

Pour citer cet article : Roth, M., & Ruf, I. (2024). Des ressources évaluatives pour les enseignant·es romand·es : une démarche intercantonale. *La Revue LEEE*, 8. <https://doi.org/10.48325/rleee.008.08>

# DES RESSOURCES ÉVALUATIVES POUR LES ENSEIGNANT·ES ROMAND·ES

## *Une démarche intercantonale*

Murielle Roth, Isaline Ruf

---

Version de la publication : mars 2024

Évaluation ouverte et collaborative

Rétroacteur·rices : Dominique Broussal et membres du réseau IEAN-CH

---

## Résumé

Cet article présente les *Pistes pour l'évaluation (PistEval)*, un site Internet sur lequel sont mises à disposition des enseignant·es romand·es des ressources évaluatives pour la 8<sup>e</sup> année (élèves de 11-12 ans) en français et en mathématiques. Elles visent à les soutenir dans l'élaboration de leurs évaluations, qu'elles servent à des fins diagnostique, formative ou certificative. Les tâches proposées sont accompagnées d'étayages théorico-didactiques et méthodologiques. Après avoir exposé le contenu des *PistEval*, leurs visées et le contexte du projet dans lequel elles s'inscrivent, nous décrivons le processus déployé et les acteurs et actrices impliqué·es dans leur élaboration. Nous nous focalisons ensuite sur leurs étayages, que nous présentons plus en détail, avant d'évoquer les éléments qui ont facilité la mise en place des *PistEval*, ainsi que les difficultés rencontrées. Finalement, nous revenons sur les principaux apports du projet, à savoir la définition de certains principes communs pour l'évaluation de la compréhension de l'écrit en français et de la résolution de problèmes en mathématiques, ainsi que sur les pistes de réflexion proposées. Nous évoquons également la perspective des *PistEval* qui est de faire évoluer les pratiques évaluatives des enseignant·es, de manière à ce qu'elles soient au service des apprentissages des élèves.

**Mots-clés :** ressources évaluatives, principes évaluatifs communs, évaluation au service des apprentissages, pratiques évaluatives, enseignant·es romand·es

## Abstract

This article presents the PistEval, a website which provides to teachers from the French-speaking part of Switzerland some assessment resources for French and Mathematics for grade 8 (pupils of 11-12 years old). They aim to support teachers in developing their assessments, whether they are used for diagnostic, formative or certification purposes. The tasks proposed on this website are accompanied by theoretical, didactic, and methodological scaffolding. First, we outline the content of the PistEval, their aims and the context of the project in which they are used. Next, we describe the process used and the actors involved in its development. We then focus on their scaffolding, that we describe in more detail. After that, we discuss the elements that facilitated the implementation of the PistEval, as well as the difficulties encountered. Finally, we come back to the main contributions of the project, namely the definition of some common assessment principles for French reading comprehension and problem solving in Mathematics, as well as the reflections proposed. We also discuss the perspective of the PistEval, which is to develop teachers' assessment practices so that they serve the pupils' learning.

**Keywords:** assessment resources, common assessment principles, assessment for learning, assessment practices, teachers from the French-speaking part of Switzerland

## 1. Introduction et contexte

La Suisse, comme de nombreux autres pays, a décidé d'assurer un monitoring basé sur la mesure des performances des élèves dans une perspective de rentabilité et d'efficacité de son système éducatif, et d'harmonisation au niveau national et régional. Pour ce faire, elle a participé aux tests internationaux PISA dès 2000 et mis en place des tests nationaux<sup>1</sup> depuis 2016. De son côté, la Suisse romande a voulu renforcer la coopération entre les cantons latins, ceux-ci étant souverains en matière d'éducation. Cette volonté s'est traduite par l'établissement d'une Convention scolaire romande (CSR), qui stipule notamment que des épreuves externes (EE) communes doivent être organisées, dans le but de vérifier si les élèves de la partie francophone du pays atteignent les objectifs du Plan d'études romand (PER) entré en vigueur dès 2010. De telles épreuves devaient également servir à s'assurer des acquisitions des élèves avec parfois des enjeux de promotion et/ou d'orientation, pouvant ainsi remplacer, dans certains cantons, les épreuves cantonales externes existantes. Derrière ces EE se retrouvent à la fois des enjeux d'efficacité et/ou d'équité des systèmes et des enjeux de régulation des actions des enseignant-es (Broadfoot, 2000).

Les effets des EE sur les pratiques des enseignant-es ont été pointés par un certain nombre de recherches, tout comme la question des bénéfices qu'une telle épreuve peut avoir sur les apprentissages des élèves. Certains travaux (Yerly, 2017b ; Yerly & Maroy, 2017) ont par exemple montré que les enseignant-es utilisent les EE pour voir quels sont les contenus à enseigner et pour ajuster leurs propres outils d'évaluation sommative (Yerly, 2017a) mais qu'ils et elles se penchent rarement sur les résultats de leurs élèves pour faire évoluer leurs pratiques. Lorsque des ajustements sont réalisés, certains effets négatifs apparaissent (Koretz, 2017) comme le « *teaching to the test* », le rétrécissement du curriculum, la fragmentation des savoirs enseignés ou encore la remise en question par les enseignant-es

<sup>1</sup> Ces tests nationaux visent à vérifier l'atteinte des compétences fondamentales (COFO).



de leurs propres compétences (Amrein-Beardsley et al., 2010 ; Au, 2007 ; Malet, 2015 ; Mons, 2009 ; Osborn, 2006).

Les réflexions menées dès 2004 au niveau romand par l'Institut de recherche et de documentation pédagogique (IRDp), mandaté par la Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP), ont donné lieu à des propositions de scénario pour une évaluation basée sur le PER (Marc et al., 2013 ; Marc & Wirthner, 2012). Elles ont également mis en évidence le panorama déjà bien rempli des dispositifs de mesure externes en Suisse romande et la nécessité – donc primordiale – de clarifier la visée que pourrait prendre une nouvelle épreuve (De Pietro et al., 2016 ; De Pietro & Roth, 2016 ; Marc & Wirthner, 2012). L'ensemble des analyses menées et les constats réalisés ont conduit les autorités éducatives romandes à finalement décider, en 2015, de développer une base de données de tâches évaluatives dont une partie serait au service des cantons pour leurs propres épreuves externes, et une autre mise à disposition des enseignant·es romand·es. Ces dernier·ères y ont accès via le site *Pistes pour l'évaluation* (*PistEval*).

Cet article a pour objectif de présenter ces *PistEval*, leur contenu, leurs visées et leur origine, le processus de leur élaboration ainsi que la contribution des acteurs et actrices impliqués·es. Nous abordons aussi les éléments ayant joué un rôle de facilitateurs ainsi que les difficultés rencontrées lors de la mise en place des *PistEval*. Puis, nous explicitons les apports des travaux menés dans le cadre de l'élaboration de ces *Pistes pour l'évaluation* : développement de principes évaluatifs communs et de pistes de réflexion pour une évaluation au service des apprentissages des élèves.

## 2. Les *PistEval* : contenu, visées et origine

Les *PistEval*, qui se présentent sous forme de pages Internet, contiennent des exemples de tâches évaluatives en lien avec les objectifs du PER, actuellement pour la compréhension de l'écrit en français<sup>2</sup> et la résolution de problèmes en mathématiques<sup>3</sup> pour des élèves de 8<sup>e</sup> année (élèves de 11-12 ans). Elles sont accompagnées d'étayages théorico-didactiques et méthodologiques sur lesquels nous nous focalisons plus particulièrement dans cet article. Ceux-ci attirent l'attention des enseignant·es sur le statut illustratif des tâches proposées, qui demandent donc aux enseignant·es d'opérer une sélection et/ou des ajustements de ces tâches pour composer leurs évaluations en classe, en fonction des objectifs poursuivis, de ce qui a été enseigné et des apprentissages visés. Ainsi, au regard des choix et ajustements effectués sur ces tâches évaluatives, les enseignant·es peuvent construire des évaluations à des fins diagnostique, formative ou certificative.

Les étayages jouent un rôle primordial puisqu'ils visent à soutenir les pratiques d'évaluation en classe, non seulement pour qu'elles soient un soutien à l'apprentissage, mais aussi en invitant les enseignant·es à (re)questionner leurs pratiques (Pasquini & DeLuca, 2022). Autrement dit, ces étayages accompagnent, d'une part, les enseignant·es dans la prise en main des tâches illustratives proposées et, d'autre part, ils mettent en lumière certaines réflexions essentielles à avoir lors du choix ou de l'élaboration de tâches pour évaluer les apprentissages réalisés par les élèves. Ces réflexions suggèrent que l'évaluation n'est pas qu'une démarche technique ou une méthode servant à la sélection (De Pietro et al., 2016) mais bien un processus dans lequel les enseignant·es conservent toute leur autonomie

<sup>2</sup> <https://bdper.plandetudes.ch/login/?next=/eprocom/francais-8/>

<sup>3</sup> <https://bdper.plandetudes.ch/login/?next=/eprocom/math-8/>



professionnelle en faisant des choix éclairés au service des apprentissages de leurs élèves. En entrant dans ce processus réflexif, ils et elles gardent le contrôle de leurs évaluations, ce qui devrait les amener à réguler leurs pratiques (Carette & Dupriez, 2009).

Les *PistEval* sont une des concrétisations d'un projet plus large mené en Suisse romande, à savoir le projet *EpRoCom-Banque d'items*, émanant de décisions politiques. Ce projet vise une certaine harmonisation de l'interprétation des objectifs du PER dans les différents cantons, par le biais de l'évaluation de ses contenus (dans la perspective du développement d'une culture évaluative commune). Dès le début des années 2000, il y a eu une volonté forte des cantons romands d'harmoniser la scolarité obligatoire avec la rédaction d'un plan d'études cadre romand - le PER. Cette volonté a été réaffirmée par le peuple et les cantons lors d'une votation populaire en mai 2006 sur de nouveaux articles constitutionnels sur la formation<sup>4</sup>. Dans ce contexte est né, en 2007, l'Accord intercantonal sur l'harmonisation de la scolarité obligatoire (art. 8, al. 2) : « Plans d'études, moyens d'enseignement et instruments d'évaluation, ainsi que standards de formation sont coordonnés entre eux » (Conférence suisse des directeurs cantonaux de l'instruction publique - CDIP, 2007). Parallèlement à cet Accord, la CSR a été élaborée afin de renforcer l'Espace romand de la formation, notamment en amenant les cantons à coopérer dans certains domaines de la scolarité obligatoire (Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin - CIIP, 2007). Les travaux de rédaction du PER ont ainsi débuté, offrant dès 2010, un référentiel intercantonal d'enseignement-apprentissage (mais pas d'évaluation), et par la suite de nombreux moyens d'enseignement romands. Au niveau de l'évaluation, le projet *EpRoCom-Banque d'items* a commencé en 2009, dans le but de répondre à l'article 15 de la CSR<sup>5</sup>.

Bien que ce projet visait, au départ, la création d'épreuves romandes communes (d'où l'acronyme *EpRoCom*), il a été réorienté en 2015 par l'Assemblée plénière de la CIIP (AP-CIIP), réunissant les sept ministres de l'éducation des cantons romands, vers la création d'une *Banque romande d'items* (ci-après *Banque*), sans pour autant être renommé. Un certain nombre de travaux exploratoires menés par l'IRD - structure scientifique de la CIIP - (notamment, de Pietro et al., 2016 ; Marc & Wirthner, 2012) ont conduit à cette réorientation. L'IRD s'est donc attelé à la création et à l'alimentation de cette *Banque*. Les tâches évaluatives insérées dans celle-ci sont caractérisées au moyen de descripteurs qui permettent leur classification. Une partie de ces tâches a ensuite été extraite et mise à disposition des enseignant-es romand-es en tant qu'illustrations sur les *PistEval*. Elles sont donc à ajuster par les enseignant-es lorsqu'ils et elles élaborent leurs évaluations en classe. Comme mentionné plus haut, elles ont ensuite été complétées par des étayages théorico-didactiques et méthodologiques.

<sup>4</sup> <https://www.ciip.ch/Espace-romand-de-la-formation/Espace-romand-de-la-formation/Espace-romand-de-la-formation>

<sup>5</sup> CSR, article 15 :

<sup>1</sup> La CIIP organise des épreuves romandes communes à l'Espace romand de la formation, en vue de vérifier l'atteinte des objectifs du plan d'études.

<sup>2</sup> En fin de cycle ou à la fin du degré secondaire I, si la discipline choisie pour l'épreuve romande commune correspond à celle d'un test de référence vérifiant un standard national, le test de référence peut servir d'épreuve commune.



### 3. Le processus d'élaboration des *PistEval*

Les exemples de tâches mises à disposition sur les *PistEval* sont puisés dans la base de données de tâches évaluatives. Pour la créer et l'alimenter, les autorités éducatives ont décidé d'initier le travail à partir d'épreuves cantonales externes (EC) des différents cantons romands, tout d'abord pour les disciplines du français et des mathématiques. Ainsi, les cantons romands ont privilégié une mise en commun de leurs expériences cantonales autour de l'évaluation en classe sachant que, dans la plupart des cantons, les EC font l'objet d'un processus de validation avec prétest(s), ce qui devrait en garantir une certaine qualité. La première étape du travail a donc consisté à rassembler les EC de 8<sup>e</sup> année en français et en mathématiques ayant déjà été soumises aux élèves dans les cantons qui les ont produites puis, pour l'équipe de projet composée de cinq chercheuses des deux disciplines, à réaliser un état des lieux de ce qui est évalué dans ces épreuves. Dans un deuxième temps, des représentant-es de tous les cantons romands ont été réuni-es via la COMEPRO, la commission de coordination du projet, pour lui présenter cet état des lieux et lui faire part des premiers constats qui en ressortaient. Sur cette base, la COMEPRO a décidé de se focaliser, dans un premier temps, sur la compréhension de l'écrit en français (cet axe étant systématiquement évalué dans tous les cantons concernés) et la résolution de problèmes en mathématiques (étant donné que cette compétence fait partie des visées prioritaires du PER et que son évaluation est reconnue comme complexe).

L'équipe de projet a donc commencé par analyser puis sélectionner un certain nombre de tâches en compréhension de l'écrit et en résolution de problèmes provenant des EC des différents cantons. Les EC de certains cantons étant utilisées à des fins d'orientation (dans des filières ou niveaux au cycle 3), l'équipe a tout d'abord vérifié que les tâches étaient en adéquation avec le PER, autrement dit qu'elles n'évaluaient pas des objectifs dépassant les attendus du curriculum de référence. À la suite de ce travail, l'équipe de projet a pu bénéficier de l'appui et de l'expérience pratique d'un groupe de résonance (GRés) composé d'un-e enseignant-e par canton romand, ainsi que de l'expertise d'un groupe de conseiller-ères didactiques (GDid) par discipline, chacun composé de trois chercheur-ses de la discipline concernée. En effet, les tâches retenues ont été soumises aux enseignant-es du GRés qui ont dû les approuver (ou non), moyennant d'éventuels ajustements, au regard de leur pertinence par rapport au public cible et à leur utilisation en classe, ainsi que de leur adéquation aux pratiques de leur canton respectif. Par ailleurs, pour certains questionnements ou points particuliers, les GDid ont aussi été sollicités, plus précisément pour discuter de la validité didactique du matériel d'évaluation. À la suite de ces échanges, une partie des tâches évaluatives analysées *a priori* a été éprouvée auprès d'une centaine d'élèves de chaque canton romand lors d'un test pilote mené en 2019. Le but premier de ce test était de vérifier la « romandité » des tâches retenues, c'est-à-dire qu'elles sont comprises et réalisées de manière similaire par des élèves de 8<sup>e</sup> année des différents cantons romands. À cette fin, les réponses aux questions pour le français et les productions des élèves pour les mathématiques ont été analysées<sup>6</sup>. Sur la base des résultats de ces analyses, une partie des tâches a été validée (après modifications dans certains cas) puis insérée dans la *Banque*. La dernière étape du processus, et non des moindres, a été la création des *PistEval*, à savoir la mise à disposition aux enseignant-es romand-es de tâches provenant de la *Banque*. Lors de cette étape, l'équipe de projet a vu la nécessité de compléter ces tâches avec un discours accompagnateur concrétisé par différents étayages

<sup>6</sup> Un rapport de travail, rédigé à l'intention des membres de l'AP-CIIP, présente les principaux résultats et constats suite au test pilote (Roth et al., 2021).



ayant été discutés avec le GRés et/ou les GDid qui en ont suggéré des ajustements avant leur mise en ligne.

## 4. Les étayages des *PistEval*

Les étayages proposés sont de différentes natures. Ils accompagnent les enseignant-es romand-es dans la prise en main des tâches présentes sur les *PistEval* et mettent notamment l'accent sur des aspects importants à prendre en compte lors de l'évaluation de la compréhension de l'écrit en français et de la résolution de problèmes en mathématiques. Les tâches y sont présentées comme des exemples de pratiques et non comme des évaluations « clés en mains », afin de limiter les effets négatifs (André, 2013 ; Au, 2007 ; Jones, 2007) et les conséquences dommageables que ces ressources externes pourraient avoir sur les pratiques évaluatives des enseignant-es (Maroy & Cattonar, 2002 ; Pasquini & DeLuca, 2022). En effet, les enseignant-es doivent en faire une sélection et éventuellement les ajuster en fonction des objectifs poursuivis et des contenus enseignés en classe. Dans le même ordre d'idée, c'est à elles et eux que revient sciemment l'attribution des points à l'évaluation. Le lien et la cohérence indispensables entre enseignement et évaluation sont ainsi soulignés. Par exemple, en mathématiques ce sont des exemples de tâches, indépendants les uns des autres, qui sont mis à disposition sur les *PistEval*. De cette manière, lorsque les enseignant-es constituent leur évaluation, ils et elles sont amené-es à les sélectionner et éventuellement les adapter au regard du travail réalisés en classe et des apprentissages visés. Pour la compréhension de l'écrit en français, il existe, dans quelques cas, des variantes de questions portant soit sur le format de questionnement, soit sur l'opération cognitive mobilisée. Ces étayages les amènent donc à s'interroger sur ce qu'ils et elles souhaitent évaluer, dans quelle perspective et de quelle manière. Ils insistent aussi sur certains aspects qui peuvent avoir un impact sur leurs évaluations lorsqu'ils et elles choisissent, élaborent et/ou analysent les tâches qui les composent (voir figures 2 et 5).

L'ambition des *PistEval* n'est pas de proposer des ressources évaluatives et des étayages exhaustifs, couvrant l'ensemble des compétences et connaissances attendues par le PER, mais de mettre à disposition des exemples de tâches explicitant les réflexions qui ont orienté les choix lors de l'élaboration de ceux-ci. Cette ambition vient d'une volonté intercantonale, à savoir de susciter certaines réflexions chez les enseignant-es dans le but qu'ils et elles se les approprient et puissent les transférer lorsqu'ils et elles évaluent d'autres compétences et connaissances. Les réflexions quant à la difficulté des textes initiées par les étayages pour « le texte qui raconte », seul regroupement de genres actuellement illustré sur les *PistEval*, peuvent par exemple être élargies à d'autres genres textuels, voire transférées à d'autres disciplines telles que l'allemand. Il est aussi espéré que ces réflexions qui concernent en premier lieu l'évaluation des apprentissages soient réinvesties pour l'enseignement-apprentissage, dans une perspective d'alignement curriculaire (Anderson, 2002 ; Biggs, 1999 ; Pasquini, 2019).

Le premier type d'étayages est d'ordre théorico-didactique : ils précisent ce qui est à entendre par *comprendre* en français et, de manière analogue, par *problème* en mathématiques (voir figure 1).



## Figure 1

Définitions de « comprendre » et « problème » sur les PistEval

<p>« <b>Comprendre</b> » est un processus complexe et multidimensionnel qui résulte de l'interaction, peu observable directement, entre le texte, le lecteur et le contexte. C'est seulement à partir des réponses à un certain nombre de questions qu'il est possible d'inférer le processus de compréhension.</p> <p><b>Comprendre, c'est arriver à avoir une image mentale cohérente du texte à partir des différentes informations du texte et de ses connaissances préalables</b> (Giasson, 2003, 2013).</p>	<p><b>Qu'est-ce qu'un problème ?</b></p> <p>Un certain consensus existe pour considérer qu'un « problème » consiste en une question plus ou moins contextualisée, dont la démarche de résolution n'est pas immédiatement disponible pour l'élève, au regard de son âge et des apprentissages réalisés.</p> <p>Ainsi une <b>résolution de problèmes</b> peut être définie comme la convocation autonome, puis l'utilisation adéquate par l'élève d'un ou plusieurs outils mathématiques.</p>
---	---

D'autres apports mettent en avant des points de vigilance qu'il est essentiel de garder à l'esprit lors de toute pratique évaluative. En français, l'accent a été mis sur le niveau de difficulté des ressources évaluatives, plus particulièrement sur les paramètres entrant en jeu dans une telle estimation – tant pour les textes (voir figure 2) que pour les questions s'y rapportant. Ces paramètres, quelque peu simplifiés afin de les rendre plus facilement accessibles, y sont explicités.

## Figure 2

Les critères pour estimer la difficulté des textes (extrait tiré des PistEval)

<p>▼ <b>LA STRUCTURE DU TEXTE</b></p> <p>Il est possible d'attribuer un indice de difficulté comme proposé ci-dessous :</p>	
<p><b>? Questions à se poser</b></p>	<p><b>Procédure pour attribuer un indice de difficulté</b> Facile, intermédiaire, difficile</p>
<p>Est-ce que la structure du texte est <b>canonique</b> (correspond-elle à ce qu'on attend du genre de texte concerné) ?</p> <p>Est-ce que des organisateurs textuels sont présents ?</p>	<p><b>Facile</b></p> <ul style="list-style-type: none"> <li>- Structure canonique sans organisateurs <b>ou</b></li> <li>- Peu éloigné de la structure canonique mais avec des organisateurs qui structurent davantage</li> </ul>
	<p><b>Intermédiaire</b></p> <ul style="list-style-type: none"> <li>- Structure canonique sans organisateurs mais avec la nécessité d'en ajouter <b>ou</b></li> <li>- Peu éloigné de la structure canonique avec des organisateurs qui perturbent <b>ou</b></li> <li>- Structure non canonique avec des organisateurs qui structurent</li> </ul>
	<p><b>Difficile</b></p> <ul style="list-style-type: none"> <li>- Structure canonique avec des organisateurs qui perturbent la compréhension car ils sont mal utilisés <b>ou</b></li> <li>- Peu éloigné de la structure canonique mais sans organisateurs <b>ou</b></li> <li>- Structure non canonique, sans organisateurs</li> </ul>
<p>► <b>LA CONSTRUCTION DES PHRASES</b></p> <p>► <b>LA FRÉQUENCE DU VOCABULAIRE UTILISÉ</b></p>	



En mathématiques, l'importance de la démarche en résolution de problèmes a été mise en évidence, ainsi que l'analyse et l'interprétation des traces des élèves. Ces observations permettent d'identifier d'éventuelles erreurs et de pouvoir proposer ensuite un feedback ou des remédiations appropriées à l'élève. D'autres étayages plus théoriques présentent une catégorisation des tâches mathématiques<sup>7</sup> (inspirées de celle proposée par Houdement, 2017).

D'autres étayages encore se veulent plutôt d'ordre méthodologique et traitent du *comment* évaluer. Tant en français qu'en mathématiques, des conseils notamment quant au choix des tâches lors de l'élaboration d'une évaluation (en compréhension de l'écrit pour le français, de manière plus générale pour les mathématiques, voir figure 3) sont proposés aux enseignant-es. En français, des clés d'estimation de la difficulté des textes et des questions associées en compréhension de l'écrit sont présentées au moyen d'indicateurs précis.

### Figure 3

*Pistes pour l'élaboration d'une évaluation de mathématiques (extrait tiré des PistEval)*

Concevoir une évaluation n'est pas chose aisée. Il s'agit tout d'abord d'être au clair sur les compétences des élèves que l'on désire évaluer. Il peut s'agir de l'évaluation de :

1. **Connaissances** ayant acquis le statut de savoirs suite au travail réalisé en classe, à l'image du nombre de côtés isométriques d'un losange ou du sens de l'expression « le tiers de ... », **ou savoir-faire « techniques »** comme lister les diviseurs d'un nombre ou réaliser une division (tâches techniques) ;
2. La **mobilisation autonome de savoirs ou de savoir-faire** dans une tâche, tels que transformer des grandeurs dans la même unité afin de les additionner (tâches d'application) ;
3. La **résolution de problèmes** requérant non seulement la mobilisation de savoirs et/ou savoir-faire mais aussi la structuration et la combinaison de ceux-ci dans une procédure de résolution (problèmes) ;
4. La **recherche et la mise en œuvre de stratégies** telles que le tâtonnement, l'ajustement d'essais successifs ou encore la mise en place d'une systématique de recherche (problèmes atypiques).

En fin de thème, ce sont principalement **les compétences 3 et 4**, en lien avec la résolution de problèmes, qui devraient être au cœur de l'évaluation. Ce sont des tâches souvent considérées difficiles par les élèves et, en particulier pour la catégorie 4, qui peuvent être chronophages. En outre, il est souvent pertinent voire nécessaire de s'assurer aussi que l'élève maîtrise certains outils (compétences 1 et 2), souvent requis mais pas forcément indispensables à la résolution de problèmes, à travers des tâches techniques ou d'application. Le cas échéant, ces types de tâches permettent d'évaluer le niveau de maîtrise des notions travaillées. Somme toute, il s'agit de **trouver un juste équilibre entre les différents types de tâches évaluatives** soumises à l'élève.

## 5. La mise en place des *PistEval*

La mise en place des *PistEval* a été d'une très grande richesse au niveau des expériences et des échanges réalisés avec les différents acteurs et actrices impliqués, incluant un certain nombre d'éléments facilitateurs présentés ci-dessous. Mais certaines difficultés ont aussi émergé. Une première difficulté a concerné le choix des tâches à retenir pour les enseignant-es romand-es, parmi l'éventail de celles disponibles dans les EC. Quant aux

<sup>7</sup> Quatre catégories ont été définies : tâches techniques, tâches d'application, problèmes et problèmes atypiques.





étayages, la plus grande difficulté a relevé de leur formulation à laquelle un soin particulier a été apporté et qui a été discutée avec le GRés. En effet, il convenait en particulier de les rédiger sous forme de « réflexions et points de vigilance à avoir quand on évalue », tout en laissant une marge de manœuvre aux enseignant·es pour qu'ils et elles restent « maitres » ou « maitresses » de leurs évaluations en pouvant faire leurs propres choix. Ce sont ainsi la forme plutôt descriptive ou l'usage du conditionnel qui ont été privilégiés, de sorte à formuler des recommandations mais de manière moins directive ou prescriptive.

## 5.1 Les éléments facilitateurs

Plusieurs éléments ont permis de faciliter la création des *PistEval*. Tout d'abord, la structure scientifique de l'IRDP offre la possibilité de rassembler divers·es acteur·rices romand·es (notamment des responsables d'épreuves cantonales, des enseignant·es, des didacticien·nes et des chercheur·ses) pour traiter des contenus qu'il convenait de mettre à disposition des enseignant·es romand·es dans la perspective d'une évaluation au service des apprentissages. Deuxièmement, les matériaux de base pour ce projet, soit les épreuves cantonales externes, représentaient un solide corpus de départ pour les analyses à mener et l'élaboration des *PistEval*. Enfin, le PER étant commun à tous les cantons, il a servi de référence pour s'assurer qu'une tâche évaluative était en adéquation avec ses objectifs et pour rédiger des étayages théorico-didactiques et méthodologiques basés sur ses intentions disciplinaires.

## 5.2 Les difficultés rencontrées

Pouvoir réunir en un même projet plusieurs personnes concernées par l'évaluation a l'avantage de permettre de confronter la diversité des points de vue, mais cela prend du temps pour prendre certaines décisions et construire un consensus. Au-delà de cette difficulté inhérente à la multiplicité des regards, plusieurs difficultés sont apparues au cours du projet.

Tout d'abord, le projet – et donc la création des *PistEval* – émane de décisions politiques. Au départ, il s'inscrit donc dans une logique top-down. Pour « compenser » en partie cette logique et tenter de répondre à des besoins concrets des enseignant·es en matière d'évaluation au service des apprentissages, il s'est avéré nécessaire d'identifier ces besoins, sans pour autant avoir la possibilité d'interroger directement le terrain. En ce sens, l'implication du GRés dans le dispositif s'est avérée capitale pour proposer des ressources qui « parlent » aux enseignant·es et soient utiles à leur pratique. Un compromis a ainsi dû être trouvé pour que les *PistEval* soient à la fois porteuses pour le terrain et que simultanément elles répondent à la demande politique de développer certains principes évaluatifs communs.

La manière de formuler les étayages des *PistEval* a également représenté une difficulté. En effet, il s'est avéré nécessaire de trouver une formulation qui soit perçue comme un soutien par les enseignant·es. Il convenait donc de leur donner des conseils, notamment par des apports théorico-didactiques et méthodologiques en matière d'évaluation, tout en les laissant disposer d'une marge de manœuvre suffisante pour qu'ils et elles sentent libres de faire leurs propres choix au regard de la réalité de leur classe. C'est dans cette perspective que les tâches mises à disposition ont été présentées comme des illustrations que les enseignant·es devraient s'approprier et ajuster à l'aune des objectifs poursuivis et de



l'enseignement dispensé, mais aussi des prescriptions cantonales en matière d'évaluation. En effet, les étayages proposés s'adressent aux enseignant-es de tous les cantons romands, alors que leurs pratiques évaluatives varient d'un canton à l'autre (voire au sein d'un même établissement). Dès lors, un ajustement des tâches proposées s'avère nécessaire, bien qu'une certaine cohérence entre les contenus des *PistEval* et les différentes directives cantonales soit assurée grâce aux retours des membres du GRés et de la COMEPRO sur ceux-ci. Par ailleurs, un soin a été apporté pour rédiger (une partie de) ces étayages de manière à susciter la réflexion chez les enseignant-es, dans une volonté de les amener à questionner leurs pratiques notamment évaluatives, voire de les faire évoluer. La terminologie utilisée a également fait l'objet de discussions pour qu'elle soit compréhensible et corresponde autant que possible à celle utilisée par les enseignant-es de tous les cantons romands. In fine, la formulation des étayages devait permettre une appropriation intuitive des tâches et donner suffisamment d'informations (mais pas trop) pour que les enseignant-es disposent des éléments essentiels pour concevoir leurs évaluations.

## 6. L'intercantonal au service du changement

La création des *PistEval* dans le cadre du projet *EpRoCom-Banque d'items* est la concrétisation d'un début de réponse à l'article 15 de la CSR. Ce chapitre traite de leurs principaux apports qui augurent des perspectives de changements au niveau intercantonal.

### 6.1 Axe 1 : développement de principes évaluatifs communs

Lors du développement des *PistEval*, différents principes évaluatifs ont été retenus pour chacune des deux disciplines. Ceux-ci ont été définis sur la base des analyses *a priori* des tâches et des résultats au test pilote, de travaux de recherche (Houdement, 2017 ; Roth & Sánchez Abchi, 2021 ; Sánchez Abchi et al., 2016 ; Sayac & Grapin, 2014) ainsi que des nombreux échanges avec les différents groupes de travail. Par ailleurs, ces principes évaluatifs, qui transparaissent sur les *PistEval*, sont en adéquation avec les principes didactiques sur lesquels s'appuie le PER.

Pour la discipline du français, le PER est basé sur le concept de genre de textes (Schneuwly, 2007 ; Schneuwly & Dolz, 1997). En compréhension de l'écrit, il est attendu des élèves qu'ils et elles comprennent des genres de textes diversifiés correspondant à des situations de communication variées. C'est donc dans cette perspective que des principes ont été retenus pour l'évaluation de la compréhension d'un texte écrit, et que la sélection de textes et de questions pour les *PistEval* a été opérée. Premièrement, le choix a été fait de sélectionner uniquement des textes « complets », servant de base à plusieurs questions (les fragments de textes ont ainsi été écartés). Deuxièmement, seules les questions relatives à un tel texte écrit nécessitant la compréhension du texte ont été conservées. Celles auxquelles il est possible de répondre sans l'avoir lu n'en font, par exemple, pas partie. En revanche, les questions évaluant l'axe thématique *fonctionnement de la langue* et contribuant à la compréhension du texte ont été retenues. Un extrait dans lequel il est demandé à l'élève d'identifier tous les mots utilisés par l'auteur-riche pour désigner un personnage en est un exemple.



D'autres principes relatifs au choix des textes et des questions en compréhension de l'écrit, dont le niveau de difficulté peut avoir un impact sur la réussite des élèves, ont été définis (Aeby et al., 2000 ; Cardinet, 1989 ; Eason et al., 2012 ; Mullis et al., 2015 ; OECD, 2019 ; Rog & Burton, 2002 ; Weiss & Wirthner, 1991). Concernant les questions, l'accent a entre autres été mis sur l'opération cognitive qui renvoie au processus de lecture mobilisé. L'emploi de questions faisant appel à différentes opérations cognitives est ainsi apparu incontournable pour rendre compte de la compréhension d'un texte écrit : l'élève doit non seulement être capable de repérer des informations dans le texte, mais également de réaliser des inférences simples et complexes ou encore des interprétations. Selon l'opération cognitive en jeu, la difficulté de la question varie et c'est un autre niveau de compréhension qui est évalué. Sur les *PistEval*, ces principes ressortent par le biais des étayages didactiques quant au soin qu'il conviendrait d'apporter lors du choix ou de l'élaboration des questions au moment de la constitution d'une évaluation (Sánchez Abchi et al., 2022) (voir figure 4).

#### Figure 4

*Pistes pour l'élaboration d'une évaluation en compréhension de l'écrit en français (extrait tiré des PistEval)*

Lors de l'élaboration ou du choix de ressources pour l'évaluation, une attention particulière devrait être portée sur le point suivant :

Non seulement **plusieurs questions** sont indispensables pour permettre d'observer la compréhension de l'élève, mais il s'agit également d'utiliser un ensemble de questions qui mobilisent **différentes opérations cognitives**. En effet, si certaines questions de repérage ou d'inférence de niveau simple peuvent permettre aux élèves d'entrer dans la compréhension du texte et de construire une compréhension plus globale (à placer donc en début d'évaluation de préférence), **le seul usage de questions de repérage ne suffit pas à attester de la compréhension**.

Ils apparaissent aussi à travers les exemples de questions se rapportant à un texte, organisées selon l'opération cognitive à laquelle elles font appel, que les enseignant-es sont amené-es à sélectionner. La présence de variantes, en termes notamment d'opération cognitive mobilisée, proposées pour certaines questions, renforce encore ce principe.

Pour la discipline des mathématiques, le PER met au centre de l'enseignement-apprentissage la résolution de problèmes comme *but* et *moyen* d'apprendre les mathématiques (Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin - CIIP, 2010)). L'importance de la résolution de problèmes est soulignée dans les visées prioritaires du curriculum, pour le domaine *Mathématiques et sciences de la nature* : « Se représenter, problématiser et modéliser des situations et résoudre des problèmes en construisant et en mobilisant des notions, des concepts, des démarches et des raisonnements propres aux Mathématiques [...] » ainsi que dans le premier chapitre de chaque axe thématique spécifique aux mathématiques<sup>8</sup> intitulé « Éléments pour la résolution de problèmes » (CIIP, 2010). Parmi ces éléments figure la « mise en œuvre d'une démarche de résolution ». C'est là le principe central dégagé pour l'évaluation de la résolution de problèmes<sup>9</sup> : l'évaluation doit porter en priorité sur la démarche mise en œuvre par l'élève, bien plus que sur (la justesse de) sa réponse. En effet, l'analyse de sa production lorsqu'il ou elle résout un problème est essentielle (Ministère de l'éducation

<sup>8</sup> Il s'agit des axes thématiques suivants : *Espace, Nombres, Opérations* et *Grandeurs et mesures*.

<sup>9</sup> Il s'agit des tâches catégorisées de « problèmes » et « problèmes atypiques » sur les *PistEval* (les autres catégories comprenant les tâches techniques et d'application).



nationale, de la jeunesse et des sports, 2022): elle permet d'identifier les outils mathématiques convoqués et de distinguer si la procédure qu'utilise l'élève est experte – c'est-à-dire lorsqu'est mobilisé l'outil le plus efficace au regard de son âge et de ses connaissances –, correcte – autrement dit lorsque l'outil convoqué permet d'arriver à la solution mais de façon moins directe – ou encore erronée. Dès lors qu'il s'agit là du cœur de la résolution de problèmes, c'est en examinant finement les productions et les éventuelles erreurs qu'il devient possible d'évaluer le niveau de compétences d'un élève en la matière. L'importance de l'analyse des démarches et des erreurs (Astolfi, 2009) est mise en avant dans les étayages des *PistEval* (voir figure 5).

### Figure 5

*L'importance de la démarche lors de l'évaluation de la résolution de problèmes (extrait tiré des PistEval)*

L'élément principal à analyser dans la production de l'élève est sa **démarche**, c'est-à-dire les outils ou les connaissances mathématiques qu'il ou elle a mobilisés et la manière dont il ou elle les a agencés pour arriver à une solution. Souvent, pour résoudre un problème, plusieurs démarches sont envisageables. Peuvent être distinguées les démarches dites "expertes" des démarches "correctes".

L'analyse de la démarche de l'élève permet à l'enseignant·e de savoir où en est l'élève par rapport à la mise en œuvre de démarches expertes vers lesquelles il s'agit de l'amener progressivement, en faisant ressortir leurs avantages. Elle offre un diagnostic sur la base duquel l'enseignant·e pourra donner un feedback précis à l'élève quant aux éléments encore à travailler (notions et savoir-faire) et envisager des remédiations ciblées au regard des types d'erreurs identifiées.

Cet aspect est renforcé par les exemples de démarches mises en œuvre par des élèves de 8<sup>e</sup> année mises à disposition pour chaque tâche illustrative. Les erreurs y sont analysées et explicitées. Par ailleurs, la présentation des tâches comprenant un encadré dédié à la démarche de l'élève, précédé de la consigne « Montre ce que tu fais pour trouver la réponse » vise à l'inciter à laisser des traces de sa réflexion (et implicitement aux enseignant·es d'en faire une attente).

En outre, des principes concernant l'élaboration d'une évaluation de mathématiques sont aussi proposés aux enseignant·es sur les *PistEval*. Ils conseillent sur le type de tâches qu'il convient de combiner en fonction des compétences à évaluer (voir figure 3).

## 6.2 Axe 2 : pistes de réflexion pour une évaluation au service des apprentissages

Les *PistEval* visent à soutenir les enseignant·es romand·es dans l'élaboration de leurs évaluations en classe, qu'il s'agisse d'évaluations diagnostiques, formatives ou certificatives. Elles ont également été conçues dans la perspective que les réflexions mises en évidence pour une tâche – grâce aux étayages – soient réinvesties par les enseignant·es dans leurs pratiques évaluatives mais aussi dans leur enseignement. En effet, bien que ces réflexions se rapportent d'abord à l'évaluation, elles devraient aussi avoir un impact sur l'enseignement-apprentissage, étant donné leur lien indissociable. De fait, évaluation et enseignement sont des éléments d'un même processus censé être cohérent de bout en bout dans une perspective d'alignement curriculaire (Pasquini, 2019). Les réflexions amenées à travers les étayages ambitionnent ainsi d'entraîner des répercussions sur les pratiques des enseignant·es de manière plus générale. Autrement dit, il est espéré qu'ils et



elles se posent les mêmes questions pour leur enseignement en classe et ajustent leurs pratiques, si nécessaire.

Pour la compréhension de l'écrit en français, les étayages attirent l'attention notamment sur des éléments qui peuvent avoir un impact sur la compréhension d'un texte, à l'image de sa structure. Par exemple, une des tâches proposées comprend un texte rédigé de façon non chronologique et donc caractérisé comme difficile. L'étayage explicite que cette difficulté s'explique par sa structure particulière. Ainsi, à partir des différents étayages, les enseignant·es devraient interroger leurs pratiques, dans ce cas précis se demander si des textes avec une telle structure ont été utilisés en classe lors du travail autour de la compréhension d'un texte, et réagir en conséquence. Dans cette perspective, les étayages proposés pour l'évaluation servent aussi à l'enseignement : si ce dernier est ajusté grâce à ceux-ci, ils contribuent à soutenir les apprentissages.

Quant aux mathématiques, les enseignant·es sont par exemple incité·es à analyser les démarches mises en œuvre et les erreurs de leurs élèves en résolution de problèmes. Il est espéré qu'à partir des constats réalisés (compétences acquises, en construction ou non acquises par exemple), ils et elles proposent une « réponse pédagogique » (des feedbacks et remédiations) basée sur les difficultés observées. Par ailleurs, tout comme c'est le cas pour le français, il est souhaité implicitement que les enseignant·es en fassent de même en phase d'apprentissage, et qu'ils et elles ajustent leur enseignement en fonction des besoins des élèves identifiés à partir de leurs productions.

Ainsi, les étayages et les exemples de tâches proposés sur les *PistEval*, que ce soit en français ou en mathématiques, mettent aussi en évidence des aspects qui devraient être travaillés en classe (et ne le sont peut-être pas suffisamment) avec pour ambition de conduire les enseignant·es à proposer des évaluations qui soient un soutien au service des apprentissages de leurs élèves.

## 7. Conclusion

Tout·e enseignant·e est tenu·e d'évaluer les compétences et connaissances de ses élèves pour savoir où ils et elles en sont dans leurs acquisitions. Sur cette base, l'enseignant·e devrait pouvoir proposer, le cas échéant, des remédiations pour que ses élèves puissent progresser dans leurs apprentissages (Mottier Lopez et al., 2012). Évaluer est une pratique nécessaire, mais difficile. C'est donc principalement dans la perspective de soutenir les enseignant·es romand·es dans cette tâche que les *PistEval* ont été créées. Celles-ci répondent également à une demande politique, à savoir une évaluation davantage harmonisée au niveau romand, passant par une interprétation partagée du PER, soit des connaissances et compétences attendues des élèves. Pour répondre à ces deux objectifs, l'implication de divers·es acteurs et actrices (responsables d'épreuves cantonales, didacticien·nes, enseignant·es, chercheur·ses, etc.) de l'ensemble de la Romandie dans le dispositif déployé s'est révélée essentielle. La multiplicité des regards portés sur les tâches évaluatives et les étayages les accompagnant a permis de s'assurer de leur cohérence avec les différentes prescriptions cantonales et avec ce qui devrait être enseigné, et leur adéquation pour les classes romandes. Bien que les étayages aient été discutés et ajustés avec les enseignant·es du GRés, ils et elles n'ont pu se prononcer qu'à l'échelle de leur classe et de leur propre établissement (voire de leur région) : les membres de ce groupe ne peuvent garantir une représentativité de l'ensemble de la population enseignante



romande. Il s'agit ici d'une limite des travaux menés, à savoir la difficulté, voire l'impossibilité, d'atteindre une plus grande proportion des établissements et des régions.

Les *PistEval* offrent un soutien aux enseignant-es romand-es lors de l'élaboration de leurs évaluations, en continuité avec les formations initiales et continues dispensées dans les instituts de formation. En effet, tant les tâches évaluatives qui ont un statut illustratif que les étayages ont pour but de leur permettre d'interroger leurs pratiques en tant que praticien-nés réflexif-ves (Cattonar & Maroy, 2000 ; Schön, 2017). Plus spécifiquement, les étayages cherchent à faire entrer les enseignant-es dans un processus réflexif et une certaine manière de réfléchir, en se questionnant notamment sur le contenu et la composition de leurs propres évaluations. En compréhension de l'écrit en français, s'interroger sur les opérations cognitives mobilisées dans les différentes questions posées aux élèves à partir de la lecture d'un texte, par exemple, permet non seulement de mieux percevoir les différents niveaux de compréhension en jeu et leur difficulté mais aussi de s'assurer que les questions choisies évaluent la compréhension du sens général du texte (Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin - CIIP, 2010). Ainsi, cela devrait permettre aux enseignant-es d'avoir une meilleure vision des compétences et connaissances de leurs élèves en jeu lors de la compréhension d'un texte. Évaluation et enseignement faisant partie d'un même processus, les étayages touchant aux pratiques évaluatives devraient, par « ricochet », amener les enseignant-es à également réfléchir à leur enseignement, et éventuellement à l'ajuster. En ce qui concerne les mathématiques, l'intention des étayages est de faire porter le regard des enseignant-es sur la démarche de l'élève, afin de lui permettre d'évaluer où il ou elle en est par rapport à la mobilisation des notions étudiées. Cette volonté d'amener les enseignant-es à s'intéresser à ce qu'a fait l'élève (bien plus qu'au résultat) vise à renforcer le lien entre évaluation et apprentissage. L'analyse du travail de l'élève se doit d'être au service de la planification de la suite du travail en classe en fonction des observations réalisées.

En somme, les étayages proposés pour la discipline du français portent plutôt sur une réflexion en amont de l'évaluation (choix des tâches évaluatives, ce qu'elles évaluent et leurs impacts sur la compréhension d'un texte), pouvant amener les enseignant-es à ajuster leur enseignement avant l'acte évaluatif, tandis que ceux définis en mathématiques interviennent davantage en aval de l'évaluation (analyse des démarches et des erreurs) et permettent de construire la suite des apprentissages. Dans les deux cas, ces apports visent une évolution des pratiques évaluatives au service des apprentissages de l'élève.

Accessibles depuis l'automne 2021, les *PistEval* portent actuellement sur la compréhension de l'écrit en français et la résolution de problèmes en mathématiques, pour des élèves de 8<sup>e</sup> année. L'évolution de ces pages Internet, facilitée par leur format numérique, passe notamment par la poursuite de leur alimentation tant en termes d'étayages (en les complétant ou les ajustant par exemple au regard de nouveaux résultats de recherche ou à partir de nouvelles expérimentations), que d'exemples de tâches évaluatives (en les élargissant à d'autres domaines, disciplines et/ou degrés scolaires). Sur la base de productions d'élèves, les *PistEval* pourraient également être étoffées avec des pistes de remédiation en s'appuyant, par exemple sur des activités proposées dans les moyens d'enseignement. Dans ce processus de développement des *PistEval*, il est selon nous nécessaire de se pencher encore plus sur les besoins concrets et les difficultés des enseignant-es. Une perspective pour les travaux à venir pourrait ainsi s'inscrire dans une logique davantage « bottom-up », en renforçant la participation du GRés dans le dispositif ou alors en développant un dispositif collaboratif, afin de pouvoir encore mieux répondre aux besoins du terrain.



La difficulté de diffusion des *PistEval* auprès des enseignant-es romand-es est une autre limite du projet. Autrement dit, comment faire pour qu'ils et elles en aient connaissance et s'y intéressent ? Dès lors que la communication d'une telle nouveauté auprès du public concerné n'incombe pas à l'IRDIP ni à la CIIP mais à chaque canton, la marge de manœuvre au niveau romand est considérablement restreinte. Au mieux, nous avons pu annoncer la mise en ligne des *PistEval* dans la rubrique « Actualité » de la plateforme du PER. Par ailleurs, nous n'exerçons aucune influence sur la manière dont ces ressources sont concrètement prises en main par les enseignant-es romand-es, ni sur les usages qu'ils et elles en font. Nous n'avons pas non plus de retour de la part de celles et ceux qui auraient exploité les *PistEval* pour l'élaboration de leurs évaluations. Pouvoir susciter et/ou solliciter des retours sur leur utilisation, et ce à différents niveaux (enseignant-es, collaborateurs et collaboratrices responsables de la discipline, etc.) représente une perspective future dans l'optique de permettre une évolution constructive de ces pages Internet.

## Références

- Aeby, S., De Pietro, J.-F., & Wirthner, M. (2000). Français 2000. Dossier préparatoire. L'enseignement du français en Suisse romande : Un état des lieux et des questions. *Institut de recherche et de documentation pédagogique*.
- Amrein-Beardsley, A., Berliner, D. C., & Rideau, S. (2010). Cheating in the first, second, and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives*, 18, 14. <https://doi.org/10.14507/epaa.v18n14.2010>
- Anderson, L. W. (2002). Curricular alignment: A re-examination. *Theory into practice*, 41(4), 255-260. [https://doi.org/10.1207/s15430421tip4104\\_9](https://doi.org/10.1207/s15430421tip4104_9)
- André, B. (2013). Les ruses des enseignants face à l'évaluation des établissements et du système scolaire. In V. Dupriez & R. Malet (Eds.), *L'évaluation dans les systèmes scolaires. Accommodements du travail et reconfiguration des professionnalités* (pp. 51-62). De Boeck.
- Astolfi, J.-P. (2009). *L'erreur, un outil pour enseigner* (9<sup>e</sup> éd.). ESF.
- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258-267. <https://doi.org/10.3102/0013189X07306523>
- Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher education research & development*, 18(1), 57-75.
- Broadfoot, P. (2000). Un nouveau mode de régulation dans un système décentralisé : L'Etat évaluateur. *Revue française de pédagogie*, 130(1), 43-55. <https://doi.org/10.3406/rfp.2000.1052>
- Cardinet, J. (1989). Évaluer sans juger. *Revue française de pédagogie*, 88, 41-52. <https://doi.org/10.3406/rfp.1989.1412>
- Carette, V., & Dupriez, V. (2009). La lente émergence d'une politique scolaire en matière d'évaluation des élèves : Quinze ans de transformations en Belgique francophone. *Mesure et évaluation en éducation*, 32(3), 23-45. <https://doi.org/10.7202/1024930ar>
- Cattonar, B., & Maroy, C. (2000). Rhétorique du changement du métier d'enseignant et stratégie de transformation de l'institution scolaire. *Éducation et sociétés*, 6, 21-42.
- Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP). (2007). *Convention scolaire romande*.
- Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP). (2010). *Plan d'études romand (PER)*. CIIP. <https://www.plandetudes.ch>



- Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP). (2007). *Accord intercantonal sur l'harmonisation de la scolarité obligatoire (concordat HarmoS)*.
- De Pietro, J.-F., & Roth, M. (2016). Pour une évaluation sous contrôle didactique. Qu'est-ce qui permet de dire qu'un élève a «compris» un texte oral? *Babylonia*, 2. <http://babylonia.ch/fr/archives/2016/numero-2/de-pietro-roth/>
- De Pietro, J.-F., Roth, M., & Sánchez Abchi, V. S. (2016). Vers un référentiel pour l'évaluation : Choix politiques et enjeux didactiques. *L'évaluation en classe de français, outil didactique et politique*, 8, 111.
- Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, 104(3), 515-528. <https://doi.org/10.1037/a0027182>
- Houdement, C. (2017). Résolution de problèmes arithmétiques à l'école. *Grand N, Revue de mathématiques, de sciences et technologie pour les maîtres de l'enseignement primaire*, 100.
- Jones, B. D. (2007). The unintended outcomes of high-stakes testing. *Journal of applied school psychology*, 23(2), 65-86. [https://doi.org/10.1300/j370v23n02\\_05](https://doi.org/10.1300/j370v23n02_05)
- Koretz, D. (2017). *The testing charade: Pretending to make schools better*. University of Chicago Press.
- Malet, R. (2015). Politiques d'accountability et devenir de la profession enseignante aux Etats-Unis. *Evaluer. Journal international de Recherche en Education et Formation (e-JIREF)*, 1(1), 27-44.
- Marc, V., & Wirthner, M. (2012). *Épreuves romandes communes: De l'analyse des épreuves cantonales à un modèle d'évaluation adapté au PER* (p. 48).
- Marc, V., Wirthner, M., & Uldry, S. (2013). *Développement d'un modèle d'évaluation adapté au PER*. IRDP, Institut de recherche et de documentation pédagogique. <http://doc.rero.ch/record/235788>
- Maroy, C., & Cattonar, B. (2002). *Professionnalisation ou déprofessionnalisation des enseignants? Le cas de la Communauté française de Belgique*.
- Ministère de l'éducation nationale, de la jeunesse et des sports. (2022). *La résolution de problèmes mathématiques au cours moyen*. <https://eduscol.education.fr/document/32206/download?attachment>
- Mons, N. (2009). *Les effets théoriques et réels de l'évaluation standardisée*. 45. <https://doi.org/10.4000/rfp.1531>
- Mottier Lopez, L., Tessaro, W., Dechamboux, L., & Villabona, F. M. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions Vives*, 6(18), 159-175. <https://doi.org/10.4000/questionsvives.1235>
- Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2015). PIRLS 2016 Reading Framework. In I. V. S. Mullis & M. O. Martin (Eds.), *PIRLS 2016 assessment framework* (2<sup>nd</sup> éd., pp. 11-30). TIMSS & PIRLS International Study Center.
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing.
- Osborn, M. (2006). Changing the context of teachers' work and professional development: A European perspective. *International Journal of Educational Research*, 45(4-5), 242-253. <https://doi.org/10.1016/j.ijer.2007.02.008>
- Pasquini, R. (2019). Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : Enjeux et perspectives. *Mesure et évaluation en éducation*, 42(1), 63. <https://doi.org/10.7202/1066598ar>





- Pasquini, R., & DeLuca, C. (2022). Standardized testing vs teachers' classroom assessment practices. *Canadian Teacher Magazine*.
- Rog, L. J., & Burton, W. (2002). Matching Texts and Readers: Leveling Early Reading Materials for Assessment and Instruction. *The Reading Teacher*, 55(4), 348-356.
- Roth, M., Ruf, I., Abchi, V. S., Soussi, A., & Weiss, L. (2021). *EpRoCom : Dispositif romand de mutualisation de tâches évaluatives : Premiers constats*.
- Roth, M., & Sánchez Abchi, V. (2021). Évaluer la difficulté d'une épreuve de compréhension de l'écrit : Expérimentation d'un outil sur une pratique enseignante. *La Revue LEE*, 3. <https://doi.org/10.48325/rlee.003.02>
- Sánchez Abchi, V., De Pietro, J.-F., & Roth, M. (2016). *Evaluer en français : Comment prendre en compte la difficulté des items et des textes*. Institut de recherche et de documentation pédagogique.
- Sánchez Abchi, V., Roth, M., & Matei, A. (2022). Estimer la difficulté des questions en compréhension de l'écrit en français Vérification empirique d'un modèle théorique. *Évaluer. Journal international de recherche en éducation et formation*, 1(8), 29-46.
- Sayac, N., & Grapin, N. (2014). Évaluer les capacités des élèves à résoudre des problèmes dans le cadre d'une évaluation externe, en France : Les spécificités de la forme QCM. *Éducation et francophonie*, 42(2), 64. <https://doi.org/10.7202/1027906ar>
- Schneuwly, B. (2007). *Genres écrits et oraux et forme scolaire. Enseignement et apprentissage de la langue première à l'école*. Presses universitaires de Namur.
- Schneuwly, B., & Dolz, J. (1997). Les genres scolaires. Des pratiques langagières aux objets d'enseignement. *Repères*, 15(1), 27-40. <https://doi.org/10.3406/reper.1997.2209>
- Schön, D. A. (2017). *The reflective practitioner: How professionals think in action*. Routledge.
- Weiss, J., & Wirthner, M. (1991). Evaluation de la compréhension de l'écrit correspondant au niveau de fin de deuxième année primaire. In J. Weiss & M. Wirthner (Eds.), *Enseignement du français. Premiers regards sur une rénovation* (pp. 265-296). Delval.
- Yerly, G. (2017a). Quel impact des politiques de responsabilisation douce sur les pratiques enseignantes ? In Y. Dutercq & C. Maroy (Eds.), *Professionnalisme enseignant et politiques de responsabilisation* (pp. 121-141). De Boeck Supérieur. <https://doi.org/10.3917/dbu.duter.2017.01.0121>
- Yerly, G. (2017b). Évaluation des apprentissages en classe et évaluation à grande échelle : Quels sont les effets des épreuves externes sur les pratiques évaluatives des enseignants ? *Mesure et évaluation en éducation*, 40(1), 33-60. <https://doi.org/10.7202/1041003ar>
- 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, 93-108. <https://doi.org/10.4000/rfp.5326>

## Remerciements

Nous remercions vivement Verónica Sánchez Abchi, Laura Weiss et Viridiana Marc, toutes trois également impliquées dans l'élaboration des *PistEval*, pour leurs relectures minutieuses et leurs précieux conseils. Nos remerciements vont également aux membres de IEAN-CH pour leurs relectures bienveillantes et leurs suggestions constructives. Toutes ces personnes-relectrices nous ont permis d'améliorer notre article de manière substantielle.



Murielle Roth

Collaboratrice scientifique à l'Institut de recherche et de documentation pédagogique (IRDP) à Neuchâtel. Chercheuse dans le domaine de l'évaluation des apprentissages et en didactique du français. Elle prépare une thèse de doctorat qui s'intéresse à la difficulté des textes et des questions utilisés pour évaluer la compréhension de l'écrit des élèves.

Isaline Ruf

Collaboratrice scientifique à l'Institut de recherche et de documentation pédagogique (IRDP) à Neuchâtel et doctorante en didactique des mathématiques à l'Université de Genève. Ses travaux et intérêts portent principalement sur la résolution de problèmes et l'évaluation des apprentissages en mathématiques.

