

EpRoCom : dispositif romand de mutualisation de tâches évaluatives

Premiers constats

Murielle Roth, Isaline Ruf, Verónica Sánchez Abchi, Anne Soussi, Laura Weiss, sous la direction de Viridiana Marc

Le projet EpRoCom

Inscrit dans la Convention scolaire romande (CSR) de 2007, l'article relatif aux *Épreuves romandes communes* (EpRoCom) a fait l'objet de travaux exploratoires par l'Institut de recherche et de documentation pédagogique (IRDp) et d'une *Task force* de la CIIP en 2012, pour déterminer les orientations à retenir. Alors que se mettaient en place des tests de référence au niveau suisse (CoFo-CDIP), l'Assemblée plénière de la Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (AP-CIIP) décidait de réorienter le projet EpRoCom vers la création d'une *Banque romande d'items** (ci-après *Banque*). En 2015, suite à la définition et l'adoption d'un Masterplan, les travaux de mise en commun de ressources évaluatives débutaient.

Les épreuves externes cantonales correspondent à toute évaluation conçue par des personnes ou organismes autres que l'enseignant.e (autorités scolaires notamment). Les différentes appellations utilisées dans les cantons pour désigner les évaluations externes ont été regroupées dans cette locution.

L'intention d'assembler ainsi du matériel évaluatif est de créer une culture commune en matière d'évaluation entre les cantons romands, au regard des apprentissages inscrits dans le Plan d'études romand (PER). Pour ce faire, les cantons ont partagé leurs épreuves externes (voir encadré) dont les tâches évaluatives ont fait l'objet d'analyses, de sélection, d'ajustements et de validation, dans le but d'en mettre certaines à disposition des cantons et du corps enseignant, ceci au cours de la rentrée 2021 sur la plateforme du PER. Le projet *EpRoCom/Banque d'items* est donc un dispositif de mutualisation de ressources évaluatives validées, pertinentes et fiables, fondées sur les *objectifs* et *progressions des apprentissages* du PER.

* Le terme technique « item » est remplacé ensuite par « tâches évaluatives » ou « questions »

Le dispositif

L'IRDp a été chargé du projet *EpRoCom/Banque d'items*. Pour le réaliser, le groupe de conception de l'IRDp est accompagné de groupes consultatifs, composés pour l'un d'enseignant.es issu.es de tous les cantons romands et, pour l'autre, de didacticien.nes de chaque discipline concernée.

Le développement d'une culture commune de l'évaluation s'est focalisé, dans un premier temps, sur deux disciplines présentes dans les épreuves externes des cantons ayant développé une telle évaluation en 8^e année, avec un focus sur certains aspects jugés d'intérêt majeur : le Français (compréhension de l'écrit) et les Mathématiques (résolution de problèmes). Une analyse minutieuse des tâches évaluatives figurant dans 24 épreuves externes cantonales de 8^e année (de 2012 à 2017) a ainsi été menée par le groupe de conception de l'IRDp. Il en a sélectionné un certain nombre qu'il a soumises à ses groupes consultatifs. Outre certaines caractéristiques de pertinence et de validité, il a retenu des tâches présentant un ancrage solide au PER. Un lot de tâches des deux domaines disciplinaires a ainsi été retenu pour réaliser un test pilote soumis à un échantillon d'élèves romand.es de 8^e année en 2019. Le but est de développer une *Banque romande d'items* rassemblant des ressources évaluatives en *compréhension de l'écrit* et en *résolution de problèmes*, ainsi que, pour chacune d'elles, des renseignements permettant de les caractériser.



© AdobeStock

Le test pilote 2019

Environ cinq classes par canton, soit près de 700 élèves de Suisse romande, se sont prêtées à l'exercice, avec pour but de vérifier si des tâches d'origines cantonales différentes pouvaient s'adresser à l'ensemble des élèves romandes de 8^e année. Il ne s'agissait pas de tester les performances de ces élèves, mais bien d'éprouver la

bonne compréhension des tâches, indépendamment du canton dont elles proviennent, en vue de leur éventuelle intégration dans la *Banque*. Les élèves ont également pu donner leur appréciation des tâches (perception de la difficulté et de la familiarité).

La compréhension de l'écrit en Français

Pour le Français, le fort intérêt quant au niveau de lecture des élèves en fin de cycle 2 a induit la focalisation sur la *compréhension de l'écrit*. « Comprendre » 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.

Comprendre, c'est arriver à avoir une image mentale cohérente du texte à partir des différentes informations du texte et de ses connaissances préalables (Giasson, 2003, 2013).

Les tâches qui permettent d'évaluer la compréhension de textes écrits comprennent à chaque fois un texte ainsi que des questions associées qui varient par leur format, les opérations cognitives que les élèves doivent mobiliser pour y répondre (repérer, inférer, interpréter) et leur difficulté, en regard de leur âge. Ainsi, estimer la difficulté à priori du texte et des questions (Sánchez Abchi, De Pietro & Roth, 2016) est crucial pour évaluer les tâches. Une grille d'analyse a pu être mise au point et a permis de caractériser les différents textes et questions, notamment du point de vue de leur difficulté.

La difficulté du texte

La difficulté du texte a été mesurée au moyen de trois indices :

- La *complexité lexicale*: proportion de mots peu fréquents dans un texte ;
- La *complexité syntaxique*: prise en compte du nombre de mots longs ainsi que du nombre de mots par phrase ;
- La *complexité de la structure*: degré de correspondance avec ce qui peut être considéré comme l'organisation textuelle « canonique » du genre discursif considéré.

La difficulté des questions

Elle a principalement été estimée grâce à deux critères : le format de questionnement et le type d'opération cognitive.

Le **format de questionnement** peut varier : QCM, question Vrai/Faux, question à réponse brève, question à réponse textuelle (qui nécessite la formulation d'une réponse plus ou moins longue), question d'ordination (où il s'agit de mettre des énoncés dans l'ordre du texte), question d'appariement (où il s'agit de relier des catégories avec des énoncés), etc.



Le **type d'opération cognitive** mobilisée par les élèves pour répondre aux questions varie en complexité; quatre types ont été observés:

Le **repérage** consiste à localiser une information clairement formulée dans un texte (mot, phrase): il s'agit d'une opération relativement simple (voir ci-contre).

[...] Sarah Marquis est une aventurière suisse née en 1972 dans le Jura. Le gout du voyage l'a amenée en Australie, mais la découverte de la marche s'est opérée en Nouvelle-Zélande. Tu vas découvrir un épisode qu'elle a vécu au départ de son périple aux États-Unis. [...]

Où Sarah Marquis a-t-elle découvert l'attrait de la marche ?

.....

1) Repérage

Les inférences nécessitent la mise en relation d'informations locales ou globales.

Les **inférences peuvent être simples** (niveau 1) quand il s'agit par exemple d'expliquer une affirmation ou de faire une déduction relativement facile en mettant en relation des informations explicites du texte. Dans l'exemple ci-dessous, il s'agit de déduire que le garçon et l'aviateur sont la même personne et que la personne qui raconte est externe à l'histoire.

► **Coche la seule fin de phrase correcte :**

Le narrateur est...

... la tourterelle

... le garçon

... un personnage externe à l'histoire

... l'instituteur

2) Inférence simple

Les **inférences peuvent être complexes** (niveau 2) quand il faut, par exemple, retrouver des informations implicites du texte ou, comme dans l'exemple (voir ci-dessous), identifier les étapes du schéma narratif.

► **Complète le tableau avec les deux résumés qui manquent.**

► **Précise clairement de qui tu parles. Évite donc d'utiliser des pronoms (il, elle, etc.).**

	Parties du schéma narratif	Résumés
1	Situation initiale
2	Complication	Un jour les animaux décident que la situation doit changer.
3	L'action entreprise par le paysan
4	Résolution	La pierre sectionne la dent et la sorcière disparaît.
5	Situation finale	Depuis ce jour, une nouvelle pointe en forme de dent domine le lac.

3) Inférence complexe

L'**interprétation** suppose la compréhension des intentions de l'auteur.e en se focalisant sur des informations locales, globales ou sur l'association de détails avec la thématique globale, le plus souvent en faisant appel aux connaissances antérieures et à la compréhension du monde du lecteur ou de la lectrice (voir ci-contre). Il s'agit de l'opération la plus complexe des quatre.

Quelle morale peut-on tirer de cette histoire ?

.....

.....

.....

4) Interprétation

Matériel testé auprès d'élèves romand.es

En tout, 77 questions portant sur quatre textes ont été testées. En cohérence avec ce qui précède concernant leur niveau de difficulté, elles variaient quant à leur format et au type d'opération cognitive mobilisée. Par ailleurs, quelques questions ont été déclinées en deux ou trois versions pour faire varier leur degré de difficulté.

Les textes sélectionnés pour le test pilote se rapportent majoritairement au regroupement de genre textuel « qui raconte » (cf. PER), regroupement le plus fréquemment

rencontré dans les épreuves cantonales analysées et, dans une moindre mesure, celui « qui relate ».

Dans le PER, les apprentissages liés à la production et la compréhension s'organisent selon les genres de textes; à des fins didactiques de généralisation, ils sont organisés en six regroupements: le texte qui raconte; le texte qui relate; le texte qui argumente; le texte qui transmet des savoirs; le texte qui règle des comportements; le texte qui joue avec la langue – texte poétique.

Premiers résultats

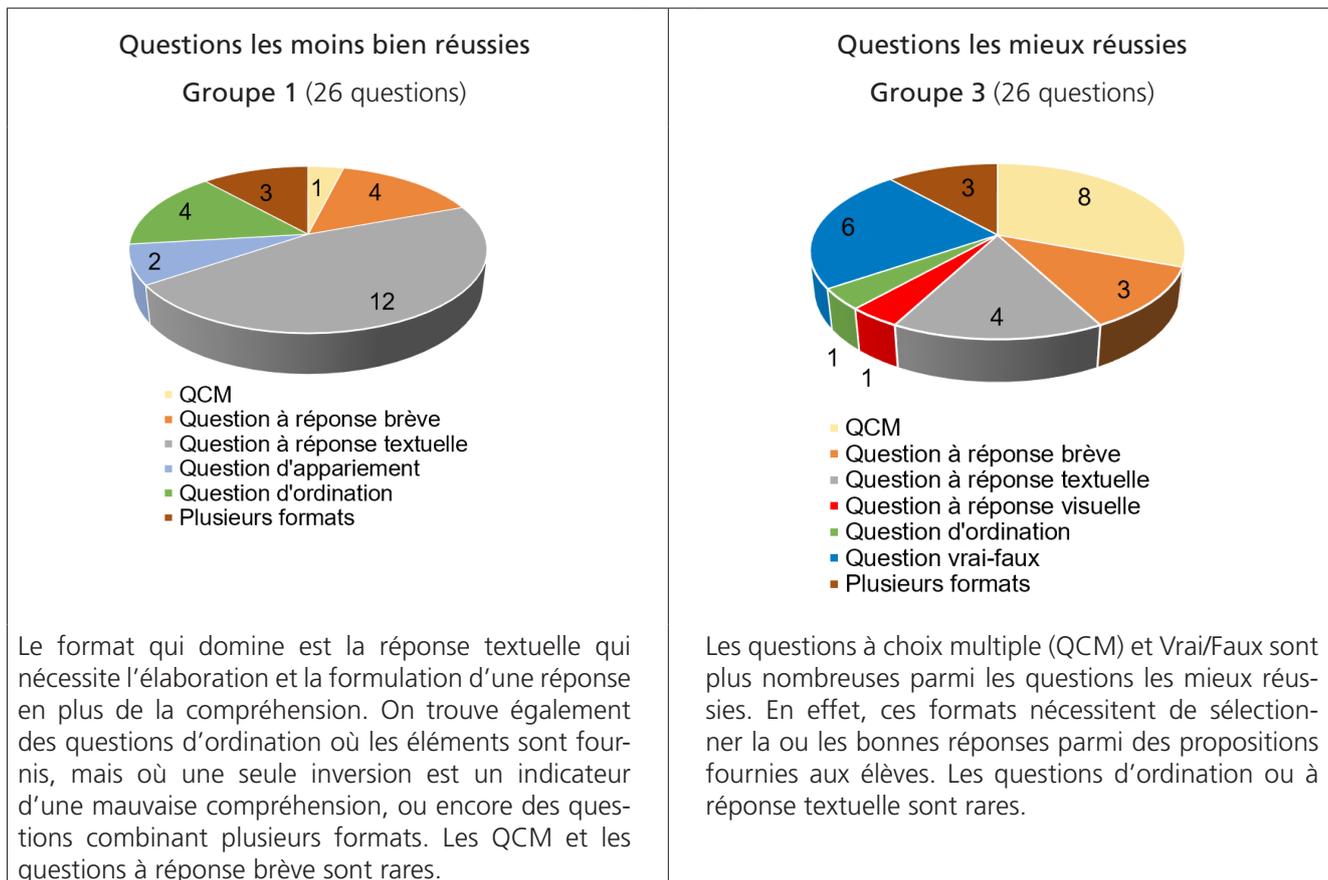
Le test pilote a permis de vérifier l'adéquation des questions aux élèves des cantons romands. Pour en rendre compte, la réussite aux questions a notamment été observée, ce qui constitue un indice important quant à la difficulté des questions. Il ressort que, dans l'ensemble, elles ont été plutôt bien réussies: plusieurs questions donnent lieu à un taux de réussite de plus de 90%. De plus, la question comportant la plus faible proportion de réponses correctes et partielles est tout de même réussie par près d'un quart des élèves.

Toutefois, la réussite aux questions varie en fonction de certains paramètres, dont le format de questionnement et le type d'opération cognitive à mobiliser.

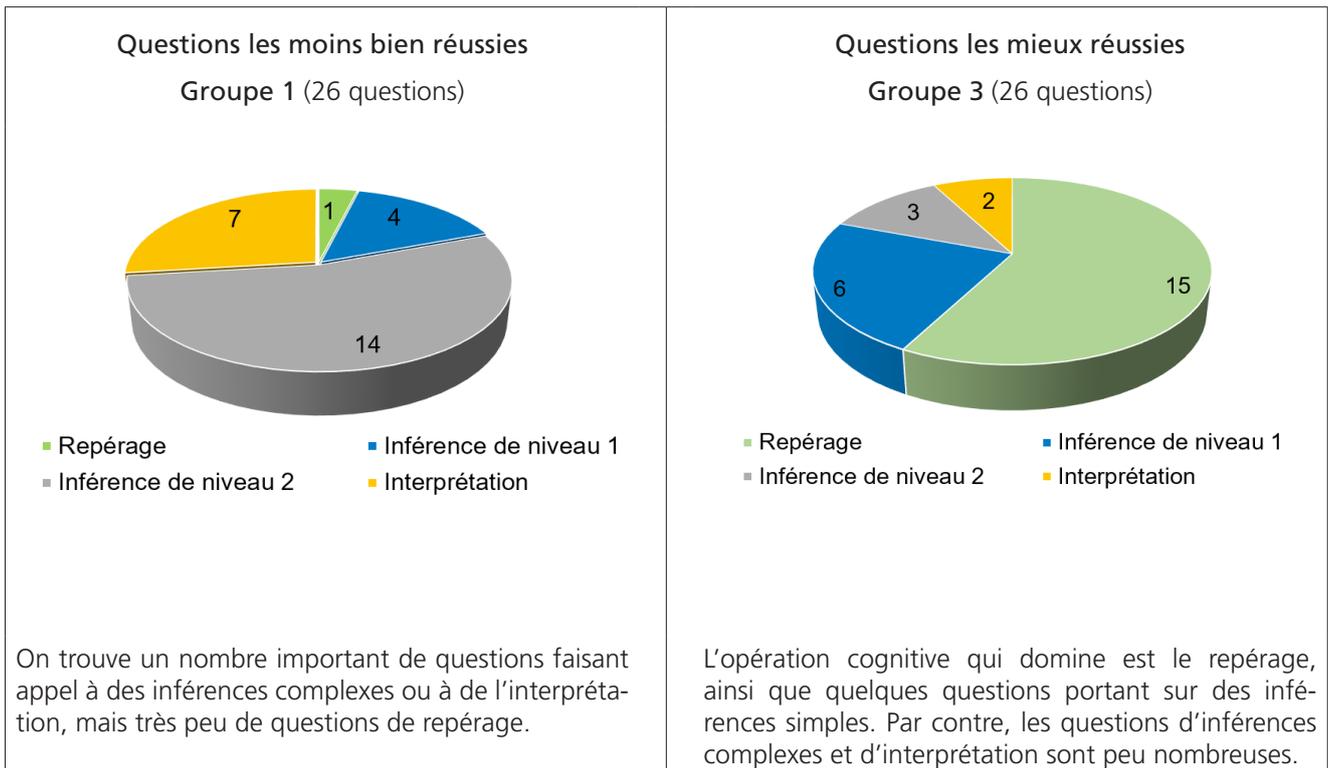
La proportion de non-réponses est également indicative de la difficulté de la question.

Pour quelques questions plus complexes, **les réponses partiellement correctes**, c'est-à-dire ne correspondant qu'en partie aux réponses attendues, ont également été traitées.

Répartition des questions selon le format de questionnement



Répartition des questions selon l'opération cognitive



Ces deux paramètres, formats de questionnement et opérations cognitives, sont liés même si la répartition du format de questionnement n'est pas homogène dans les différents types d'opérations cognitives. Ainsi, un QCM de repérage est mieux réussi qu'un QCM nécessitant une interprétation. De même, une question d'inférence complexe sera mieux réussie si elle est présentée par exemple sous forme de question V/F que sous forme de question d'ordination ou d'appariement.

Le lien avec le PER a également fait l'objet d'une analyse afin de déterminer quelles étaient les *progressions des apprentissages* concernées par les questions. Il s'est cependant avéré plus difficile à les mettre en correspondance avec la réussite. En effet, quand un lien avec certains éléments de *progression des apprentissages* a pu être établi de façon claire, il est apparu qu'un nombre important de questions étaient en relation avec la même *progression*, ne permettant ainsi pas de constituer une gradation dans la difficulté de ces dernières.



© AdobeStock

La résolution de problèmes en Mathématiques

Pour les Mathématiques, le choix s'est porté sur la résolution de problèmes, nécessitant de définir ce qui était entendu par « problème ». 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 à un moment donné, au regard de son âge et des apprentissages réalisés. Ainsi :

Une résolution de problèmes peut être définie comme la convocation autonome, puis l'utilisation adéquate par l'élève d'un ou plusieurs outils mathématiques.

Dans ce cadre, l'élève construit et justifie sa réponse à l'aide de ses connaissances et savoir-faire.

L'objet du test pilote étant la résolution de problèmes, et non uniquement la maîtrise d'algorithmes ou de connaissances mathématiques, l'intérêt de l'expérimentation ne porte pas tant sur la réponse correcte ou non, mais bien sur les « démarches » mises en œuvre par les élèves pour parvenir à résoudre ces problèmes. Les procédures et les outils attendus sont-ils bien mobilisés par les élèves romand.es? Quelles sont les éventuelles sources d'erreurs? Ainsi, la réponse seule n'est pas suffisante pour analyser leurs productions. Comme pour le Français, la focalisation sur les démarches s'inscrit dans la volonté de s'assurer que les tâches conviennent aux élèves de toute la Suisse romande.

Matériel testé auprès d'élèves romand.es

67 problèmes mathématiques issus des épreuves externes cantonales, touchant tous les axes du PER, ont fait l'objet du test pilote. Ils ont été résolus sans l'aide de la calculatrice. Les productions des élèves ont été codées, en prenant en compte :

- l'exactitude de la réponse: correcte, incorrecte, absente et, dans quelques cas, partiellement correcte;
- la démarche de résolution: experte, correcte, non pertinente, absente;

- le type d'erreur: de calcul, de non prise en compte de certaines données de l'énoncé ou de contraintes de la situation, de mathématisation de la situation ou encore d'autres erreurs.

Les **erreurs de mathématisation** correspondent à l'utilisation d'un outil mathématique erroné. Par exemple, l'élève utilise l'addition dans une situation de proportionnalité pour « transformer » une recette prévue pour six personnes en une recette pour huit personnes, l'élève ajoute 2 à chaque ingrédient.

Une « démarche experte » est beaucoup plus fiable pour arriver à la réponse et elle est également plus directe. Une « démarche correcte », parfois préférée par les élèves, car ils et elles s'y sentent plus à l'aise, permet de résoudre le problème posé au prix d'un travail souvent plus laborieux et comportant plus d'étapes. Dans l'exemple ci-dessous, pour trouver le nombre d'entrées minimal pour que l'abonnement proposé à la piscine soit plus avantageux que des entrées individuelles, l'élève A fait une division (une seule opération), tandis que l'élève B recherche cette même information suite à des multiplications lacunaires successives (plusieurs opérations).

A Espace pour ta démarche et tes opérations :

Réponse : Tadeo doit aller à la piscine 38 fois au minimum.

B Espace pour ta démarche et tes opérations :

Réponse : Tadeo doit aller à la piscine 38 fois au minimum.

Premiers résultats

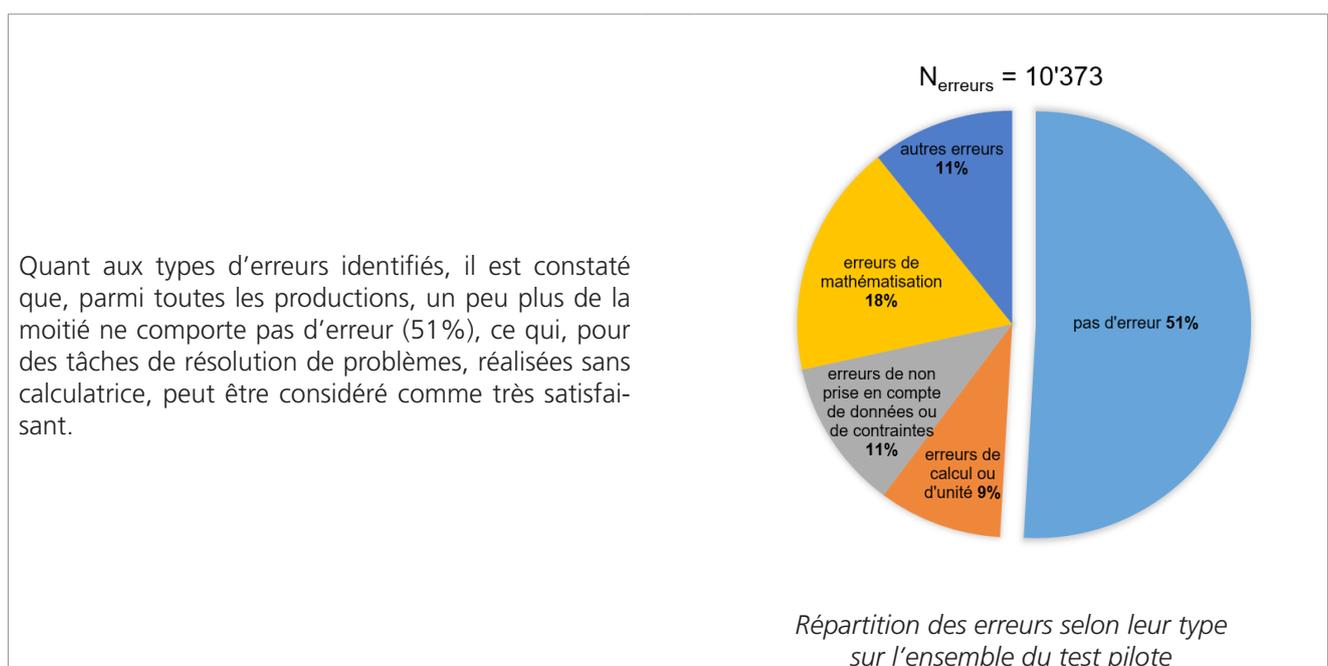
Toujours dans le but de vérifier l'adéquation des tâches aux élèves de Suisse romande, la réussite aux problèmes testés montre, à posteriori, la difficulté très variable de ceux-ci.

Afin de permettre des analyses plus fines, dix regroupements de tâches (par exemple problèmes additifs et soustractifs; problèmes de proportionnalité; figures géométriques planes et solides; repérage dans le plan), ancrés au PER et à ses chapitres, ont été créés. Ils corres-

pondent à la notion mathématique principale à laquelle il est fait appel pour résoudre le problème. Le fait que plusieurs problèmes testent des notions mathématiques similaires est particulièrement intéressant. En effet, une notion comporte plusieurs facettes et ce n'est qu'à travers la mise en œuvre de ces différentes facettes qu'il devient possible de vérifier si l'élève a compris, avec une certaine profondeur, la notion en jeu et peut mobiliser les outils correspondants.



En ce qui concerne les démarches mises en œuvre par les élèves, on constate, dans l'ensemble, que celles considérées comme expertes (55%) sont plus nombreuses que les correctes (7%). Néanmoins, une démarche correcte souvent observée est l'utilisation de multiplications lacunaires successives, certainement pour contourner la division (voir exemple de démarche correcte ci-dessus). Notons que certaines tâches ne peuvent se résoudre par une démarche autre qu'experte (par exemple un problème additif en 8^e année).



Quant aux types d'erreurs identifiés, il est constaté que, parmi toutes les productions, un peu plus de la moitié ne comporte pas d'erreur (51%), ce qui, pour des tâches de résolution de problèmes, réalisées sans calculatrice, peut être considéré comme très satisfaisant.

Conclusion et perspectives

Bilan de la première étape

Le dispositif mis en place participe, grâce à la mutualisation des pratiques évaluatives, au développement d'une culture évaluative commune pour les cantons romands et leur corps enseignant, notamment par un soutien à l'évaluation des compétences définies par le PER et par l'enrichissement d'une *Banque romande d'items*.

Grâce à ce dispositif, ces travaux, qui vont se poursuivre dans le temps, ont permis :

- de mettre en commun des pratiques évaluatives externes de Suisse romande, de les analyser et de les discuter, de sélectionner certaines tâches évaluatives et de les soumettre à des élèves. Le test pilote de 2019 a notamment permis de vérifier l'adéquation d'une série de tâches évaluatives à un échantillon d'élèves de Suisse romande de 8^e année, dans la perspective de les mettre à disposition du corps enseignant ;
- en Français, de mettre en évidence, au niveau du contenu, que les tâches évaluatives en *compréhension de l'écrit* étaient adéquates par rapport à ce qui peut être attendu d'un.e élève de 8^e année. Restent à définir les critères d'analyse permettant de déterminer à partir de quel moment il est possible de dire qu'un.e élève a compris un texte, autrement dit ce qui est nécessaire pour déterminer le niveau de compréhension d'un.e élève ;
- en Mathématiques, de constater que la résolution de problèmes restait globalement difficile pour des élèves romand.es. Toutefois, le test pilote a permis de repérer la variété des démarches mises en œuvre, dont l'analyse est indispensable pour déterminer si une tâche convient aux élèves de tous les cantons romands.

Perspectives

Au-delà de la mise à disposition pour les enseignant.es, au cours de la rentrée 2021, de ressources évaluatives ayant été validées lors du test pilote de 2019, les prochaines étapes envisagées sont les suivantes :

- La poursuite de la récolte et de l'analyse de tâches évaluatives provenant des cantons romands dans d'autres sous-domaines en vue d'étendre la mise à disposition au corps enseignant relative à :
 - la production de l'écrit, la compréhension et la production de l'oral en Français ;
 - des tâches techniques en Mathématiques.
- À plus long terme, étendre le dispositif à d'autres *objectifs d'apprentissage* du PER, à d'autres disciplines, voire à d'autres années scolaires en lien avec les besoins des cantons et des enseignant.es.

Enfin, en écho à de premiers travaux initiés sur le plan romand (Marc & Wirthner, 2012) et au vu des développements des tests de référence suisses ou PISA, une réflexion mériterait d'être menée afin de dégager les avantages et les limites de tâches et de questions sur support numérique.



PistEval – Issues de ces travaux, des pistes pour l'évaluation sont disponibles sur la plateforme du PER, accessibles directement depuis les *objectifs d'apprentissage*. Elles offrent des ressources et des indications pour les utiliser de manière optimale :

>> [pour le Français](#)

>> [pour les Mathématiques](#)

(accessibles dès le 23 septembre 2021)

POUR EN SAVOIR PLUS...

Roth, M., Ruf, I., Sánchez Abchi, V., Soussi, A. & Weiss, L. (2020). *Dispositif romand de mutualisation de tâches évaluatives*. Rapport de travail EpRoCom - banque d'items. Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP).

Sánchez Abchi, V., De Pietro, J.-F. & Roth, M. (2016). *Évaluer en français : comment prendre en compte la difficulté des items et des textes*. Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP).

Marc, V. & Wirthner, M. (2012). *Épreuves romandes communes : de l'analyse des épreuves cantonales à un modèle d'évaluation adapté au PER : rapport final du projet EpRoCom*. Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP).

Julo, J. (2002). *Des apprentissages spécifiques pour la résolution de problèmes ?* *Grand N*, 69, 31-52.