



Zinguinian, M. & André, B. (2021). L'impact d'un instrument d'évaluation sous forme d'échelles descriptives sur les évaluations formatives et certificatives de futurs enseignants du secondaire obligatoire en Suisse. *Formation et pratiques d'enseignement en questions*, 27, 137-155. <https://doi.org/10.26034/vd.fpeq.2021.303>

This article is published under a *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY)*:
<https://creativecommons.org/licenses/by/4.0>

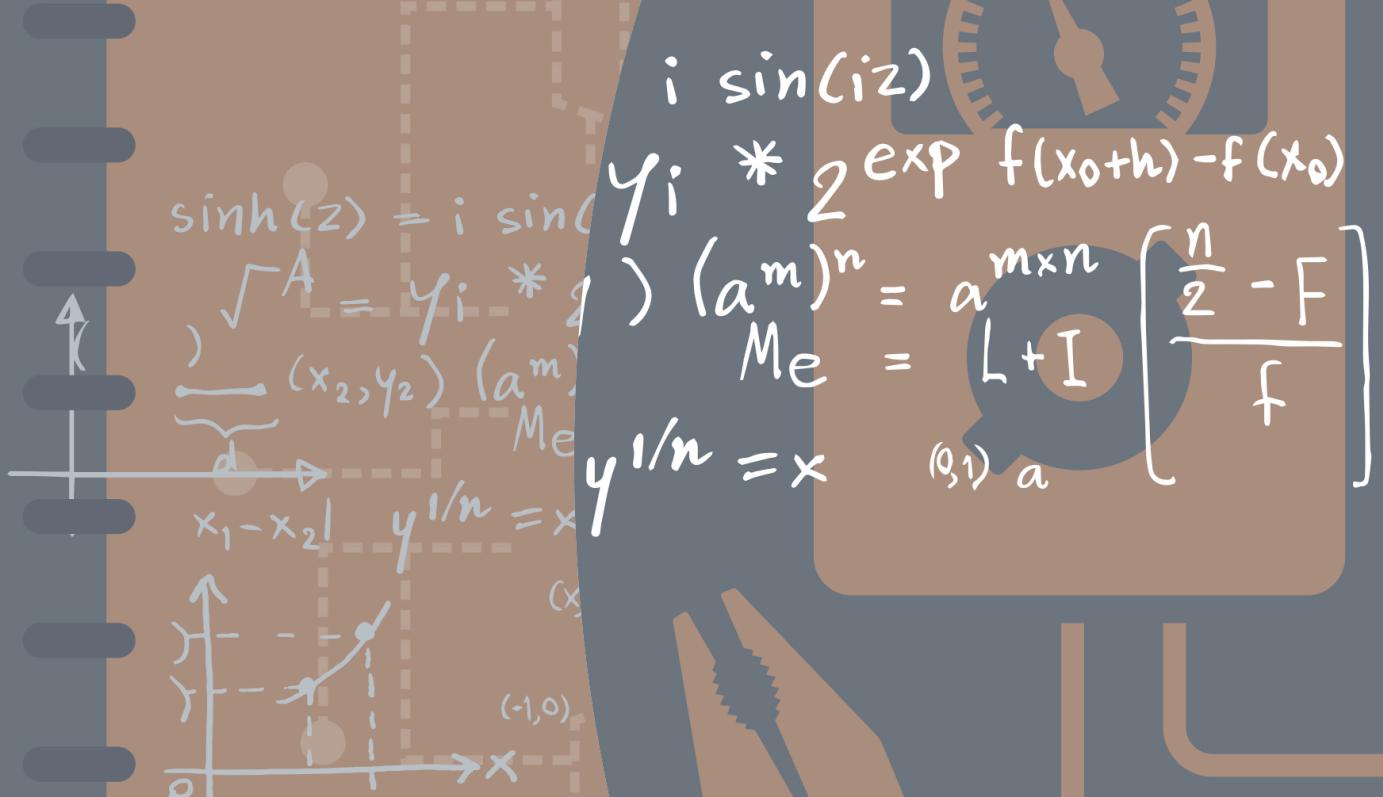


© Méliné Zinguinian, Bernard André, 2021



Revue des **HEP** et institutions assimilées de Suisse romande et du Tessin

Lorsque l'alternance vient soutenir la professionnalisation



Comité de rédaction

Catherine Audrin HEP Vaud
Isabelle Caprani, IFFP
Pierre-François Coen, HEP Fribourg
Stefano Losa, SUPSI
Fabio Di Giacomo, HEP Valais
Deniz Gyger Gaspoz, HEP BEJUNE
Christophe Ronveau, UNIGE/ FPSE
Edmée Runtz-Christan, CERF, Université de Fribourg

Comité scientifique

Bernard Baumberger, HEP Lausanne
Jonathan Bolduc, Université d'Ottawa
Gérard Sensevy, IUFM de Bretagne
Cecilia Borgès, Université de Montréal
Evelyne Charlier, Facultés universitaires Notre Dame de la Paix de Namur
Serge Dégagné, Université Laval
Marc Demeuse, Université de Mons-Hainaut
Jacques Ducommun, HEP BEJUNE
Jean-François Desbiens, Université de Sherbrooke
Hô-A-Sim Jeannine, IUFM de Guyane
Jean-François Marcel, Université de Toulouse II
Lucie Mottier Lopez, Université de Genève
Danièle Périsset Bagnoud, HEP du Valais
Philippe Le Borgne, IUFM de Franche-Comté
Sabine Vanhulle, Université de Genève

Coordinatrices du N° 26

Christophe Gremion / Christophe.Gremion@iffp.swiss
et Méliné Zinguinian / meline.zinguinian@hepl.ch

Rédacteur responsable

Pierre-François Coen / coenp@edufr.ch

Secrétariat scientifique

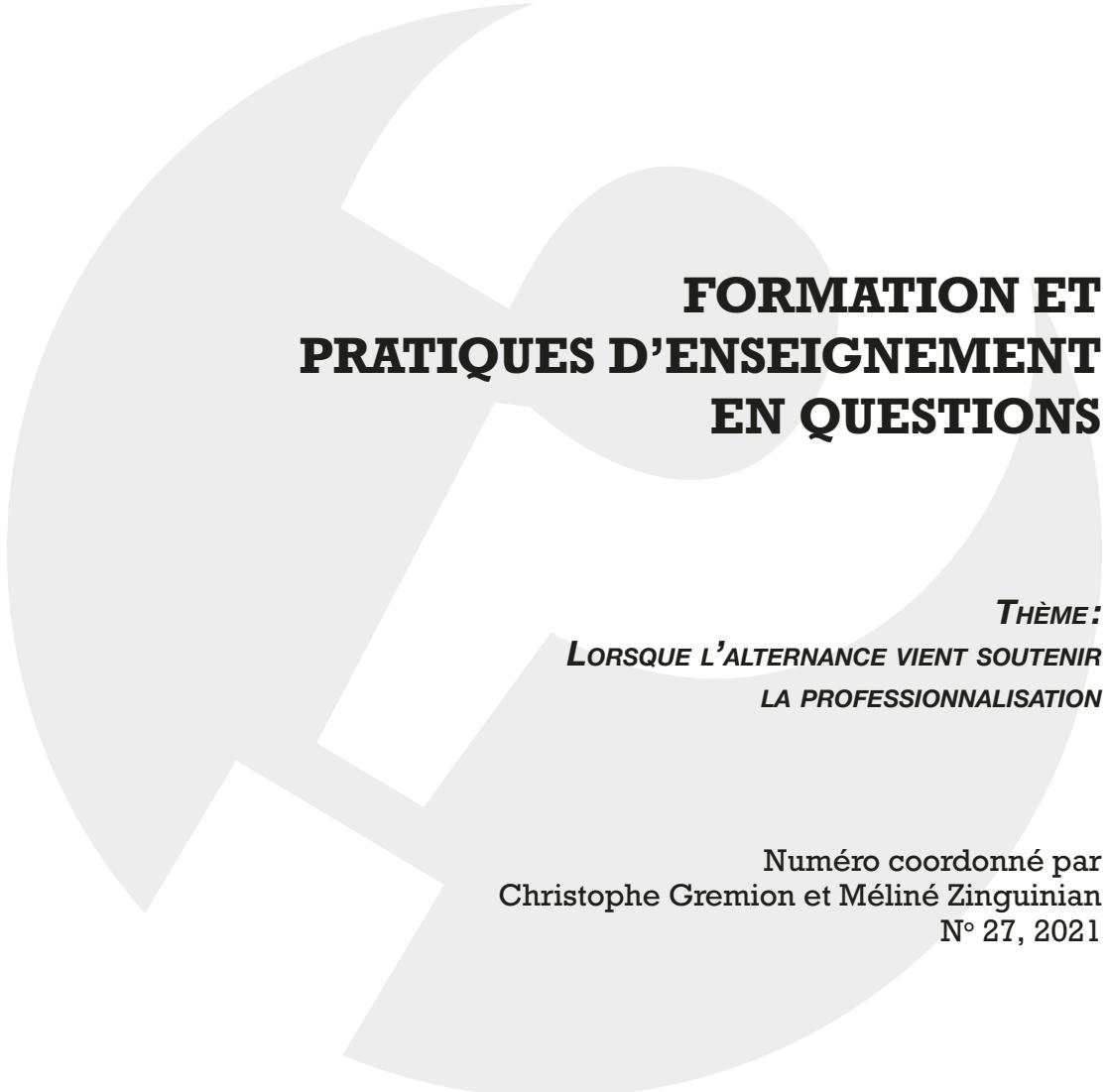
Sarah Boschung / boschungsa@edufr.ch

Secrétariat de la revue

Revue « Formation et pratiques d'enseignement en questions »
Haute école pédagogique de Fribourg
Rue de Morat 36
CH - 1700 Fribourg
www.revuedeshep.ch

Edition

Conseil académique des Hautes écoles romandes en charge de la formation
des enseignant.e.s (CAHR)



FORMATION ET PRATIQUES D'ENSEIGNEMENT EN QUESTIONS

THÈME :

***LORSQUE L'ALTERNANCE VIENT SOUTENIR
LA PROFESSIONNALISATION***

Numéro coordonné par
Christophe Gremion et Méliné Zinguinian
N° 27, 2021

Comité de lecture

René Barioni, HEP Vaud (Suisse)
Francine Chainé, Université Laval (Canada)
Anne Clerc, Haute école pédagogique du canton de Vaud (Suisse)
Marie-Noëlle Cocton, Université Catholique de l'Ouest (France)
Frédéric Darbellay, Université de Genève (Suisse)
Jean-Rémi Lapaire, Université de Bordeaux (France)
Valérie Lussi Borer, Université de Genève (Suisse)
Françoise Masuy, Université de Louvain-La-Neuve (Belgique)
Danielle Périsset, Haute école pédagogique du Valais (Suisse)
Marie Potapushkina-Delfosse, Université Paris-Est Créteil (France)
Sar Savrak, Haute Ecole d'Ingénierie et de Gestion du canton de Vaud (Suisse)
Gabriele Sofia, Université Paul Valéry Montpellier 3 (France)
Stéphane Soulaine, Université de Montpellier (France)
Katja Vanini De Carlo, Université de Genève (Suisse)

Le contenu et la rédaction des articles n'engagent que la responsabilité de leurs auteurs.

© Conseil académique des hautes écoles romandes en charge de la formation des enseignant.e.s (CAHR)

ISSN 1660-9603

Secrétariat scientifique : Sarah Boschung
Rédacteur responsable : Pierre-François Coen
Conception graphique : Jean-Bernard Barras
Mise en page : Marc-Olivier Schatz



Thème : Lorsque l'alternance vient soutenir la professionnalisation

Numéro coordonné par
Christophe Gremion et Méliné Zinguinian

TABLE DES MATIERES

<i>Introduction</i> Christophe Gremion et Méliné Zinguinian	7
<i>L'évaluation des compétences professionnelles sur le terrain : De la formation à la certification, une progressivité à interroger</i> Catherine Tobola Couchepin et Danièle Périsset	15
<i>De la transmission à l'acculturation scientifique dans les approches d'utilisation de connaissances issues de la recherche par des enseignants associés en contexte d'encadrement du stagiaire</i> Josianne Caron, Liliane Portelance et Abdellah Marzouk	35
<i>Présence enseignante lors de stages supervisés à l'aide du numérique : analyse qualitative de pratiques de supervision à distance</i> Matthieu Petit, Julie Babin et Marie-Ève Desrochers	57
<i>Élaboration d'un cadre d'analyse de pratiques de supervision de stage en enseignement</i> Lucie Dionne, Claudia Gagnon et Matthieu Petit	75
<i>La collaboration dans l'organisation des stages en enseignement : rapprochement de points de vue des acteurs universitaires et scolaires</i> Liliane Portelance et Josianne Caron	95
<i>La coévaluation est-elle un espace privilégié pour favoriser les liens entre la théorie et la pratique dans la formation des enseignants ?</i> Olivier Maes, Agnès Deprit, Marc Blondeau, Stéphane Colognesi et Catherine Van Nieuwenhoven	117
<i>L'impact d'un instrument d'évaluation sous forme d'échelles descriptives sur les évaluations formatives et certificatives de futurs enseignants du secondaire obligatoire en Suisse</i> Méliné Zinguinian et Bernard André	137
<i>Transition vers une notation succès/échec pour l'évaluation de stages en enseignement au Québec : regards croisés de divers acteurs impliqués</i> Mylène Leroux, Vincent Boutonnet et Karina Lapointe	157
<i>Qualité des stages en milieu de travail réalisés dans le cadre d'une formation professionnelle : validation d'une échelle de mesure en français auprès d'élèves en alternance travail-études</i> Mylène Beaulieu, Yves Chochard, Annie Dubeau, Camille Jutras-Dupont et Isabelle Plante	175



<i>La formation des enseignants associés au Québec: soutien à leur développement professionnel</i>	195
Glorya Pellerin, Liliane Portelance, Isabelle Vivegnis et Geneviève Boisvert	
<i>Quelle alternance pour quelle professionnalisation ? Etude de deux dispositifs de formation professionnelle</i>	211
Christophe Gremion	
<i>Les récits projectifs et rétroprojectifs comme dispositifs de réflexion collective dans la formation en alternance</i>	227
Marcos Maldonado	
<i>L'entretien d'autoconfrontation comme voie de développement professionnel des étudiants ?</i>	245
Marc Blondeau, Anaïs Laurent et Catherine Van Nieuwenhoven	
<i>Appropriation des savoirs théoriques et postures identitaires dans les textes réflexifs</i>	267
Léa Couturier	
<i>La formation initiale prépare-t-elle à la vie réelle d'un enseignant ? Perspectives de finissants québécois</i>	293
France Dufour, Liliane Portelance, Glorya Pellerin et Geneviève Boisvert	
<i>Évaluation des effets d'une formation professionnelle en alternance sur la qualité de l'emploi de travailleurs québécois de l'industrie manufacturière</i>	311
Yves Chochard, Tetyana Ryabets, Caroline Turcotte, Félix-Antoine Cloutier et Marc-Antoine Vallée Lebouthillier	



**LORSQUE L'ALTERNANCE VIENT SOUTENIR
LA PROFESSIONNALISATION**



L'impact d'un instrument d'évaluation sous forme d'échelles descriptives sur les évaluations formatives et certificatives de futurs enseignants du secondaire obligatoire en Suisse

Méliné ZINGUINIAN¹ (Haute école pédagogique du canton de Vaud, Suisse) et **Bernard ANDRÉ²** (Haute école pédagogique du canton de Vaud, Suisse)

En retraçant les étapes et les résultats de la mise en oeuvre d'un instrument d'évaluation des compétences des étudiants en stage pour l'enseignement au secondaire obligatoire en Suisse, cet article met en lumière les enjeux de l'évaluation. L'analyse porte principalement sur les notes mises dans les bilans de stage et sur les retours des évaluateurs en vue d'améliorer l'instrument. Les principaux résultats montrent que le nouvel instrument d'évaluation sous forme d'échelles descriptives remplit son rôle de balise sans pour autant amener une baisse massive des notes et une augmentation significative du nombre d'insuffisances comme les évaluateurs le présageaient. La multiplicité des acteurs concernés par le résultat d'une évaluation certificative permet de comprendre que cet instrument semble influencer davantage le suivi formatif des étudiants en stage, qui est finalement au coeur de leur professionnalisation durant leurs études.

Mots-clés: Evaluation, échelles descriptives, formation pratique, compétences, enseignement

Introduction

La formation à l'enseignement dans le canton de Vaud s'effectue en alternance entre des cours donnés à la Haute école pédagogique (HEP) et des stages réalisés dans des établissements scolaires sous la supervision d'un enseignant³ mandaté et formé par cette institution pour cette fonction. Ces enseignants-formateurs de terrain sont volontaires et bénéficient en principe d'au moins dix ans d'expérience de l'enseignement. Outre la socialisation aux gestes du métier d'enseignant, leur mandat inclut l'évaluation des compétences des étudiants en stage. Pour remplir cette deuxième partie de leur mandat, ils observent l'étudiant enseigner et comparent leurs observations aux attentes qui figurent dans les documents d'évaluation qui leur sont fournis par la HEP.

1. Contact: meline.zinguian@hepl.ch

2. Contact: bernard.andre@hepl.ch

3. Pour des raisons de lisibilité, le masculin généraliste est utilisé dans cet article.



Les formateurs de terrain évaluent les compétences des étudiants en stage, dans une visée formative d'abord puis certificative au terme du stage. Ils sont également amenés à formuler des prédictions quant au potentiel des étudiants à progresser au-delà du stage (Zinguinian, 2019). Bien qu'elles se chevauchent parfois avec des enjeux de certification traités au moment de l'évaluation formative ou des pronostics effectués dans le cadre de l'évaluation certificative (Zinguinian, 2019), ces trois dimensions de l'évaluation (formative, certificative, pronostique) n'ont ni les mêmes enjeux, ni les mêmes visées, ni les mêmes destinataires (voir Tableau 1) :

Tableau 1 : Distinctions entre l'évaluation formative, certificative et pronostique

	Formative	Certificative/sommative	Pronostique
Finalités	Régulation des apprentissages	Sanction positive ou négative d'une prestation ou d'un ensemble de prestations sur la base de toutes les informations recueillies y compris pour les évaluations formatives	Appui à la prise de décision en situation d'orientation ou de sélection
Temporalité	En cours de session d'apprentissage	Au terme d'une session d'apprentissage	Avant l'entrée dans un nouveau cursus de formation ou dans le métier
Relations	Coopération entre l'évaluateur et l'évalué	Intérêts divergents : Attestation d'un niveau de maîtrise versus obtention de son diplôme. Tension relationnelle pour l'évaluateur qui certifie pour la société celui qu'il a formé et avec qui il a construit une relation (Laveault, 2008; Lehmann, 2009)	-

Source : (Perrenoud, 2001a, 2001b)

Dans une évaluation formative, le résultat de l'appréciation reste le plus souvent entre l'évaluateur et l'évalué alors que lorsqu'il s'agit d'une évaluation certificative, les destinataires se multiplient: évalué, acteurs de l'institution qui délivre le diplôme, recruteurs, etc. De cette multiplicité d'acteurs apparaissent les conséquences importantes d'une telle évaluation pour l'évalué: prolongation d'études voire refus d'accès à la profession en cas d'échec définitif par exemple. La multiplicité des acteurs concernés par les résultats de l'évaluation participe à en accroître les enjeux, car l'évaluation ne prend son plein effet qu'une fois communiquée à l'évalué. La dimension sociale d'une évaluation entendue comme une interaction, une négociation (Weiss, 1991) et comme un acte de communication (Mottier Lopez & Allal, 2008), en impacte les résultats.

C'est ce dont nous discuterons dans cet article : l'évaluation certificative a plusieurs destinataires, ce qui la fait sortir du cadre de la relation évaluateur-évalué, avec un ensemble de conséquences sur les décisions prises et le résultat des évaluations posées par les formateurs de terrain.

Jusqu'en 2017, la HEP utilisait, pour cadrer l'évaluation sur les lieux de stage, un référentiel de onze compétences, chacune déclinée en plusieurs niveaux de maîtrise à atteindre en fin de stage. Ces niveaux de maîtrise prenaient la forme d'énoncés, par exemple : «Manifester une compréhension critique



de son cheminement culturel et en apprécier les potentialités et les limites» (HEP Vaud, 2013, p. 5). L'évaluateur était ensuite invité à attribuer à chaque compétence ainsi déclinée une note à partir d'une échelle uniforme allant de A à F, le A représentant un *excellent niveau de maîtrise*, le B un *très bon niveau de maîtrise*, le C un *bon niveau de maîtrise*, le D un *niveau de maîtrise suffisant*, le E un *niveau de maîtrise passable* et le F une *maîtrise insuffisante*. Il revenait donc aux évaluateurs de traduire les niveaux de maîtrise en gestes et attitudes observables et de définir ce qui distinguait un *bon*, d'un *très bon niveau de maîtrise* par exemple.

Même s'ils avaient l'avantage, pour un formateur de terrain rencontré, d'offrir de la finesse et de la précision dans l'évaluation, ces documents présentaient, pour d'autres, plusieurs difficultés notamment liées à la compréhension des compétences figurant dans le référentiel, à leur traduction en gestes et attitudes observables et à la définition du seuil entre les notes (André, 2015). Le processus faisait peser une charge parfois lourde sur les épaules des évaluateurs, d'autant plus lorsqu'il s'agissait de sanctionner des insuffisances (Zinguinian & André, 2015) et ceci plus particulièrement pour les formateurs de terrain au début de leur fonction (Zinguinian, 2019). En découlaient, parfois, des tensions dans les jurys d'évaluation et des difficultés pour l'étudiant à s'autoévaluer, la référence pouvant fluctuer considérablement d'un évaluateur à l'autre. Une autre conséquence a parfois été observée : l'attribution d'une note *passable* à des stages qui auraient dû, selon les propos des évaluateurs, être sanctionnés par un échec. En effet, certaines conditions pouvaient amener le formateur à ne pas arriver à assumer les conséquences d'une telle évaluation pour l'étudiant (prolongation d'études, refus d'accès à l'emploi, etc.) comme pour lui (déclenchement d'une procédure jugée lourde, visibilisation de son travail au risque d'y perdre la face, etc.) (André & Zinguinian, 2016 ; Zinguinian, 2019 ; Zinguinian & André, 2016).

C'est pourquoi, en 2015, la direction de la HEP a mis sur pied un groupe de travail dont la tâche consistait à repenser le référentiel de l'évaluation de manière à ce qu'il soutienne le jugement professionnel des évaluateurs, leur offre l'appui du collectif dans les prises de décision relatives à l'attribution des notes de stage et facilite le travail de mise en relation des observations avec les attentes décrites dans le document d'évaluation. C'est ainsi qu'en 2017, un nouvel instrument sous forme d'échelles descriptives a pu être mis en œuvre. Une échelle descriptive se distingue d'une échelle uniforme, car elle décrit ce qui est attendu pour l'obtention de chaque note (Scallon, 2004). Après un premier article qui a montré le travail de conception de cet instrument et les débats qui l'ont traversé (Zinguinian & André, 2018), ce texte présente sa mise en œuvre : les débats qui ont eu lieu dans le cadre de la formation à son utilisation, jusqu'à l'analyse des évaluations qui ont été produites grâce à lui.

La question qui a présidé à cette recherche était la suivante : dans quelle mesure les échelles descriptives sont-elles effectivement un soutien pour les évaluateurs dans leurs prises de décision relatives à l'évaluation des compétences des stagiaires ? L'hypothèse à la base de tout le travail d'élaboration



des échelles descriptives était que la formulation de balises plus précises permettrait de réduire l'ambiguité dans l'attribution de notes à partir d'une échelle uniforme et de faciliter le rendre compte des résultats, spécialement lors de l'attribution d'une note insuffisante.

Nous observions, avec l'instrument d'évaluation précédent, une distribution globale des notes qui se polarisait vers le haut de l'échelle (environ 75% des notes certifiaient d'un très bon voire d'un excellent niveau de maîtrise entre 2009 et 2012). Nous faisions deux hypothèses. Primo, cette distribution était en partie le résultat des difficultés rencontrées par les formateurs de terrain à utiliser l'instrument d'évaluation et à porter seuls la définition du seuil entre les notes. Secundo, cette dispersion se trouverait, par voie de conséquence, modifiée par l'introduction d'un nouvel instrument d'évaluation offrant des balises plus précises.

S'il ne s'agit pas ici de dénier quoi que ce soit aux compétences des stagiaires ni de rechercher une distribution plutôt qu'une autre, étudier la dispersion des notes a finalement permis d'explorer les enjeux de l'évaluation certificative et de se questionner sur le rôle des notes en contexte de stage. Nous montrerons que si les échelles descriptives offrent effectivement des balises plus précises, il semble qu'elles influencent davantage la formation des étudiants que les évaluations certificatives faites. Nous interprétons ce résultat comme étant la conséquence de la multiplicité des destinataires de l'évaluation certificative, ce qui implique de facto des enjeux importants, différents, voire contradictoires, qui contraignent fortement les résultats des évaluations.

Après une présentation de l'instrument et des étapes de sa mise en œuvre, nous montrerons, par l'analyse des débats qui ont eu lieu durant la formation à son utilisation et par celle des notes attribuées durant les deux premières années de son existence, que le nouvel instrument d'évaluation adopté ne peut, à lui seul en tout cas, permettre le dépassement des tensions liées à une évaluation qui a plusieurs destinataires.

Méthode et corpus

Pour appréhender l'utilisation qui est faite des échelles descriptives par les formateurs de terrain et l'impact de leur introduction sur les résultats des évaluations, nous avons recouru à des données qualitatives et quantitatives. Les données qualitatives ont été essentiellement construites à partir des questions et commentaires formulés par les participants à la formation destinée aux utilisateurs des échelles descriptives entre 2017 et 2019. Nous les avons pris en note durant la formation et avons ensuite analysé leur contenu (Bardin, 2007) en vue de les catégoriser. Plusieurs lectures globales des données ont permis leur regroupement en six catégories: conception de la progression, place accordée aux différents gestes du métier, rôle des notes, validité de l'instrument, protection de l'évalué, clarté du document. Cette formation a été dispensée à des formateurs de terrain de l'enseignement primaire et secondaire en même temps qu'à des formateurs de la HEP effectuant des visites ponctuelles de stage. Nous traiterons, dans cet article, les



formateurs de manière homogène, c'est-à-dire sans procéder à des comparaisons qui mettraient au jour d'éventuelles différences entre eux selon leur profil. Une telle analyse, bien qu'intéressante, nous éloignerait de notre question de recherche.

Pour élargir la partie qualitative de notre corpus, nous avons également utilisé les appréciations et commentaires rédigés dans un questionnaire envoyé aux formateurs de terrain et aux formateurs de la HEP qui effectuent des visites après un semestre d'utilisation des échelles descriptives. Ces données sont venues compléter la catégorisation déjà évoquée. Analyser le contenu des remarques des utilisateurs a permis de saisir le sens qu'ils donnent à l'instrument d'évaluation, à leurs pratiques et aux situations qu'ils rencontrent lors de l'évaluation (Quivy & Van Campenhoudt, 1988). Ces données ont servi l'interprétation des données quantitatives.

Afin de comprendre comment l'instrument d'évaluation peut influencer les notes attribuées, nous avons également utilisé des données quantitatives. Elles sont constituées des notes mises dans 312 bilans de stage rédigés à l'aide des échelles descriptives. 123 formateurs de terrain ont rempli les 147 bilans analysés pour la première année de mise en œuvre et 141 formateurs ont rempli les 165 bilans pour la seconde année. Nous nous sommes penchés uniquement sur les bilans de stage des étudiants du secondaire obligatoire, car ce sont les seuls bilans à notre disposition qui concernent plusieurs promotions d'étudiants. Nous avons également analysé la dispersion globale des notes mises dans 74 bilans de stage rédigés pour une promotion d'étudiants du secondaire obligatoire à partir de l'instrument d'évaluation précédent. Nous avons procédé, pour cette partie de la recherche, à des analyses de fréquences et des tests de chi-carré, bien que ces analyses ne tiennent pas compte du fait que les mêmes formateurs peuvent figurer dans les deux corpus de données (1^{re} et 2^e année) et plusieurs fois dans chacun d'eux lorsqu'ils accompagnent deux étudiants simultanément. Ceci pourrait influencer les résultats. Ces tests ont permis de repérer l'existence ou l'absence de liens entre les notes attribuées et l'instrument d'évaluation utilisé (N'Da, 2015). Ces données nous ont surtout permis de tester l'hypothèse, construite à partir des commentaires des évaluateurs, selon laquelle les notes allaient baisser de manière significative avec l'introduction des échelles descriptives.

Nous avons également utilisé des statistiques donnant à voir l'évolution depuis 2012 du nombre de signalements. Par signalement, nous entendons l'annonce faite à la HEP d'une suspicion d'échec au stage ou de difficultés significatives, par un formateur de terrain ou par un formateur de la HEP ayant effectué une visite de stage. Un signalement intervient généralement avant la possible prononciation d'une insuffisance dans un bilan et déclenche l'organisation de visites de stage par des formateurs de la HEP, si elles n'ont pas déjà eu lieu, et, éventuellement, la mise sur pied d'une conférence intermédiaire qui réunit l'étudiant, le ou les formateurs de terrain du stage en cours, les formateurs de la HEP concernés et le responsable du service chargé de l'organisation des stages à la HEP. Cette rencontre, comme les



premières visites de stage qui font suite à un signalement, est à visée formative et diagnostique et permet de faire le point sur la situation et de discuter les moyens à la disposition de l'étudiant pour atteindre les objectifs de la formation d'ici l'évaluation certificative. Analyser l'évolution des signalements avant et après l'introduction des échelles descriptives a permis de constater un possible impact de l'instrument non pas sur la dispersion des notes, mais sur le suivi formatif proposé aux étudiants.

De l'élaboration des échelles descriptives à leur mise en œuvre

Le groupe de travail, composé de formateurs de la HEP et de formateurs de terrain, a conçu dix échelles descriptives pour l'évaluation des compétences des étudiants qui se forment au métier d'enseignant. Elles prennent la forme d'un tableau (voir Figure 1).

Échelle 1	L'étudiant-e maîtrise les contenus enseignés.			
	Compétence 1 : Agir en tant que professionnel critique et porteur de connaissances et de culture			
	F	E	C	A
Année 3		<p>L'étudiant-e maîtrise les connaissances enseignées qui sont en lien avec le plan d'études (PER).</p> <p>ET</p> <p>La maîtrise des connaissances rend possible l'analyse et l'anticipation des difficultés liées aux contenus enseignés.</p>	<p>La maîtrise des connaissances rend possible la mise en œuvre d'une progression des apprentissages des élèves pertinente.</p>	<p>La maîtrise des connaissances permet à l'étudiant-e de susciter chez les élèves des liens significatifs entre les divers contenus.</p>

Source : Échelle descriptive de fin de formation pour la formation à l'enseignement primaire (HEP Vaud, 2019)

Figure 1 : Échelle descriptive n° 1

L'objectif d'un tel instrument est double :

1. Servir de guide pour le choix des notes alors que les instruments d'évaluation, selon Tillema (2009), ne le sont pas toujours.
2. Permettre des évaluations moins dépendantes de la subjectivité de chaque évaluateur, ce qui passerait par la triangulation entre plusieurs sources d'informations (Merle, 2007) et par le recours à des pratiques d'évaluation collectives (Chauvigné, 2016), voire à des calibrations entre évaluateurs (Chung Wei & Pecheone, 2010). Dans tous les cas, il s'agit d'avoir une définition collective du niveau de maîtrise attendu (Mayen, 2005, cité dans Tourmen, 2015).

Le travail d'élaboration de cet instrument a été l'occasion de débats autour des compétences attendues d'un enseignant en début de carrière, des attentes acceptables pour des étudiants aux différents stades de leur formation et des caractéristiques d'un travail de conception d'un instrument d'évaluation «bien fait» (Zinguinian & André, 2018). Contrairement à ce que les chercheurs anticipaient, la difficulté pour atteindre un consensus n'a pas



eu lieu lors des débats sur les gestes et attitudes d'un «bon» enseignant, mais a porté bien davantage sur trois points: le niveau d'exigence à avoir pour évaluer les compétences d'un enseignant en formation; la place à accorder aux gestes relatifs à l'instruction des élèves par rapport aux gestes favorisant leur socialisation; et enfin la manière de concevoir la progression, à savoir les critères permettant l'obtention de la note minimale pour réussir le stage et ceux permettant l'obtention d'une note plus élevée (Zinguinian & André, 2018).

Le groupe de travail a formalisé une première version des échelles descriptives qui instituait deux changements notables par rapport aux pratiques d'évaluation qui avaient cours chez certains formateurs de terrain en tout cas:

1. Une démarche d'évaluation désormais ascendante là où l'analyse des commentaires mis en regard des notes attribuées à deux promotions d'étudiants entre 2009 et 2012 montrait une démarche d'évaluation plutôt descendante. Nous entendons par là que le formateur commence par comparer les performances observées aux critères décrits pour l'obtention de la note minimale de réussite. Si la performance satisfait ces critères, l'évaluateur regarde si elle satisfait également les attentes pour l'obtention de la note supérieure. Si la réponse est à nouveau affirmative, il s'assure que la performance remplit encore les critères d'obtention de la note suivante jusqu'à l'obtention de la note maximale. La logique descendante correspond à l'idée que le stage est réussi avec la note maximale jusqu'à preuve du contraire. Plus l'évaluateur observe de lacunes et plus la note diminue.
2. La non-compensation des critères inscrite dans l'ergonomie même de l'instrument. Auparavant, certains formateurs de terrain éprouvaient une difficulté à mettre en échec un étudiant qui satisfaisait une majorité des critères mentionnés pour l'obtention de la note minimale. Avec les échelles descriptives, la logique même de la démarche ascendante vise à empêcher la compensation des critères.

Cette première version des échelles a été mise en consultation à plusieurs reprises auprès des formateurs de terrain, des formateurs de la HEP, des étudiants, de la direction de l'institution et a fait l'objet de communications dans plusieurs colloques. Ce qui a permis sa révision sous la forme d'une deuxième version mise en œuvre plus largement. La procédure d'ajustement de l'instrument se fait de manière itérative depuis 2017 et vraisemblablement jusqu'en 2022, date à laquelle des échelles adaptées à chaque programme de formation auront été élaborées, mises en consultation et introduites progressivement (voir Figure 2):

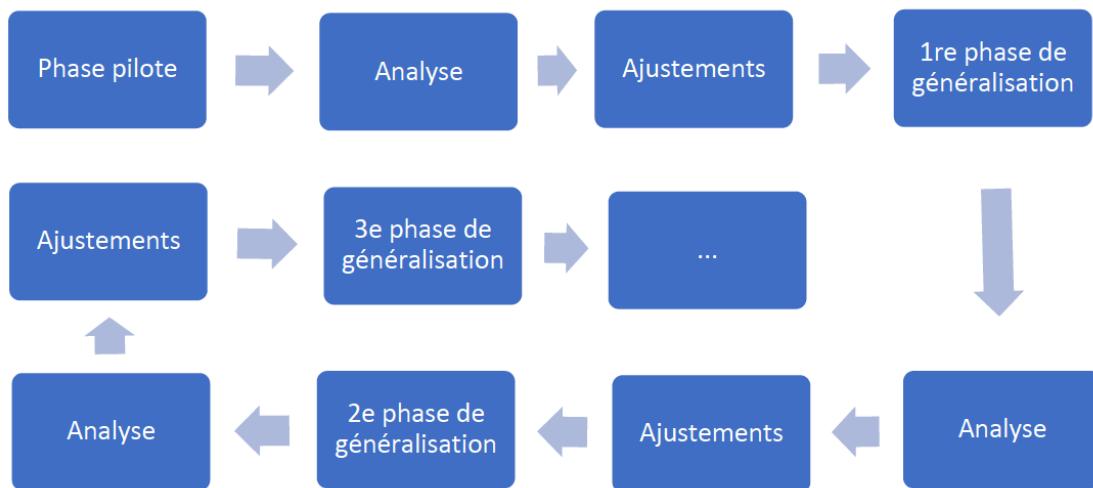


Figure 2: Étapes de la mise en œuvre des échelles descriptives à la HEP Vaud

À chaque mise en œuvre auprès d'une nouvelle population, les évaluateurs concernés sont formés à l'utilisation de ce nouvel instrument. Cette formation est aussi le lieu pour eux d'échanger sur leurs pratiques. Elle a touché, entre 2017 et 2019, un millier de formateurs (de terrain et de la HEP) et a permis aussi le recueil de précieuses informations en vue d'ajuster l'instrument.

Dans le même but d'ajustement, un questionnaire a été systématiquement envoyé aux évaluateurs après un semestre d'utilisation et les notes attribuées ainsi que les commentaires rédigés dans les bilans de stage à l'aide de cet instrument d'évaluation ont été analysés de manière à repérer les incompréhensions par rapport aux attentes décrites et les difficultés dans l'utilisation de ces échelles descriptives.

Résultats et discussion

De l'analyse des données, nous retenons cinq résultats qui seront décrits et discutés dans les parties suivantes.

Les sujets mis en débat par les utilisateurs des échelles descriptives

Les enjeux de la certification apparaissent dans les questions et remarques formulées durant la formation à l'utilisation des échelles descriptives, comme dans les commentaires rédigés dans le questionnaire envoyé après un semestre d'utilisation de cet instrument. Ces questions et commentaires se regroupent en six catégories (voir Tableau 2):



Tableau 2 : Fréquence d'apparition de chacune des catégories de commentaires*

Catégorie	Fréquence en formation	Fréquence dans le questionnaire	Exemple**
1. Conception de la progression	17	8	<i>Échelle 4, pas assez exigeante en fin de formation</i>
2. Place accordée aux différents gestes du métier	15	3	<i>Pas assez de place dans les échelles pour la gestion de classe</i>
3. Rôle des notes	3	3	<i>Pourquoi mettre des notes et ne pas utiliser ces échelles à des fins formatives</i>
4. Validité de l'instrument	17	27	<i>Échelles réductrices, trop épurées, enfermantes</i>
5. Protection de l'évalué	21	13	<i>Importance d'informer toutes les directions du canton pour éviter les risques à l'embauche</i>
6. Clarté du document	50	10	<i>Les commentaires ne vont pas assez à l'essentiel et il est difficile d'en saisir le cœur</i>

* Les sujets de débats qui apparaissaient dans plusieurs journées de formation à destination du même public (formateurs de terrain pour l'enseignement primaire versus secondaire obligatoire versus secondaire post-obligatoire) n'apparaissent qu'une seule fois dans le tableau. Le rôle des notes est ainsi largement sous-estimé par rapport à la place qu'y occupe le débat sur la pertinence de mettre une note globale au stage.

** Seuls les commentaires invitant à des ajustements ont été pris en notes.

Nous observons, dans ce tableau, que les remarques et questions de compréhension et d'amélioration du document, que ce soit en termes de clarté de ce qui est écrit, de présence des différents gestes du métier et de conception de la progression, ont occupé une large place lors de la formation alors que dans le questionnaire envoyé quelques mois plus tard, les commentaires portaient davantage sur la validité du document.

En revanche, la place de la protection de l'évalué du pouvoir qu'a l'évaluateur, ou l'institution qui délivre les diplômes, a occupé la deuxième place dans les débats quel que soit le moment et la manière de recueillir l'information (de vive voix en collectif ou anonymement et individuellement). Si nous voyons déjà les enjeux de la certification dans les questions qui renvoient à l'importance d'un instrument d'évaluation valide ainsi que dans la conception de la progression qui peut induire des notes basses, l'essentiel de l'enjeu réside, selon nous, dans le rapport de pouvoir constitutif d'une évaluation (Berrebi-Hoffmann, 2010; Zarka, 2009) et des conséquences que les décisions prises peuvent avoir sur l'évalué comme sur l'évaluateur.

Le pouvoir de l'évaluateur et de l'institution apparaît lorsqu'il est question du niveau d'exigence de l'instrument d'évaluation et comme, déjà évoqué, de sa validité et de la conception de la progression. Ces sujets de débats, comme la plupart de ceux qui ont eu lieu lors de la formation à l'utilisation des échelles, étaient déjà apparus lors de la conception de l'instrument et lors des premières consultations⁴. Nous constatons donc une transversalité des points

4. Parmi les points de débats qui ne se sont pas retrouvés chez les concepteurs et semblent donc propres aux utilisateurs figure celui sur le rôle des notes et la pertinence d'en mettre. Nous pouvons comprendre cela, car les règles de fonctionnement du groupe de travail des concepteurs interdisaient la remise en question du mandat, sur lequel ils n'avaient pas prise et dans lequel figurait l'utilisation de notes. La seule exception qui a été faite lors de l'élaboration des échelles, à cette demande de conserver une échelle de note en six positions, touche l'évaluation du respect de l'éthique professionnelle et de la déontologie qui se fait à l'aide d'une échelle en deux positions «réussi/échoué» (Zinguinian et André, 2018).



d'attention peut-être liée au fait que les modifications du dispositif sont telles, sur certains aspects, qu'une communication importante sur les choix opérés est nécessaire auprès de chaque nouvel acteur pour espérer le rallier.

Le niveau d'exigence des échelles descriptives

Les échelles descriptives proposées sont globalement considérées comme exigeantes. Certains évaluateurs estiment que les notes les plus élevées de l'échelle ne pourraient même pas être atteintes par des enseignants expérimentés ou que la note minimale requiert parfois la maîtrise de gestes experts, comme la prise en compte de la diversité de la classe et des besoins particuliers des élèves. Lors de la formation, il s'est agi de clarifier ce qui était attendu en termes de prise en compte de la diversité pour prétendre à cette note, à savoir d'adapter le matériel et d'anticiper le soutien spécifique à apporter aux différents élèves et non l'élaboration de plusieurs scénarios pédagogiques par exemple.

La conception de la progression

La conception de la progression d'une note à l'autre qui a été finalement retenue après de longs débats parmi les concepteurs, à savoir la progression par l'adjonction de gestes professionnels plus complexes et non sur le mode «complexification», a aussi fait l'objet de débats parmi les utilisateurs. Dans une progression par complexification, ce qui décrit les différentes notes d'une même échelle utilise le même registre de gestes avec pour conséquence qu'un étudiant qui ne satisfait pas les critères d'obtention d'une note suffisante ne pourra pas maîtriser les gestes relatifs à une note supérieure. Alors que lorsque la progression est pensée comme une «adjonction», les gestes décrits pour chaque note sont autonomes. Il est donc possible qu'un étudiant satisfasse aux critères minimaux et à ceux permettant l'obtention de la note maximale de l'échelle sans parvenir à satisfaire ceux de la note intermédiaire. Dans ce cas, l'étudiant obtient la note minimale, si l'on respecte la logique ascendante à la base des échelles descriptives proposées. Là encore, il a été nécessaire de réexposer les arguments ayant mené à ce choix ainsi que le fait que l'utilisation régulière des échelles descriptives durant le stage permettrait de pallier le risque de voir un étudiant obtenir la note minimale pour la raison évoquée ci-dessus, car cela permet d'orienter, à certains moments, les efforts de l'étudiant et du formateur vers l'acquisition des gestes nécessaires à la satisfaction des critères de la note intermédiaire.

Il en a été de même avec l'option prise de penser l'évaluation à mi-stage comme un point de situation par rapport aux attentes de fin de stage. Si une telle approche n'a pas fait débat pour les stages qui durent un semestre, il en a été tout autrement pour ceux qui durent une année. Pour les concepteurs, procéder ainsi permet à l'étudiant de se projeter dans la suite de son stage en prenant confiance tout en sachant sur quels aspects du métier porter plus particulièrement son attention. Pour les évaluateurs, mettre, en début de formation, des notes faibles, voire insuffisantes, fait courir le risque de décourager les étudiants alors même que le temps restant avant la certification et leur investissement leur permettront sans aucun doute d'atteindre de meilleures notes.



La validité de l'instrument d'évaluation proposé

L'importance de disposer d'un instrument d'évaluation avec une formulation de critères clairs afin de limiter les interprétations possibles a été soulevée par les concepteurs comme par les utilisateurs. Le débat a aussi porté sur le niveau de précision adapté pour décrire chaque note : assez précis pour limiter les interprétations, pas trop précis pour ne pas risquer de voir l'instrument devenir inutilisable parce que ne permettant plus à l'évaluateur de tenir compte du contexte du stage. Par exemple, certains utilisateurs auraient apprécié que l'institution fixe le nombre d'erreurs orthographiques justifiant l'insuffisance à l'échelle relative à la maîtrise de la langue d'enseignement, ce qui n'a pas été fait pour la raison que nous venons d'évoquer.

Outre l'élaboration d'un instrument d'évaluation clair, effectuer un travail de conception bien fait afin d'aboutir à un instrument d'évaluation de qualité consistait encore, pour les concepteurs, à s'assurer que l'étudiant ne soit pas sanctionné plusieurs fois pour la même insuffisance, parce que les mêmes critères réapparaîtraient dans plusieurs échelles. Il importait pour eux également qu'il soit possible pour l'évaluateur de prendre en compte les facteurs contextuels d'exercice de la profession, que les échelles soient fondées sur des apports tant scientifiques qu'expérientiels, qu'elles soient en cohérence avec les réglementations qui cadrent la profession et qu'elles soient d'un niveau d'exigence acceptable (Zinguinian & André, 2018).

Parmi les évaluateurs, le souci d'un travail bien fait est également très présent au travers de la question de la validité de l'instrument et de la recherche d'une compréhension partagée des attentes de l'institution qui délivre les titres. Il est aussi largement présent dans les interventions qui visent à protéger l'évalué du pouvoir de l'évaluateur. D'un souci de ne pas accabler l'étudiant en le sanctionnant à de multiples reprises, l'objet du débat s'est déplacé sur la possibilité pour les étudiants d'obtenir des notes élevées et surtout de ne pas être désavantagés sur le marché de l'emploi sachant qu'ils vont s'y retrouver en concurrence avec des personnes ayant peut-être obtenu leur titre plusieurs années auparavant à l'aide d'un instrument d'évaluation apparemment moins sélectif. Le spectre des acteurs concernés par les résultats de l'évaluation, et donc susceptibles d'influer sur les décisions, s'élargit pour inclure, en plus de l'évalué, de l'évaluateur, des élèves et de l'institution qui délivre les titres, le recruteur.

Protéger l'évalué : rapports de pouvoir et recrutement

Dans le canton de Vaud, le directeur d'établissement joue un rôle important dans la sélection des enseignants en participant à leur recrutement, puis, en tout cas durant la première année d'engagement, en évaluant leurs compétences. Le formateur de terrain se retrouve ainsi à l'interface du stagiaire et du recruteur. D'autant plus que, comme le montrent Losego et Girinshuti (2019) être recommandé par un formateur de terrain facilite l'insertion des enseignants novices de même que, selon Broyon *et al.* (2013), la réussite d'un stage, surtout en emploi⁵.

5. À la HEP Vaud, les étudiants peuvent effectuer leur formation pratique soit dans la classe d'un formateur de terrain soit en emploi. Dans ce dernier cas, ils sont engagés au sein d'un établissement scolaire et sont donc responsables de leurs classes.



Le niveau d'exigence des échelles descriptives et l'anticipation par les évaluateurs d'une baisse des notes liées à la démarche ascendante qui y est inscrite ont amené à plusieurs reprises des questionnements à propos du recrutement.

Les directions d'établissement ont été informées, afin de limiter ce risque, de l'introduction de ces échelles qui leur ont été présentées par la direction de la HEP. Elles ont été très bien accueillies parce qu'elles offrent une définition collective des attentes vis-à-vis d'un enseignant novice. Le fait que cet instrument arrive dans un contexte où les directions d'établissement scolaire expriment le manque d'un tel instrument pour l'évaluation des enseignants une fois engagés, nous laisse penser que ces échelles pourraient, à l'avenir, servir aussi l'évaluation des compétences des enseignants en début de carrière.

L'analyse des notes mises à l'aide de cet instrument d'évaluation, décrite et discutée dans les deux parties suivantes, montre que finalement l'anticipation d'une baisse importante des notes et surtout celle de voir le nombre d'insuffisances augmenter considérablement ne s'est pas réalisée.

L'impact des échelles descriptives sur l'évaluation certificative

Nous avons observé, lors de la phase pilote, une différence significative de notes avec l'introduction des échelles descriptives⁶ (voir Figure 3) :

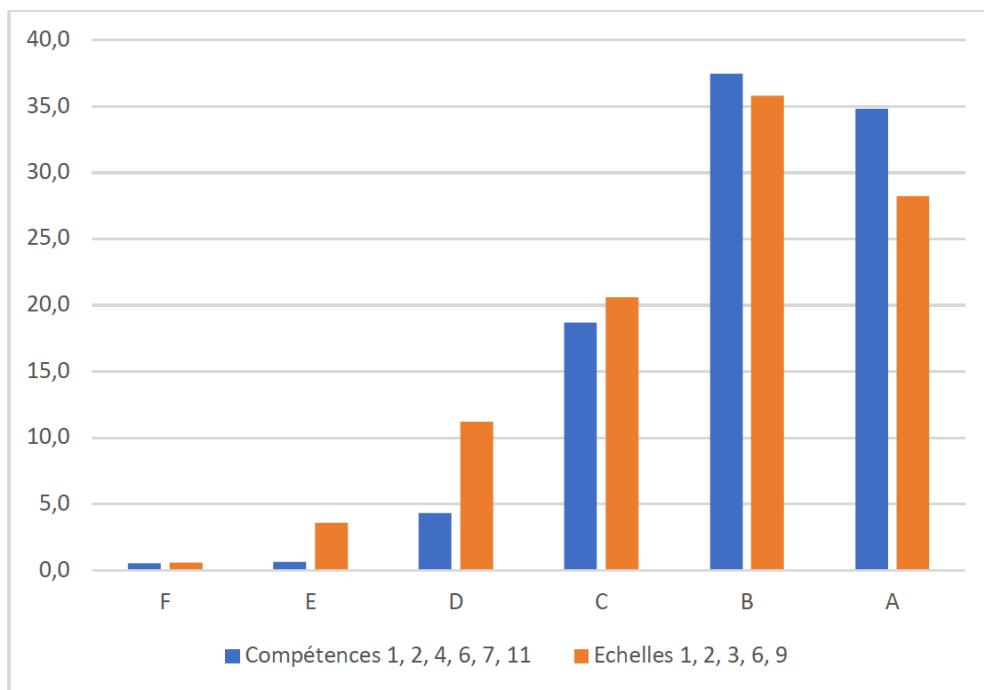


Figure 3 : Répartition des notes attribuées au bilan certificatif de fin de 2^e semestre de formation selon l'instrument d'évaluation utilisé (en pour cent)*

6. Le test du chi-carré montre un lien significatif entre les notes mises et l'instrument d'évaluation utilisé que les insuffisances soient ou non incluses. En effet, leur faible nombre tend à fausser les résultats (chi-carré (4, 1227) = 38.739, p = .000).



* Le nouvel instrument d'évaluation sous forme d'échelles descriptives compte dix échelles et l'instrument d'évaluation précédent structuré autour des compétences du référentiel en comptait onze. Nous avons mis en correspondance les échelles descriptives servant l'évaluation des stages en fin de 1^{re} année de formation (1. Maîtrise des contenus, 2. Planification de la leçon, 3. Animation de la leçon, 6. Conduite de la classe, 9. Maîtrise de la langue d'enseignement), en orange dans le graphique, avec les compétences permettant l'évaluation des mêmes aspects du métier dans l'instrument d'évaluation précédent (1. Professionnel critique et porteur de culture, 2. Engagé dans son développement professionnel, 4. Conception et animation de la leçon, 6. Conduite de la classe, 7. Adaptation de ses interventions, 11. Maîtrise de la langue d'enseignement), représentées en bleu. Cette comparaison, si elle a l'intérêt d'offrir un panorama général a, cependant, deux limites non négligeables. D'une part, les critères de l'évaluation avant et après l'introduction des échelles descriptives sont sensiblement différents, le formateur de terrain n'évalue donc pas la même chose avec un instrument et avec l'autre. D'autre part, ce ne sont ni les mêmes évaluateurs, ni les mêmes évalués. Ces deux éléments, entre autres, sont à considérer dans l'interprétation de ce résultat.

Le nombre de E et de D a augmenté au détriment du nombre de A. De nombreuses variables participent à expliquer cette différence, nous n'observons, dans tous les cas, pas la baisse de notes généralisée que les formateurs de terrain anticipaient ni l'augmentation massive du nombre d'insuffisances. Comme précédemment, la dispersion montre qu'une majorité des notes attribuées sont des A et des B.

Avoir des balises plus précises a été salué par nombre de formateurs de terrain, comme avoir un instrument d'évaluation qui accompagne l'évaluateur dans l'opérationnalisation des énoncés textuels qui figurent dans le référentiel de compétences et qui les soutient dans la définition des attentes et des seuils entre les notes (voir Figures 4 et 5). Néanmoins, les enjeux de la certification ainsi que les représentations que les évaluateurs ont des notes participent vraisemblablement au fait qu'elles soient polarisées vers le haut de l'échelle dès qu'il s'agit de certifier des compétences.

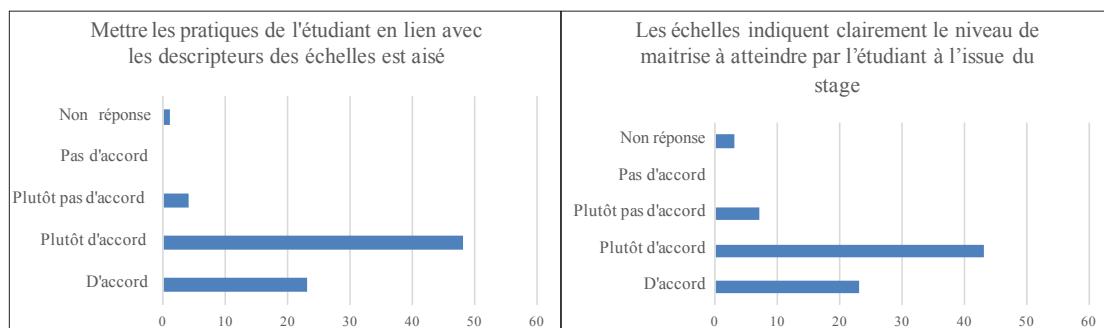


Figure 4 :
Facilité d'utilisation des échelles*

* En pour cent

Figure 5 :
Clarté du niveau de maîtrise attendu*

Les formateurs comme les étudiants ont, pour beaucoup, l'idée que la note C, qui devrait témoigner d'un bon niveau de maîtrise, est une note tout juste suffisante. Dès lors, mettre une note inférieure enverrait le signal d'un dysfonctionnement.

Pour limiter l'effet des représentations liées à la signification de chaque note, le programme informatique de saisie des bilans d'évaluation a été conçu de manière à faire apparaître, dans un premier temps les attentes de la note E. Si l'évaluateur estime que les pratiques observées satisfont ces attentes, ils cliquent sur le mot «OUI», dans le cas contraire sur le «NON». En cas de réponse affirmative apparaissent les attentes de la note supérieure et ainsi



de suite jusqu'à l'obtention de la note maximale. En cas de réponse négative, il est demandé à l'évaluateur de commenter son évaluation avant le passage à l'échelle suivante. Il n'est pas aisé pour les acteurs de l'évaluation de se distancier de l'idée qu'ils ont à attribuer des notes en fonction des représentations qu'ils en ont. Dans la démarche présentée ici, il ne leur est pas demandé de mettre des notes, mais d'attester l'atteinte des objectifs de la formation, la note découlant automatiquement de leurs réponses. Ce renversement de perspective n'est pas évident pour certains évaluateurs qui trouvent, dès lors, la mise des notes trop compliquée. De tels renversements requièrent sans doute du temps avant d'espérer pouvoir en évaluer les éventuels effets.

Les échelles semblent être globalement prises au sérieux par les évaluateurs dans le sens où le suivi formatif des étudiants a été impacté, nous y reviendrons. Cela laisse penser qu'elles jouent le rôle qui leur est dévolu : celui de balise. En revanche, les échelles descriptives, à elles seules, ne permettent pas de dépasser certaines difficultés, largement documentées, liées à l'évaluation des pratiques professionnelles, comme la relation de proximité qui existe entre l'évaluateur et l'évalué dans une formation qui met en scène un formateur de terrain pour un voire deux étudiants (Lebel, 2007 ; Lehmann, 2009 ; Surdez, *et al.*, 2013), la relation de pouvoir qui existe entre évaluateur et évalué ou entre évaluateurs dans le cas où il existe une forme de coévaluation (Berrebi-Hoffmann, 2010 ; Pasche Gossin, 2012 ; Perrenoud, 1996), le fait que l'évaluation est un acte de communication (Mottier Lopez & Allal, 2008), une négociation (Weiss, 1991). Si elles ne permettent pas le dépassement de ces difficultés, les échelles ont eu le mérite de permettre le débat entre formateurs autour de ce qui est considéré comme des critères importants de la professionnalité des enseignants, de ce qui est signe d'une pratique experte ou d'une pratique novice, de ce qui est minimalement attendu pour pouvoir confier une classe à un enseignant débutant, du rôle de l'école et de la place à accorder à chaque aspect du métier dans l'évaluation. Elles peuvent, aussi, servir la communication entre les différents acteurs de l'évaluation, ces échelles étant utilisées par tous : formateur de terrain, formateur de la HEP qui effectue des visites, étudiant et, peut-être, direction d'établissement scolaire qui évalue les compétences des enseignants en fonction.

L'impact des échelles sur l'évaluation formative

Bien que les échelles descriptives n'aient finalement pas influencé la dispersion des notes attribuées autant que les commentaires des utilisateurs le laissaient penser, elles semblent avoir eu un effet non négligeable sur la quantité de signalements⁷ effectués, c'est-à-dire sur l'annonce à la HEP d'une suspicion d'échec au stage ou de difficultés significatives (voir Tableau 3) :

7. Pour rappel, un signalement déclenche l'organisation de visites de stage et, éventuellement, la mise sur pied d'une conférence intermédiaire, c'est-à-dire une réunion à laquelle participent les évaluateurs et l'évalué, dans une visée formative et diagnostique pour faire le point sur la situation.



Tableau 3: Évolution du nombre de signalements et de conférences depuis l'introduction des échelles descriptives

	2016	Phase pilote 2017	1 ^{re} phase de géné- ralisation 2018	2 ^e phase de géné- ralisation 2019
Nombre de stages	3433	3617	3695	3962
Signalements	38	70	120	124
Conférences intermédiaires ou de régulation	30	44	61	84

Source : statistiques réalisées par le service chargé de l'organisation des stages de la HEP

Dans ce tableau, nous voyons que le nombre de placements en stage augmente chaque année (entre +2% et +7% par rapport à l'année précédente) en cohérence avec l'augmentation du nombre d'étudiants que compte la HEP. Cependant, les signalements ainsi que les conférences intermédiaires ou de régulation⁸ se sont multipliés dans une proportion plus importante. Les signalements ont quasi doublé l'année de la phase pilote d'introduction des échelles descriptives. L'année suivante, ils doublent quasiment à nouveau avec la première phase de la généralisation des échelles et l'année suivante, bien que la généralisation continue, le nombre de signalements semble se stabiliser. En ce qui concerne les conférences intermédiaires et de régulation, leur croissance se situe entre +37% et +46% par rapport à l'année précédente avec une tendance au ralentissement également. Ces résultats pourraient être liés à la «nouveauté», les évaluateurs ayant peut-être davantage besoin du soutien de l'institution les premières années de mise en œuvre d'un tel instrument. Afin de s'en assurer, il s'agira de regarder si la tendance se renverse à l'avenir.

Nous en déduisons que les échelles descriptives ont essentiellement influencé la formation des étudiants et le suivi qui leur est proposé. Il en ressort qu'elles pourraient, d'une part, avoir finalement touché ce qui constitue le cœur de la professionnalisation des étudiants en stage et, d'autre part, avoir leur plein effet dans une évaluation formative, auquel le signalement participe, qui est bien moins porteuse d'enjeux que la certification, notamment parce qu'elle ne compte pas les futurs recruteurs comme destinataires.

Les enjeux considérables de l'évaluation certificative notamment celui du recrutement des diplômés et l'importance pour les évaluateurs de ne pas désavantager des étudiants qu'ils jugent compétents sur le marché du travail participant, selon nous, à expliquer ce résultat. Au vu de ce qui précède, les notes ont-elles vraiment leur place dans une évaluation certificative ?

8. Lorsque le signalement concerne des difficultés relationnelles entre le formateur de terrain et l'étudiant ou des difficultés organisationnelles liées au stage, une conférence de régulation est organisée en lieu et place d'une conférence intermédiaire.



Est-ce que mettre des notes a finalement du sens ?

Le fait d'avoir plus de deux positions à l'échelle de notes (réussi/échoué) pour évaluer les compétences des étudiants en stage en enseignement est parfois remis en question. C'est l'orientation prise par Portelance et Caron (2017), dans leur article sur le rapport que les étudiants entretiennent à l'évaluation et ce qu'elle peut générer. L'évaluation provoquerait, en suivant les résultats de la recherche de Maunier (2015) que ces auteures présentent, une possible orientation davantage vers la réussite que vers un réel développement de ses compétences. Portelance et Caron (2017) aboutissent à la conclusion que signifier le succès ou l'échec aux différentes compétences permettrait d'éviter la quête de performance plutôt que celle d'apprentissages plus profonds et durables par les étudiants. Se contenter d'une notation dichotomique serait d'autant plus pertinent dans le cadre de l'évaluation certificative, surtout du bilan final de la formation, qui est au cœur de forts enjeux (Bain, 2014). Cependant, il nous faudrait encore explorer si les enjeux de l'évaluation ne se reporteraient pas, dès lors, sur la rédaction des commentaires associés à la réussite ou à l'échec.

En outre, certains étudiants, surtout lorsqu'ils peinent à estimer si leur prestation est juste suffisante ou alors largement suffisante, expriment le souhait de recevoir une note en plus des commentaires rédigés. La note serait donc une indication supplémentaire qui aiderait l'étudiant à interpréter les commentaires. Le même besoin de notes est parfois aussi exprimé par les évaluateurs. À la HEP Vaud, une seule note, le F, permet de sanctionner l'insuffisance. Ceci peut représenter une difficulté pour un formateur de terrain qui peine à faire entendre à l'étudiant que ses prestations sont largement en deçà des attentes. Le seul usage de mots rédigés sous la forme d'un commentaire associé à une unique note ne permet pas toujours de le faire de manière suffisamment claire. L'utilisation ou non de notes fait, selon nous, l'objet d'un dilemme professionnel.

Pour avancer, néanmoins, sur ce sujet, une remarque a traversé la quasi-totalité des journées de formation à l'utilisation des échelles descriptives : le sens de mettre une note globale au stage alors que chaque échelle est déjà notée. Cette contrainte est liée au fait qu'une note unique doit figurer sur le relevé semestriel des étudiants. Cette note globale est critiquée, car elle est choisie par l'évaluateur sans étalement du collectif et est donc susceptible d'être très largement marquée par la subjectivité de chacun. Ces utilisateurs proposent un dispositif hybride : un dégradé de notes pour chaque échelle (A à F) puis la conversion du résultat selon une échelle en deux positions sur le relevé de notes semestriel.

Au-delà de ce débat, ce qui nous paraît une voie intéressante pour amoindrir les effets indésirables de la certification, tout en maintenant un système de notes, est la réduction du nombre d'évaluations certificatives afin de donner davantage de place à la formation.



Conclusion

En retracant les étapes de la mise en œuvre d'un instrument d'évaluation des compétences des étudiants en stage pour l'enseignement au secondaire obligatoire en Suisse et ses effets, cet article a mis en lumière que le nouvel instrument d'évaluation sous forme d'échelles descriptives remplit son rôle de balise et semble pris au sérieux. En effet, le suivi formatif des étudiants a été globalement impacté par son introduction. Il n'amène, néanmoins, pas une baisse massive des notes ni une augmentation significative du nombre d'insuffisances comme les utilisateurs le présageaient.

Les difficultés du travail de certification, mais surtout, selon nous, la multiplicité des acteurs concernés par le résultat d'une évaluation certificative (évalué, évaluateur, institution qui délivre les diplômes, recruteurs, etc.) permet de comprendre que les notes, au moment de la certification, n'aient pas connu la baisse anticipée par les évaluateurs.

Cependant, l'introduction des échelles descriptives semble avoir influencé la formation des étudiants en stage et le suivi qui leur est proposé. Étant donné qu'il ne compte pas le recruteur comme destinataire potentiel, le suivi formatif est bien moins porteur d'enjeux. Il en ressort que ce nouvel instrument d'évaluation toucherait finalement le cœur de la professionnalisation des étudiants en stage.

Cet article a l'intérêt de montrer l'influence que peuvent avoir sur les résultats de l'évaluation un de ses enjeux, le recrutement, et certains obstacles au travail de certification comme la relation de proximité qui se crée entre l'évaluateur et l'évalué (Lebel, 2007), les relations de pouvoir qui traversent l'évaluation (Berrebi-Hoffmann, 2010) et le fait qu'elle est un acte de communication (Mottier Lopez & Allal, 2008). Ces obstacles traversent finalement bien d'autres, si ce n'est toutes les formations professionnalisantes. Ceci amène à se questionner sur le sens d'attribuer des notes.

Plusieurs recherches restent néanmoins à faire pour affiner ces résultats, les mettre en perspective et explorer l'utilisation qui est faite de cet instrument d'évaluation : une recherche auprès des directions d'établissement et une autre auprès des étudiants.



Références

- André, B. (2015). L'usage du référentiel de compétences par les formateurs de terrain pour évaluer les étudiants en stages. *Evaluer. Journal international de recherche en éducation et formation*, 1(3), 57-72.
- André, B. et Zinguinian, M. (2016). L'absence de cadre comme obstacle dans l'évaluation des stagiaires par les formateurs de terrain dans les formations à l'enseignement. *Phronesis*, 6(4), 10-21.
- Bain, D. (2014). Généralisabilité et évaluation des compétences: pistes et fausses pistes. Dans C. Dierendonck et al. (dir.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (p. 141-165). De Boeck Supérieur.
- Bardin, L. (2007). *L'analyse de contenu*. Presses universitaires de France.
- Berrebi-Hoffmann, I. (2010). Évaluation et élitisme: d'une alliance à l'autre. *Cahiers internationaux de sociologie*, 1(128-129), 79-90.
- Broyon, M.-A., Canevascini, S., Forster, M., Girinshuti, C., Gremion, F., Losego, P., Rey Pellissier, J. et Zampieri, S. (2013). *Insertion professionnelle des enseignants de Suisse Romande et du Tessin*. INSERCH
- Chauvigné, C. (2016). Évaluer des compétences: un jugement d'adaptabilité. *Éducation et francophonie*, 44(2), 86-108.
- Chung Wei, R.R. et Pecheone, R.L. (2010). Assessment for learning in preservice teacher education. Dans M. Kennedy (dir.), *Teacher assessment and the quest for teacher quality: a hanbook*. Jossey-Bass.
- HEP Vaud. (2013). *Formation des enseignantes et enseignants. Référentiel de compétences professionnelles* (2^e éd.). Haute école pédagogique du canton de Vaud.
- HEP Vaud. (2019). *Évaluation du stage de 3^e année. Échelles descriptives et commentaires. Année 2019-2020*. Haute école pédagogique du canton de Vaud.
- Laveault, D. (2008). Le jugement professionnel: foyer de tensions et de synergies nouvelles en évaluation scolaire. *Revue suisse des sciences de l'éducation*, 30(3), 483-500.
- Lebel, C. (2007). Les formateurs de terrain et l'évaluation de la pratique. Dans L. Bélair, D. Laveault et C. Lebel (dirs.), *Les compétences professionnelles en enseignement et leur évaluation* (p.85-102). Presses de l'Université d'Ottawa.
- Lehmann, R. (2009). *Évaluer les gestes professionnels des futurs enseignants: l'impossible mission des formateurs en établissement?* [Mémoire de maîtrise en sciences de l'éducation, Université de Genève]. Archive ouverte UNIGE. <https://archive-ouverte.unige.ch/unige:3831/ATTACHMENT01>
- Losego, P. et Girinshuti, C. (2019). Trouver un emploi dans l'enseignement: l'insertion professionnelle des enseignants en Suisse romande. Dans H. Durler et P. Losego (dir.), *Travailler dans une école. Sociologie du travail dans les établissements scolaires en Suisse romande*. Alphil-Presses universitaires suisses.
- Maunier, S. (2015). *L'évaluation formative en classe, dans les interactions entre enseignants et étudiants, selon une perspective ethnométhodologique*. [Mémoire de maîtrise, Université du Québec].
- Mayen, P. (2005). *L'évaluation des compétences: un pari sur l'avenir*. Séminaire des responsables de formation CNFEPJJ, Vaucresson. <http://www.intefp-sstfp.travail.gouv.fr/datas/files/SSTFP/Resp%20Comp%C3%A9tence%20est%20elle%20mesurable.pdf>
- Merle, P. (2007). *Les notes: secrets de fabrication*. Presses universitaires de France.
- Mottier Lopez, L. et 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.
- N'Da, P. (2015). *Recherche et méthodologie en sciences sociales et humaines. Réussir sa thèse, son mémoire de master ou professionnel, et son article*. L'Harmattan
- Pasche Gossin, F.M. (2012). *Agir en formateur dans une formation des enseignants par alternance: une analyse de l'activité des formateurs dans trois dispositifs de formation par alternance* [thèse de doctorat en sciences de l'éducation, Université de Genève]. Archive ouverte UNIGE. <https://archive-ouverte.unige.ch/unige:24167/ATTACHMENT01>
- Perrenoud, P. (1996). L'obligation de compétences: une évaluation en quête d'acteurs. *Éducateur*, 11, 23-29.
- Perrenoud, P. (2001a). Évaluation formative et évaluation certificative: des postures définitivement contradictoires? *Formation professionnelle suisse*, 4, 25-28.
- Perrenoud, P. (2001b). Les trois fonctions de l'évaluation dans une scolarité organisée en cycles. *Éducateur*, 2, 19-25.
- Portelance, L. et Caron, J. (2017). Logique du stagiaire dans son rapport à l'évaluation. *Phronesis*, 6(4), 85-98.
- Quivy, R. et Van Campenhoudt, L. (1988). *Manuel de recherche en sciences sociales*. Dunod



- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. De Boeck.
- Surdez, M., Sainsaulieu, Y. et Poglia Milet, F. (2013). L'investissement du rôle chez des directeurs des ressources humaines dans les secteurs de l'horlogerie et de la banque en Suisse. Dans M. Perrenoud (dir.), *Les mondes pluriels de Howard S. Becker* (p. 29-47). La Découverte.
- Tillema, H.H. (2009). Assessment for learning to teach. Appraisal of practice teaching lessons by mentors, supervisors, and student teachers. *Journal of teacher education*, 60(2), 155-167.
- Tourmen, C. (2015). L'évaluation des compétences professionnelles : apports croisés de la littérature en évaluation, en éducation et en psychologie du travail. *Mesure et évaluation en éducation*, 38(2), 111-144.
- Weiss, J. (1991). *L'évaluation: problème de communication*. Congrès de l'Association pour le développement des méthodologies d'évaluation en éducation (ADMEE), Neuchâtel, 24-26 septembre 1990. Delval.
- Zarka, Y.-C. (2009). L'évaluation : un pouvoir supposé savoir. *Cités*, 1(37), 113-123.
- Zinguinian, M. (2019). Former la relève. Problèmes de prédition dans l'évaluation des stagiaires. Dans H. Durier et P. Losego (dir.), *Travailler dans une école. Sociologie du travail dans les établissements scolaires en Suisse romande* (p. 147-168). Alphil-Presses universitaires suisses.
- Zinguinian, M. et André, B. (2015). Approche comparative de deux instruments d'évaluation relatifs à l'aptitude à enseigner. *Évaluer. Journal international de recherche en éducation et formation*, 1(3), 73-93.
- Zinguinian, M. et André, B. (2016). Certifier les stagiaires : le «dur travail» voire le «sale boulot» du formateur en établissement scolaire. *Phronesis*, 6(4), 71-84.
- Zinguinian, M. et André, B. (2018). La construction d'échelles descriptives pour évaluer des stages en enseignement : analyse d'un processus à visée consensuelle. *Formation et pratiques d'enseignement en questions*, 23, 229-247.