



**CONSEIL SCIENTIFIQUE
DE L'ÉDUCATION NATIONALE**



POURQUOI DES ÉVALUATIONS EN CP-CE1 ?

par Stanislas Dehaene, Président du Conseil scientifique de l'éducation nationale, Liliane-Sprenger-Charolles, Johannes Ziegler, Franck Ramus et Michel Fayol, membres du groupe de travail « Evaluations et interventions » du Conseil scientifique de l'éducation nationale

Un des rôles du Conseil scientifique de l'éducation nationale (CSEN) est de rappeler les acquis des recherches dans le domaine des apprentissages, d'offrir un espace de dialogue où ces acquis sont confrontés au regard des enseignants et, ainsi, de leur permettre d'améliorer leur pratique et leur formation afin que tous les enfants, particulièrement ceux issus de milieu défavorisé, bénéficient des meilleures approches pédagogiques.

Le présent article a pour but de rappeler certains des acquis de la recherche en éducation, d'expliquer pourquoi les évaluations mises en place par la Direction de l'évaluation, de la prospective et de la performance (Depp) du Ministère de l'Education nationale et de la Jeunesse sont utiles, et de répondre aux critiques récentes de certains syndicats enseignants¹.

¹ Voir le dossier de presse de rentrée 2019 du SNIUpp : <https://www.snuipp.fr/system/resources/W1siZiIsIjIwMTkvMDGvMjMvNXhoZ2sxdmtsN19Eb3NzaWVyXzlwMTIzMV9yZV9wYXJ0aWUucGRmIl1d/Dossier%202019%201%C3%A8re%20partie.pdf>

RAPPEL : EVALUER POUR MIEUX IDENTIFIER ET AIDER LES ÉLÈVES EN DIFFICULTÉ

La mise en place des évaluations nationales de CP-CE1 a été motivée par une idée simple : « Evaluer pour mieux Aider » (d'où le nom du dispositif : EvalAide). Ce type d'évaluation, qui est la première phase d'un dispositif de réponse à l'intervention, vise à (1) raccourcir le délai entre la détection d'un besoin de l'élève et l'intervention pédagogique correspondante (2) vérifier si l'élève répond bien à cette intervention pédagogique, ou s'il faut l'intensifier. Ce type de dispositif n'est pas seulement « anglo-saxon », mais il a été adopté par de très nombreux pays (développés et en voie de développement) pour aider le plus rapidement possible les élèves en difficultés, en particulier ceux issus des milieux les moins favorisés. En effet, de nombreux travaux signalent que les inégalités de réussite dues à des facteurs sociaux, qui sont très précoces, s'accroissent dans le temps, ce qui est surtout dû à l'augmentation des exigences en fonction des niveaux scolaires². En l'absence d'aides pédagogiques spécifiques destinées aux élèves les plus fragiles, les inégalités relevées en début de scolarité ne peuvent donc que s'accroître au fil des années.

La nécessité de mettre en place, en France, ce type de dispositif tient surtout à l'évolution négative des résultats de nos élèves en compréhension de la langue écrite et en mathématiques d'après les enquêtes internationales PISA, PIRLS et TIMSS. En effet, de 2000 à 2016, leurs résultats se sont détériorés et la France est devenue progressivement le pays dans lequel l'écart entre les meilleurs élèves et les plus faibles est le plus fort³.

Enfin, comme le notent certains chercheurs français⁴, les différences de performances entre élèves peuvent être masquées par un biais méthodologique : l'absence d'épreuves standardisées au niveau national. En effet, les établissements qui notent le plus sévèrement les élèves sont aussi ceux qui scolarisent les meilleurs, et inversement les établissements qui scolarisent les plus faibles sont souvent les plus indulgents en matière de notation, ce qui conduit à sous-estimer les inégalités d'acquis scolaires. L'objectivité du dispositif EvalAide doit permettre de rétablir une forme d'égalité dans les ambitions pédagogiques.

ORIGINALITÉ DU DISPOSITIF EVALAIDE : UNE DIVERSITÉ DE CONTENUS

L'originalité du dispositif EvalAide réside dans l'examen d'un vaste ensemble de compétences cognitives évaluées chez tous les élèves français en début et en milieu de CP ainsi qu'en début de CE1. Son objectif est de fournir des repères précis sur les acquis et les progrès de chaque élève dans différents domaines du langage et des maths. Le Conseil scientifique de l'éducation nationale a publié en mai 2019 un texte détaillé qui présente le cadre de référence de ce dispositif et les principales contraintes pratiques qui ont présidé à sa conception, tant pour la lecture que pour les mathématiques⁵.

² Baudelot C. & Establet R. (2009). L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales, Seuil ; Broccolichi S. & Sinthon R. (2012). Comment s'articulent les inégalités d'acquisition scolaire et d'orientation ? Relations ignorées et rectifications tardives. *Revue Française de Pédagogie*, 175, 15-38.

³ Voir, pour PISA, Baudelot & Establet (2009), op. cit. et, pour PIRLS, Depp, Note d'Information, 17-24 (2017)

⁴ Par exemple, Baudelot & Establet (2009, op. cit.) et Broccolichi & Sinthon (2012, op. cit.)

⁵ CSEN (2019). EvalAide (Évaluer pour mieux aider) : Un dispositif scientifique de prévention des difficultés en lecture et en mathématiques au CP et au CE1. Disponible sur <https://www.reseau-canope.fr/conseil-scientifique-de-leducation-nationale.html>

Concernant **la lecture**⁶, l'objectif est bien entendu que les enfants parviennent à une bonne compréhension de ce qu'ils lisent. Or, d'après les travaux de recherche, le niveau de compréhension du langage écrit dépend, d'une part, du degré d'automatisation des procédures d'identification des mots écrits, et d'autre part, du niveau de compréhension du langage oral⁷. En conséquence, si l'on veut découvrir les raisons pour lesquelles certains élèves ne comprennent pas bien ce qu'ils lisent, il faut évaluer le degré de maîtrise du décodage (correspondances graphème-phonème et compétences reliées, visuelles et phonologiques), mais aussi de la compréhension du langage oral. C'est ce cadre théorique qui a déterminé le choix des épreuves d'évaluation. Une des caractéristiques majeures d'EvalAide est donc l'étendue des compétences mesurées, contrairement à ce qui est écrit dans le texte du SNIUpp. Ainsi, pour la lecture, en fonction du niveau scolaire, les évaluations portent :

- d'une part, sur la compréhension du langage oral et/ou écrit, du mot au texte en passant par la phrase. Par exemple, sont examinées les capacités de compréhension littérale, la maîtrise des termes spatiaux, celle des anaphores (Luc => ce garçon => il) ou encore la capacité d'effectuer des inférences, c'est-à-dire, de déduire, à partir du contexte, une information implicite.
- d'autre part, sur la maîtrise du principe alphabétique : capacités phonologiques (capacités de segmentation phonémique) et visuo-attentionnelles (ordre et forme des lettres : « lion » versus « loin », « f » versus « t »), et à partir de mi-CP, capacités de décodage (maîtrise des correspondances graphème-phonème et lecture).

Une approche similaire a été adoptée **pour les mathématiques**. Là encore, l'objectif est de faire en sorte que les enfants puissent se servir des nombres et d'autres concepts mathématiques pour résoudre des problèmes nouveaux – un domaine dans lequel les comparaisons internationales montrent à nouveau un déficit prononcé chez les enfants français. La recherche suggère que cette capacité repose sur des fondamentaux que la Depp a choisi de mesurer, car la recherche a démontré leur caractère prédictif d'éventuelles difficultés ultérieures : (1) la connaissance des symboles des nombres ; (2) le passage rapide des symboles à leur sens et vice-versa (dénombrement, comparaison des quantités) ; (3) le concept de ligne numérique, une droite graduée où les nombres occupent des positions successives de la gauche vers la droite ; (4) la résolution de petits problèmes arithmétiques adaptés au niveau de l'élève ; et (5) la maîtrise de certains concepts de base dans le domaine de la géométrie. Là encore, le dispositif se caractérise par la diversité des épreuves proposées.

Quant au choix de démarrer les évaluations dès le début du CP, puis de les reprendre en milieu de CP, il permet aux enseignants d'évaluer les progrès de leurs élèves, donc l'efficacité de leurs interventions pédagogiques et, si nécessaire, de les réviser. Les

⁶ CSEN (2019), op. cit.; Sprenger-Charolles L. & Ziegler J. (2019). Apprendre à lire: contrôle, automatismes et auto-apprentissage. In l'Apprentissage de la lecture, Nathan.

⁷ Gentaz E., Sprenger-Charolles L. & Theurel A. (2015). Differences in the predictors of reading comprehension in first graders from low socio-economic status families with either good or poor decoding skills. PLoS ONE. n°10(3): e0119581 ; Goigoux R., Cebe S. & Pironom J. (2016). Les Facteurs explicatifs des performances en lecture-compréhension en fin de CP. Revue Française de Pédagogie, n°196,67-84. Pour une synthèse : Sprenger-Charolles L., Desrochers A. & Gentaz E. (2018). Apprendre à lire-écrire en français, Langue française 199-3, 51-67, On peut aussi se référer à l'exposé de Franck Ramus : <https://www.youtube.com/watch?v=SfPHLhq9qY>.

évaluations finales de début CE1 permettent de déterminer, pour les élèves qui manifestent des difficultés persistantes en dépit des interventions pédagogiques effectuées, si un bilan et un accompagnement avec l'appui de personnels spécialisés doit être envisagé.

RÉPONSE À CERTAINES CRITIQUES DU DISPOSITIF EVALAIDE

Références « anglosaxonnes » d'un dispositif « controversé »

La recherche est internationale, et ses résultats sont convergents en France et ailleurs. La langue maintenant privilégiée pour les échanges entre chercheurs n'est plus le français, ni le latin, mais l'anglais. Qu'il y ait des débats est une bonne chose, mais il ne faut pas priver nos élèves des avancées majeures qui ont montré leur efficacité dans des expérimentations contrôlées sous prétexte qu'elles proviennent de l'étranger, ou qu'une telle ou qu'un tel refuse leurs conclusions. Nous avons la chance, en France, d'avoir d'excellents enseignants-chercheurs et chercheurs capables de suivre la littérature internationale, et d'y contribuer. Ne nous privons pas de cette richesse intellectuelle et sortons définitivement de ce « divorce partiel, dangereux et pénible, entre le milieu des praticiens et celui des scientifiques » souligné lors d'un colloque sur « Lecture/ Ecriture : Acquisition » qui a eu lieu en...1992.⁸

L'article précité permet surtout souligner que, si depuis les travaux sur la lecture publiés lors de ce colloque de nouvelles découvertes ont été faites (entre autres grâce à l'apport de nouveaux moyens d'investigations)⁹, il y a aussi dans ce domaine des résultats anciens très robustes, en particulier ceux concernant les relations entre compréhension et décodage. Ces résultats, qui ont d'abord observés chez des anglophones et, par la suite, reproduit chez des francophones¹⁰, ne sont malheureusement encore pas bien connus par les enseignants.

Importance accordée à la fluence en lecture

La fluence en lecture désigne la précision et la rapidité avec laquelle un lecteur parvient à identifier les mots écrits. Selon le SNUIPP, « les défenseurs » de la fluence « peinent ... à convaincre qu'elle serait la voie royale pour construire la compréhension en lecture » (p.11). La fluence est pourtant un des meilleurs prédicteurs de la compréhension en lecture au début de cet apprentissage, y compris d'après l'étude du seul chercheur cité par le SNUIPP¹¹. C'est la raison pour laquelle, au Royaume Uni, le *phonics check*, dispositif équivalent à EvalAide et mis en place dès 2012, ne comprend qu'une seule épreuve de fluence en lecture de mots et de pseudomots. Cette épreuve est passée par tous les élèves

⁸ Colloque financé par le Ministère de l'éducation, alors dirigé par Jack Lang. Voir Morais J. (1993). Compréhension / décodage et acquisition de la lecture, in *Lecture / écriture : acquisition (Actes de la Villette)*, Paris, Nathan, 10-21.

⁹ Voir les travaux de S.Dehaene et L.Cohen sur l'importance d'une zone cérébrale spécifique (la zone de la forme visuelle des mots) dans la lecture et son apprentissage : cf. Kolinsky R., Morais J., Cohen L. & Dehaene S. (2018). Les bases neurales de l'apprentissage de la lecture, *Langue Française*, 199, 17-33

¹⁰ Sur les relations entre compréhension et décodage, Voir, en plus de l'article de Morais dans les Actes du Colloque de la Villette et des publications récentes citées dans la note 7 (Gentaz et al., 2015 ; Goigoux et al., 2016), les synthèses diffusées par l'Observation National de la Lecture (par exemple, celle publiée en 2004 par Fayol M. et Morais J. : <http://onl.inrp.fr/ONL/publications/publi2004/evolution/>). Voir également les recommandations actuelles du Conseil Scientifique de l'Education Nationale

¹¹ Goigoux et al. (2016), op. cit.

à la fin de la première année¹². Des exercices individualisés sont proposés à ceux qui n'arrivent pas à lire correctement la plupart des mots présentés. Ces élèves repassent la même épreuve un an plus tard pour s'assurer qu'ils ont comblé leurs lacunes. Dans le cas contraire, des exercices plus ciblés leur sont proposés. Ce dispositif a fait ses preuves. En effet, le niveau des élèves anglais sur ce test a augmenté d'année en année et surtout, d'après PIRLS, leurs scores en compréhension de la langue écrite se sont améliorés entre 2011 et 2016 (+7 points : de 552 à 559). Pendant la même période, ceux des élèves français se sont détériorés (-9 points : de 520 à 511).

Il importe donc que, en France, comme maintenant dans de très nombreux autres pays, les enseignants disposent de normes de fluence en lecture de mots isolés, et puissent identifier rapidement les enfants qui éprouvent des difficultés.

Choix des seuils d'alerte

Les tests proposés fournissent des mesures quantitatives qui permettent de positionner chaque élève sur un continuum. En outre, pour chaque test, la Depp, en consultation avec le CSEN, a proposé deux niveaux d'alerte : un seuil bas en deçà duquel les élèves ont un besoin immédiat d'intervention, et un seuil moyen qui définit les élèves dont les acquis sont fragiles. Au-delà de ce dernier niveau, on considère que les compétences sont suffisantes pour permettre aux élèves d'entrer sereinement dans les apprentissages. Les seuils ont été choisis par la Depp après consultation d'enseignants de CP et CE1, des conseillers pédagogiques, des inspecteurs de l'éducation nationale, et des chercheurs, étant donné les attendus. Toute classification de ce type conserve bien entendu une part d'arbitraire, mais les enseignants ont également à leur disposition les résultats détaillés de leurs élèves qui leur permettent de porter un diagnostic plus fin.

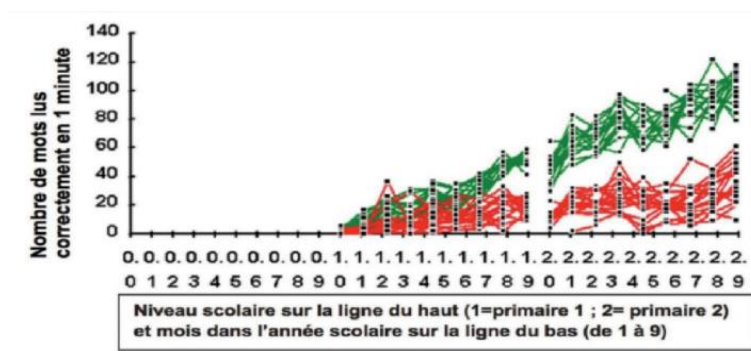
L'une des mesures les plus intéressantes est le nombre de mots lus correctement en une minute – une mesure simple et rapide de la fluence et donc de l'automatisation de la lecture. De nombreux enseignants s'en servent déjà pour vérifier régulièrement les progrès des élèves. Les seuils Depp sont placés respectivement à 16 et 30 mots par minute en début de CE1, ce qui en 2018 a conduit à identifier 12% d'élèves « à besoin » et 19,7% d'élèves « fragiles ». Ces chiffres n'ont rien de choquant, sachant qu'il s'agit non pas d'étiqueter, mais d'épauler ces élèves par des interventions pédagogiques.

Il est légitime pour un système éducatif de se fixer des objectifs élevés (à condition bien sûr de s'en donner les moyens). Dans ses attendus de fin de CP, le ministère souhaite que les élèves atteignent 50 mots par minute. La barre est haute, et l'objectif ne sera probablement pas atteint dans l'immédiat, mais elle n'est pas excessive puisque 34,1 % des élèves l'ont dépassée en 2018 (et 26,4% des élèves en REP).

Dans les années à venir, les résultats des évaluations nationales permettront progressivement d'affiner ces seuils en vérifiant leur corrélation avec la réussite ultérieure des enfants. A titre d'exemple, le graphe présenté ci-dessous montre l'évolution des scores en lecture entre la 1^{ère} et la 2^{nde} année du primaire d'élèves anglophones plus ou moins bons lecteurs. Ceux qui, au début de l'étude, lisaient plus de 40 mots en une minute (en vert),

¹². National curriculum (2013), English programs of study: key stages 1 and 2.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/335186/PRIMARY_national_curriculum_-_English_220714.pdf

progressent rapidement, tandis que ceux qui en lisaient moins de 40¹³ ne progressent guère en seconde année : l'écart se creuse massivement avec le temps. L'objectif des évaluations est de détecter ces difficultés et d'intervenir avant qu'elles ne soient définitivement installées.



Difficultés de compréhension fine des textes

Selon le SNUipp, puisque les difficultés des élèves français se manifestent plus fortement sur des questions impliquant une compréhension fine de ce qu'ils lisent, elles ne peuvent pas s'expliquer par des problèmes de décodage. En réalité, nous n'en savons rien, car ni dans PIRLS, ni dans PISA, les capacités de décodage des élèves n'ont été évaluées, pas plus que leurs capacités de compréhension de la langue orale. Ce sont deux limites de PIRLS et PISA qui doivent être prises en considération pour interpréter les résultats.

Dans la mesure où l'identification des mots est une condition nécessaire de la compréhension en lecture, il est plausible que la piètre performance des élèves français en compréhension résulte, au moins en partie, de la difficulté qu'éprouvent certains dans le décodage et la fluence de lecture. Bien entendu, ces résultats invitent aussi les enseignants à renforcer l'enseignement explicite des stratégies de compréhension de texte, un domaine qui est probablement négligé. En effet, les enseignants français disent ne pratiquer que rarement certains exercices permettant une compréhension approfondie des textes. En témoignent leurs réponses à la question « A quelle fréquence demandez-vous aux élèves de faire les choses suivantes pour les aider à développer leurs aptitudes ou leurs stratégies de compréhension de l'écrit » (voir le tableau ci-dessous) :

¹³ Good R.H., Simmons D.C. & Smith S.B. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review*, 27(1), 45-56.

Depp, Note d'information 17-24, déc 2017	Moyenne France	Pays de l'UE (21) avec un score supérieur à celui de la France	Ecart
Retrouver des informations dans un texte	99	96	+3
Expliquer ou argumenter pour montrer ce qu'ils ont compris	91	94	-3
Dégager les idées principales du texte	89	93	-4
Généraliser ou élaborer des inférences à partir du texte	64	79	-15
Prévoir ce qui va se passer dans la suite du texte	59	72	-13
Comparer le texte à des lectures antérieures	50	69	-19
Comparer ce qu'ils ont lu à des faits qu'ils ont vécus	41	82	-41
Décrire le style ou la structure du texte	41	60	-19
Déterminer la perspective ou les intentions de l'auteur	35	56	-20

Le Conseil scientifique de l'éducation nationale souligne la nécessité d'améliorer ces points en approfondissant la compréhension des textes, d'abord à l'oral (dès la maternelle), puis à l'écrit, une fois que la lecture est devenue fluente. La compréhension fait l'objet de recommandations détaillées dans le document « Pédagogies et manuels pour l'apprentissage de la lecture : Comment choisir ? » publié par le CSEN en mai 2019.¹⁴ Cependant, ce serait une erreur d'opposer décodage et compréhension. La recherche montre en réalité que les deux dimensions sont essentielles et doivent être toutes deux travaillées.¹⁵

Liberté pédagogique

Les évaluations nationales donnent aux enseignants des informations fiables sur leurs élèves, et les invitent à mettre en œuvre des interventions pédagogiques pour ceux en difficulté. Dans ce but, la Dgesco met à leur disposition un éventail de ressources pédagogiques. Les enseignants sont entièrement libres de choisir parmi ces ressources celles qui leur conviennent (et conviennent à leurs élèves), d'en utiliser d'autres, ou aucune. Ils sont simplement incités à intervenir de façon ciblée par tous les moyens qui leur paraissent appropriés. Par conséquent, ce dispositif respecte intégralement la liberté pédagogique des enseignants.

¹⁴ Disponible sur <https://www.reseau-canope.fr/conseil-scientifique-de-leducation-nationale/outils-pedagogiques.html>

¹⁵ Bianco M., Pellenq C., Lambert E., Bressoux P., Lima L., & Doyen A.-L. (2012). Impact of early code-skill and oral-comprehension training on reading achievement in first grade. *Journal of Research in Reading*, 35(4), 427–455. Voir aussi les synthèses citées dans la note ¹⁰

POUR CONCLURE

Les évaluations nationales ne s'arrêtent pas à la mesure d'une seule compétence, comme la vitesse de lecture, mais incluent une diversité de compétences dont la maîtrise prédit fortement la réussite des élèves en lecture et en mathématiques. Ces évaluations devraient permettre aux enseignants d'avoir une photographie très précise des compétences à renforcer et leur donner des hypothèses éclairées concernant l'origine des difficultés éprouvées par certains de leurs élèves.

Indubitablement, le dispositif EvalAide peut et doit encore être amélioré. Au fil des années, les statistiques montreront probablement que certains tests sont redondants et peuvent être supprimés. A la rentrée 2019, certaines épreuves ont été enlevées, raccourcies ou simplifiées. D'autres manquent sans doute, notamment sur l'évaluation de l'écriture et de l'orthographe. Nous souhaitons vivement que les critiques des enseignants conduisent à affiner le dispositif. Mais nous espérons également que les explications présentées dans cet article ainsi que dans le texte « EvalAide » émanant du CSEN permettront de lever les résistances de certains enseignants. Grâce à ce dispositif d'évaluation, couplé à des interventions pédagogiques immédiates et ciblées, la France devrait cesser d'être le « pays du grand écart »¹⁶ et parvenir à réduire enfin les inégalités sociales entre élèves.

¹⁶ Baudelot & Establet (2009). op. cit.