117. <https://doi.org/10.7202/1065166ar>
- Vial, M. (2012). L'évaluation située. In M. Vial (Ed.), *Se repérer dans les modèles de l'évaluation* (pp. 347-403). De Boeck Supérieur.
- Yerly, G. (2014). *Les effets de l'évaluation externe des acquis des élèves sur les pratiques des enseignant·e·s : analyse du regard des enseignant·e·s du primaire* [thèse de doctorat, Université de Fribourg]. Edudoc.ch. <https://edudoc.ch/record/120691>
- Yerly, G. (2022). La collaboration des enseignants autour de l'évaluation des apprentissages comme moyen de piloter le système ? Le cas de la politique « évaluation commune » dans le secondaire supérieur en Suisse. *Les dossiers des sciences de l'éducation*, 44, 77-97.



- Yerly, G., & Maroy, C. (2017). La gouvernance par les résultats est-elle un mode de régulation de l'école légitime aux yeux des enseignants ? Une enquête qualitative dans 4 systèmes scolaires. *Revue française de pédagogie*, 198(1), 93-108. <https://doi.org/10.4000/rfp.5326>
- Younès, N. (2020). *Vers une évaluation écologique dans l'enseignement supérieur : dispositifs et transformations* [Habilitation à Diriger des Recherches, Université de Lorraine]. HAL archives-ouvertes. <https://hal.science/tel-03032417>
- Wiggins, G. (1989a). A true test: Toward more authentic and equitable assessment. *The Phi Delta Kappan*, 70(9), 703-713. <https://www.jstor.org/stable/20404004>
- Wiggins, G. (1989b). Teaching to the (Authentic) Test. *Educational Leadership*, 46(7), 41-47. https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_198904_wiggins.pdf
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgment practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59-75. <https://doi.org/10.1080/09695940903565610>

Tristan Donzé

Responsable des publications des éditions HEP-BEJUNE (Haute école pédagogique des cantons de Berne, Jura et Neuchâtel), chargé d'enseignement et membre de l'équipe de recherche Apprentissage et évaluation au sein de la HEP-BEJUNE, il a été didacticien de philosophie et est aujourd'hui doctorant à l'Université de Genève.

Françoise Pasche
Gossin

Professeure à la Haute Ecole pédagogique de BEJUNE, elle explore les questions d'apprentissage et d'évaluation dans différents systèmes d'éducation et de formation. Ses travaux s'intéressent aux évolutions et aux problématiques de l'enseignement en général, ainsi qu'aux questions relatives à l'égalité des chances et à la diversité.

