

FAIRE MENTIR LES CHIFFRES, EN PÉDAGOGIE AUSSI

« Les premiers résultats sont là »

Le 7 mai 2019, en conclusion du Grand débat, Jean-Michel Blanquer, ministre de l'Éducation nationale et de la Jeunesse, a écrit à tous les professeurs des écoles pour saluer le travail accompli sous sa houlette : « *les premiers résultats sont là* ». Pour étayer son affirmation, il s'appuie sur les résultats des évaluations CP et CE1 publiés quelques jours plus tôt¹ par la DEPP (Direction de l'évaluation, de la prospective et de la performance, MEN). Dans le domaine de la lecture, un progrès spectaculaire est mis en valeur pour administrer la preuve de son succès : « *si, en début de CP, 23 % des élèves n'identifiaient que la moitié des lettres et des sons qui leur étaient soumis, ils ne sont plus que 3,3 % au mois de janvier* ».

Les chiffres sont exacts. La DEPP relève que, au mois de septembre 2018, 23 % des élèves ont échoué au test intitulé « Connaitre le nom des lettres et le son qu'elles produisent ». Dans cette épreuve, à dix reprises, les enfants devaient entourer, parmi une suite de cinq lettres imprimées (par exemple « p b d a q »), celle qui correspondait au son qu'ils entendaient au début d'un mot monosyllabique prononcé par le maître (p. ex. « bulle »). Autrement dit, ils devaient être capables de discriminer un phonème en position initiale (« bulle » commence par le son /b/) puis de sélectionner la lettre correspondant au phonème qu'ils venaient d'isoler. Ceci impliquait de combiner une connaissance de la valeur sonore des lettres (« le B fait Beu ») et une habileté phonologique complexe, généralement hors de portée des enfants à qui on n'a pas encore appris à déchiffrer². Est-ce que, pour autant, cet échec signifie que les élèves sont « en difficulté » comme l'écrit la DEPP ?

Si l'on consulte le décret qui fixe les objectifs de l'école maternelle, on peut constater que les compétences ainsi évaluées ne sont pas au programme. Quoi de plus normal, par conséquent, qu'à la rentrée bon nombre d'élèves ne disposent pas de connaissances qui ne leur ont pas été enseignées ? Et quoi de plus normal, quatre mois plus tard, qu'ils sachent ce qu'on leur a appris ? En d'autres termes, en quoi le résultat de janvier peut-il être considéré comme la preuve que « *l'École de France sait être réactive et déterminée pour se placer aux avant-postes des politiques sociales de notre pays* », c'est-à-dire de la politique sociale du gouvernement ?

Cet argument, élaboré par les services de communication du ministère et déjà relayé par plusieurs DASEN (« *le nombre d'élèves en difficulté baisse de 20 points* » osent-ils écrire dans les mails adressés aux écoles), est mensonger. Ne prendrait-on pas les professeurs des écoles (et peut-être aussi les journalistes) pour des crétins ?

Comme nous ne pouvions pas nous résigner à le croire, le ministre ayant fait précéder sa signature de la mention : « *Avec toute ma confiance* », nous avons entrepris d'examiner de plus près les publications ministérielles. Dans l'article qui suit nous allons présenter le résultat de nos investigations à la recherche des mécanismes de fabrication de « bobards » tels que celui-ci. Le dernier, dévoilé en conclusion, ne sera pas le moindre.

¹ <https://www.education.gouv.fr/cid141364/evaluations-2018-2019-en-cp-et-ce1-des-premiers-resultats-encourageants.html>

² José Morais (1984/1999), *L'art de lire*, éditions Odile Jacob. Morais est un cognitiviste reconnu par le Conseil Scientifique de l'Éducation Nationale (CSEN).

PREMIERE PARTIE : LES ÉVALUATIONS NATIONALES

L'évaluation : un cheval de Troie

✚ La clé de voute du dispositif d'évaluation : « la réponse à l'intervention »

Les trois notes d'information publiées par la DEPP officialisent ce que la DGESCO avait laissé dans le flou jusqu'à présent, suscitant la plus grande incompréhension chez les enseignants de cycle 2 : les évaluations CP-CE1 n'étaient que la première partie d'un dispositif plus vaste appelé « la réponse à l'intervention » (**RAI**). Seule une connaissance de cette opération aurait permis aux maîtres de comprendre que les choix des tests puis des seuils de réussite avaient été déterminés par les modalités de remédiation préalablement arrêtées.

La RAI est la version francisée, par les québécois³, d'un dispositif américain appelé « *Tiered approach* », c'est-à-dire une approche de prévention par étages (ou par niveaux)⁴. L'objectif est d'intervenir le plus tôt possible pour prévenir les difficultés d'apprentissage des élèves, sans attendre que l'échec s'installe et sans se soucier, dans un premier temps, des causes de ces difficultés. Il suffit de les identifier en définissant des seuils d'alerte lorsque le niveau attendu n'est pas atteint.

- Au **premier niveau** et dans un premier temps, tous les élèves bénéficient d'un enseignement collectif et identique fondé, dans la mesure du possible, sur les recommandations issues de la recherche en éducation et/ou de leur employeur. En France, cela signifie que les maîtres doivent faire classe en respectant scrupuleusement les consignes du guide orange.
- Au **deuxième niveau** et dans un second temps, les élèves signalés « en difficulté » à l'issue de tests standardisés bénéficient d'un enseignement supplémentaire dispensé par l'enseignant titulaire de la classe. On considère que ces élèves n'ont pas suffisamment « répondu à l'intervention », au sens où ils n'ont pas bénéficié comme on l'attendait de l'enseignement dispensé en collectif.
- Au **troisième niveau** et dans un troisième temps⁵, les élèves qui n'ont pas réalisé les progrès escomptés aux deux premiers niveaux sont destinataires d'interventions complémentaires, souvent délivrées par des enseignants spécialisés ou d'autres professionnels de la rééducation⁶.

Dans cette approche du traitement des difficultés d'apprentissage, les enseignants sont invités à se référer à des normes nationales, et non à leurs évaluations personnelles réalisées au cours des activités ordinaires d'enseignement. Fixer ces normes et comparer les élèves sont donc les premières fonctions des évaluations ministérielles CP-CE1.

³ https://savoirs.usherbrooke.ca/bitstream/handle/11143/10274/015_SILE2015_Desrochers_Laplante_Brodeur.pdf?sequence=3

⁴ Sugai, G., & Horner, R. H. (2009). Responsiveness-to-intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality*, 17(4), 223-237.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe learning disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, 34, 33-58.

⁵ Parfois simultanément ou à la place du deuxième.

⁶ Franck Ramus, par exemple, considère que les élèves qui ne progressent pas au milieu du CE1 alors qu'ils bénéficient d'une intervention de niveau 3 devraient être orientés vers des bilans médicaux, notamment pour diagnostiquer d'éventuelles dyslexies.

De l'évaluation à la remédiation : enseigner ce qui est évaluable

Dans le dispositif RAI, les enseignants doivent s'appuyer sur les résultats des évaluations pour 1° identifier les élèves à risque de difficultés d'apprentissage (ou de troubles d'apprentissage au troisième niveau) et 2° adapter l'intensité et la nature des remédiations « à la réponse des élèves aux interventions effectuées⁷ ».

Le dispositif promeut donc une individualisation très exigeante de la prise en charge pédagogique. Les évaluations répétées (trois fois par an) servent à évaluer les progrès réalisés et à ajuster le dispositif de remédiation en conséquence (« *analyser item par item l'évolution des élèves, tout particulièrement lorsque des items sont communs aux deux temps d'évaluation* », écrit la DEPP). Elles doivent aussi et surtout inciter les enseignants à renforcer leurs interventions sur les compétences jugées déficitaires.

Le choix des tests joue donc un rôle primordial dans l'élaboration des tâches de remédiation et, plus largement, des tâches d'enseignement. Une grande partie des protestations des enseignants, des formateurs et des chercheurs vient de là. Le ministère ayant minoré ou carrément ignoré des pans entiers des apprentissages scolaires du français (par exemple, la calligraphie, la copie ou la production de textes ; la compréhension de textes entendus mi-CP...), les maîtres sont incités à en faire autant. Cela s'appelle piloter l'enseignement par l'évaluation ; ce qui n'est pas évalué perd sa valeur, les maîtres sont incités à enseigner l'évaluable.

Si l'intention de la RAI – intervenir avant que l'échec s'installe – est louable, divers aspects de son efficacité à court et moyen terme et de sa mise en œuvre sont discutés sur le plan international⁸. Plusieurs conditions doivent être réunies pour qu'elle ait une chance d'être bénéfique : 1) la qualité et la pertinence des tests choisis. 2) L'adhésion des enseignants au dispositif et sa possible intégration dans leur culture professionnelle : les impositions autoritaires échouent. Sur ces deux premiers points, l'opération française conduite en catimini et sans concertation est très mal engagée. 3) La possibilité de dégager du temps supplémentaire pour les élèves en difficulté sans que cela se fasse au détriment des autres enseignements au programme et sans que leurs camarades soient pénalisés. 4) La qualité des outils de remédiation et leur capacité à couvrir l'ensemble des compétences au programme⁹. 5) La complémentarité des informations apportées par les tests avec celles dont disposent déjà les enseignants qui travaillent au quotidien avec les élèves. Aucune de ces trois dernières conditions n'est remplie en France pour le moment, pas plus que n'est réglée la question de la pertinence des seuils choisis par la DEPP.

La DEPP a défini trois sous-groupes d'élèves pour chaque domaine de compétences évaluées : les élèves « à besoins » (en dessous du seuil n° 1), les élèves « fragiles » (entre le seuil n° 1 et le seuil n° 2), les élèves sans difficultés (au-dessus du seuil n° 2). On notera que, dans ses publications, elle groupe les deux derniers sous-groupes, situés en dessous du seuil n° 2, sous le vocable **d'élèves « en difficulté »**¹⁰. Ceci indique que le MEN a choisi, dans un premier temps, de privilégier le niveau 2 de la *réponse à l'intervention*

⁷ Cf. note 3

⁸ Note d'habilitation à diriger les recherches de Caroline Viriot-Goedel : *Aider l'apprenti-lecteur en difficulté. Analyse du cas français à la lumière de la Réponse à l'intervention*. Université de Paris 8, juin 2017. Garant : Pr. Jacques Crinon

⁹ Haager, D. E., Klingner, J. E., & Vaughn, S. E. (2007). *Evidence-based reading practices for response to intervention*. Paul H Brookes Publishing.

¹⁰ La notion de « difficulté » ou de « risque de difficulté » repose donc sur le choix d'épreuves testant des compétences prédictives de la réussite et sur l'établissement de seuils pertinents. https://cache.media.education.gouv.fr/file/2019/96/1/depp-ni-2019-19-13-Evaluations-reperes-2018-debut-CP-premiers-resultats_1114961.pdf

en groupant les élèves « à besoin » et « fragiles ». Il ne sait pas encore comment organiser le niveau 3. La circulaire de rentrée indique que, l'année prochaine, les enseignants devront proposer des dispositifs de remédiation aux élèves « en difficulté » (seuils 1 et 2 confondus) et rendre compte – à leur hiérarchie et aux parents – des progrès accomplis.

Jusqu'à présent, le déploiement de cette nouvelle politique n'a pas été discuté ni expliqué, pas plus aux IEN qu'aux enseignants qui ont dû faire passer des évaluations sans la comprendre. Ce n'était pas vraiment un secret, mais l'information n'a pas été diffusée. Pour la connaître plus tôt, il fallait assister aux colloques du SNUipp¹¹ ou consulter les blogs personnels de certains membres du CSEN¹². On pourrait en déduire que les évaluations CP-CE1 jouent le rôle d'un cheval de Troie pour pénétrer dans l'école de la défiance par une porte dérobée.

Des tests contestables

Masquer les échecs en abaissant les seuils

À la rentrée 2018, Les enseignants de CP ont beaucoup protesté contre les épreuves qu'ils jugeaient inadaptées et qui plaçaient beaucoup de leurs élèves en situation d'échec dès les premiers jours d'école. L'analyse des résultats leur donne entièrement raison. Le MEN saura-t-il le reconnaître ?

Les tests choisis étaient ceux que les psychologues cognitivistes utilisent pour leurs recherches fondamentales : en laboratoire, ils sont proposés à des élèves plus âgés et administrés en tête-à-tête, de manière à encourager les enfants et à pouvoir les interrompre en cours de passation après plusieurs échecs consécutifs. Ils n'ont pas pour but d'évaluer les compétences définies par les programmes scolaires et ne couvrent qu'une toute petite partie des habiletés enseignées à l'école¹³.

Ces tests ont ensuite été transposés pour des passations collectives (sur livret) mal adaptées à des enfants sortant de l'école maternelle : les problèmes de repérage dans l'espace graphique de la page en temps limité expliquent bien des échecs. Certains tests étaient si mal conçus ou si inappropriés à l'âge des enfants que la DEPP a dû **établir des seuils extrêmement bas** pour éviter que cela soit trop visible et que trop d'élèves soient rangés dans la catégorie « en difficulté » (cf. tableau 1).

¹¹ Conférence prononcée au colloque du SNUipp le 28 novembre 2018 à Paris. Vidéo en ligne : <https://www.facebook.com/snuipp/videos/200541127498699/UzpfSTeZMDQ3NTc4Mzk6MzA2MDYxMTI5NDk5NDE00jc1OjE1MTQ3OTM2MDA6MTU0NjMyOTU5OTotMzQ1NzU4ODQ1NjAzNTAyMzU3Mw/?id=1304757839>

¹² Cf. le blog de Franck Ramus <http://www.scilogs.fr/ramus-meninges/approche-reponse-a-lintervention-difficultes-lecture/>

¹³ Les tests ont aussi été choisis en fonction des préoccupations scientifiques et des objets d'étude de membres du conseil scientifique de l'Education nationale ; ce qui explique par exemple l'étonnante épreuve de comparaison de suites de lettres chère à Johannes Ziegler. Ce qui explique aussi les impasses faites sur des sujets non étudiés par des membres du CSEN

Tableau 1 : définition des seuils pour chaque épreuve et pourcentage d'élèves les ayant atteints au début du CP

L = D x C	Compétences	Tests	Score sur	Seuil 1 "Besoin"	% d'élèves	Seuil 2 "Fragile"	% d'élèves
D	Reconnaissance des lettres et phonographie	Comparer des suites de lettres	24	≤ 3	5,5	≤ 8	13,1
		Reconnaître des lettres parmi des signes	4	≤ 1	8,4	≤ 3	28,3
		Reconnaître les différentes écritures d'une lettre	7	≤ 2	14,3	≤ 4	24,6
		Connaître le nom des lettres et le son qu'elles produisent	10	≤ 2	6,8	≤ 5	16,2
	Phonologie	Manipuler des phonèmes	15	≤ 2	5,4	≤ 5	16,5
		Manipuler des syllabes	15	≤ 5	6,2	≤ 8	12,8
C	Compréhension orale (ou plutôt d'écrits oralisés)	Comprendre des mots lus	15	≤ 6	5,7	≤ 10	20,9
		Comprendre des phrases lues	14	≤ 6	3,5	≤ 10	14,4
		Comprendre des textes lus	18	≤ 4	3,2	≤ 7	6,7

Source : Roland Goigoux à partir des données éparées de la DEPP (barèmes d'une part et résultats d'autre part)

Trois exemples sont, selon nous, très significatifs :

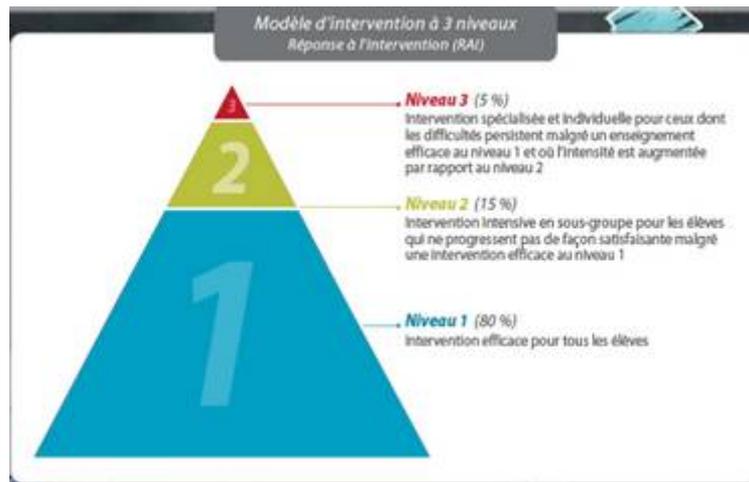
1. l'exercice de comparaison de suites de lettres était si difficile que la DEPP a fixé le seuil minimal à **3 réussites seulement sur 24 items**. Grâce à ce procédé, elle a pu réduire le nombre d'élèves « à besoin » (seuil n° 1) à 5,5 % de la population et à 13,1% la proportion de « fragiles » (seuil n° 2 : 8 réussites sur 24 suffisaient).
2. De la même manière, le seuil n° 1 de l'épreuve de manipulation de phonèmes a été fixé à **2 réussites sur 15 items** (5,4 % d'élèves « à besoin ») pour masquer son inadéquation aux compétences attendues à la fin de l'école maternelle. Au total, 23 % des élèves sont jugés « en difficulté » lorsqu'ils réussissent moins de 6 items sur 15. Cette épreuve, remplie de pièges sémantiques et phonologiques, visait elle aussi à évaluer la maîtrise de compétences dépassant largement les attendus du programme. Seuls les élèves déjà quasi-lecteurs pouvaient la réussir entièrement.
3. Inversement, pour ne pas avoir à reconnaître que les élèves avaient besoin d'un enseignement de la compréhension de récits à l'école maternelle et au CP, le ministère a **artificiellement fait baisser les seuils de réussite** dans l'épreuve de compréhension de quatre textes entendus au début du CP. En considérant qu'il suffisait de répondre correctement à 7 questions sur 18 pour passer au-dessus du seuil n° 2, le ministère a pu conclure que cette épreuve était réussie par plus de 90 % des élèves. Inutile d'en faire une priorité d'enseignement à l'école maternelle, ni de remédiation au cours préparatoire, puisque seuls 9,8 % d'élèves étaient en difficulté ! CQFD.

✚ La définition des seuils de réussite

On l'aura compris, la définition du public d'élèves relevant des divers dispositifs de remédiation dépend de la définition des seuils de réussite attendue. Les enseignants doivent être informés que la définition de ces seuils n'a rien de scientifique : elle est le fruit de tâtonnements réalisés par la DEPP avec l'aide de quelques professionnels de terrain en fonction de consignes politiques. Pour chaque test, il faut trouver où placer le curseur pour qu'il n'y ait ni trop, ni trop peu, d'élèves signalés en difficulté. C'est pour cette

raison que ces seuils ne peuvent pas être communiqués aux enseignants avant l'analyse des résultats. C'est le ministère, pas la science, qui détermine après coup (une fois les résultats recueillis) le pourcentage d'élèves à signaler aux enseignants. Et qui gère sa communication politique sur la base de chiffres apparemment indiscutables : au début du quinquennat, tout allait mal mais l'amélioration est en marche.

Ce que nous venons de montrer à l'entrée au CP est également vrai mi-CP et au début du CE1. Au CE1 par exemple, le ministère voulait pouvoir identifier une population « en difficulté » avoisinant les 20 % d'une classe d'âge comme c'est le cas dans les autres dispositifs de RAI à l'étranger¹⁴.



Pour y parvenir (cf. infra tableau n° 3), il fallait que les élèves connaissent au moins **12 mots sur 15** au test lexical (22 % des élèves au-dessous du seuil n° 2), mais répondent seulement à **3 questions sur 9** à l'épreuve de lecture-compréhension (lecture autonome du texte et des questions ; 22 % des élèves au-dessous du seuil n° 2). Les taux de réussite pris en compte sont donc très différents d'une épreuve à l'autre (ici 4/5 et 1/3) selon les conclusions qu'on veut en tirer. En lecture par exemple, exiger la réussite à 4 questions sur 9 modifierait considérablement la proportion d'élèves « en difficulté ».

La communication ministérielle, ornée de tous les oripeaux statistiques imaginables, reste donc avant tout une communication politique : **il suffit de jouer avec les taux de réussite attendus pour influencer les pourcentages d'élèves en difficulté**. Lorsque le ministre affirme devant la presse qu'un quart des élèves est en difficulté, cela signifie simplement qu'il a choisi de juger « en difficulté » les 25 % des élèves qui ont obtenu les scores les plus faibles à une épreuve (non étalonnée). Sont-ils pour autant « en difficulté » ? Parfois oui, parfois non. Il ne s'agit pas de nier la difficulté scolaire mais de s'assurer de ses contours afin de lui apporter une réponse appropriée.

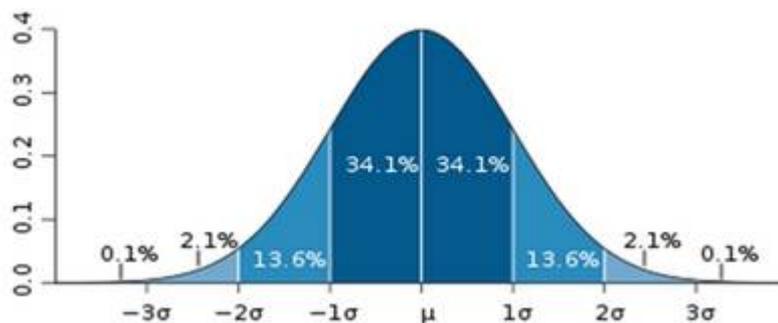
Cette méthode permet d'identifier les élèves les plus faibles de n'importe quel groupe : à Polytechnique et à l'ENA aussi, on peut isoler le quart des étudiants les moins performants. Mais peut-on affirmer qu'ils sont en difficulté ?

¹⁴ Le schéma placé ci-dessous a été emprunté à la commission pédagogique de La Jonquière, Québec.
<https://www.csjonquiere.qc.ca/donnees/media/1/fichiers/presentation%20EHDA.pdf>

Trois solutions pour établir rigoureusement des seuils

Comment résout-on habituellement la question des seuils ? Les systèmes scolaires ont le choix entre trois solutions.

1. La plupart des recherches en éducation et en sciences cognitives choisissent d'identifier les élèves « en difficulté » par comparaison avec l'ensemble des élèves. La population de référence est caractérisée par sa moyenne et son écart-type, les élèves jugés faibles (p. ex. « faibles décodeurs » ou « faibles compreneurs »)¹⁵, sont ceux dont les performances sont inférieures à un écart-type sous la moyenne. Cette borne établie après analyse statistique des performances obtenues est le seuil recherché. Celui-ci permet d'identifier les élèves les moins performants au test soit environ 15 % de la population de référence. Il est donc construit **à postériori** mais il peut l'être avec la même rigueur pour tous les tests quelle que soit leur nature.



Dans cette distribution normale, μ est la moyenne et σ l'écart-type.
Un peu plus de 15 % de la population se situe au-dessous du seuil de -1σ

L'intérêt de cette technique est sa standardisation scientifique et le consensus qui l'accompagne. Elle permet aussi les comparaisons entre études et par conséquent le cumul des connaissances. Les systèmes éducatifs peuvent choisir d'autres seuils standardisés (p. ex. $-0,8\sigma$) pour identifier une proportion plus importante d'élèves (p. ex. 20 % d'une classe d'âge) et s'y tenir pour toutes les épreuves. Ils peuvent alors examiner ce que savent (ou savent faire) les élèves les moins performants et se demander si leur résultat est inquiétant ou pas.

Son principal inconvénient, pour notre ministère, est qu'on obtient toujours la même proportion d'élèves « en difficulté » (environ 15 %) et que cela ne permet pas de communiquer sur les effets d'une politique. Pour cela, il faudrait décrire de quoi sont capables les 15 % des élèves les plus faibles à différentes échéances successives et montrer qu'ils ont progressé.

2. La seconde solution consiste à définir les seuils **à priori**, en fonction des exigences du système éducatif. Décider, par exemple, qu'à la fin du cours préparatoire, les élèves doivent être capables de lire à haute voix 30 mots par minute. Une fois ce seuil fixé, on peut calculer le pourcentage d'élèves qui ne l'atteignent pas et les déclarer « en difficulté ». L'avantage est que le procédé est transparent et qu'un débat peut s'instaurer avec les professionnels (et les citoyens) sur la définition des compétences requises aux différents paliers de la scolarité.

¹⁵ P.ex. : Goigoux, R., Cèbe, S. & Pironom, J. (2016). Les facteurs explicatifs des performances en lecture-compréhension à la fin du cours préparatoire. *Revue française de pédagogie*, (3), 67-84, <https://journals.openedition.org/rfp/5076>

Son inconvénient est la difficulté à obtenir un consensus sur les objectifs assignés à l'école.

3. La troisième solution consiste à établir un seuil à l'aide de la première technique puis à refaire passer le même test quelque temps plus tard en conservant le seuil établi (donc un **à postériori** qui devient un **à priori**) et appliquer la seconde technique, c'est-à-dire calculer la proportion d'élèves qui se trouvent encore sous le seuil fixé.

On peut alors commenter les progrès réalisés mais cela n'est légitime que si l'on tient compte du fait qu'entre deux évaluations l'enseignement s'est déployé et que les autres élèves, pendant le même intervalle de temps, on aussi fait des progrès (*cf.* l'exemple développé en introduction de cet article).

Si **la plus grande suspicion règne** aujourd'hui (on notera le silence assourdissant du CSEN sur le sujet), c'est que les résultats ont été bricolés en mélangeant ces trois logiques. Les seuils retenus sont *ad hoc*, déterminés par les stratégies de communication politique et par les priorités didactiques qui avaient été fixées au préalable. Ce ne sont pas les chiffres qui mentent.

Conséquences pratiques pour les écoles

Quelles sont les conséquences pédagogiques de ce choix stratégique ?

Les conséquences pour les écoles sont prévisibles si l'on se réfère à ce qui se passe à l'étranger dans les dispositifs de « réponse à l'intervention » :

- il faut ajouter un temps d'enseignement pour les élèves jugés « en difficulté », en plus du temps dispensé en collectif.
- Cet enseignement supplémentaire doit être organisé en petits groupes (de 4 à 8 élèves, par ex.).
- Il doit être intensif et régulier.
- Il doit bénéficier d'un horaire important : de 20 à 45 minutes par jour, 3 à 5 fois par semaine, pendant 8 à 15 semaines (c'est le cas en Ontario, par ex.).
- Il doit être dispensé par l'enseignant titulaire de la classe, pas par un enseignant spécialisé.
- Il doit répondre aux besoins identifiés par les évaluations.

Ce dernier point peut générer de graves malentendus s'il n'est pas orchestré par des didacticiens compétents. Les psychologues qui conseillent la DGESCO suggèrent pour leur part que les tâches de remédiation découlent directement des tâches d'évaluation qu'ils ont eux-mêmes élaborées : ils recommandent par exemple la multiplication d'exercices phonologiques pour remédier aux échecs constatés aux épreuves de manipulation de phonèmes. Or, cette solution est très insuffisante : les enfants qui discriminent mal les sons ont surtout besoin d'un enseignement explicite et structuré de l'encodage phonographique, c'est-à-dire de tâches d'écriture, absentes des évaluations de septembre. C'est la nécessité de segmenter la chaîne orale pour la transcrire qui donne sens et méthode à l'analyse phonémique. Mais ces tâches phono-graphémiques sont la spécialité d'autres chercheurs, didacticiens ou linguistes, non consultés ou écartés par la DGESCO.

✚ Des pratiques imposées aux enseignants ?

En Amérique du nord, les modalités d'enseignement au niveau 1 (pour tous) sont fortement prescrites par l'employeur et standardisées. Ce sont des pratiques d'**Instruction directe** (« *Direct instruction*¹⁶ ») mais ses promoteurs francophones préfèrent parler « d'enseignement explicite », ce qui crée la confusion en France puisque les programmes en vigueur définissent cette notion tout autrement¹⁷. Elles sont **modulaires** au sens où les habiletés élémentaires sont travaillées séparément, dans des modules étanches et fortement structurés. La complexité de la tâche est réduite au maximum et l'activité est découpée en sous-tâches. Dans ce modèle, les élèves ne sont pas encouragés à tâtonner, bien au contraire : c'est le professeur qui leur indique les bonnes procédures. Les auteurs canadiens qui sont à la fois les promoteurs de l'Instruction directe et de la réponse à l'intervention (Bissonnette & al. par exemple¹⁸) appellent cela du « **modelage** »¹⁹ : l'enseignant doit montrer les stratégies à utiliser pour réaliser tel ou tel exercice et donner à voir son expertise en « pensant à haute voix ». Il avance dans sa planification après avoir vérifié que chaque compétence travaillée à l'étape antérieure est maîtrisée. Les résolutions de problème ne peuvent donc être abordées que dans un second temps, parfois très éloigné, au terme d'un parcours d'acquisitions de procédures élémentaires.

Dans les dispositifs RAI, les modalités d'enseignement au niveau 2 sont voisines de celles qui sont recommandées pour tous au niveau 1 mais elles sont encore plus répétitives et multiplient les occasions de s'exercer. Elles réduisent encore plus la complexité des tâches proposées au niveau 1 et accentuent le modelage procédural par l'enseignant. Les élèves, moins nombreux, bénéficient de plus de *feed-back* immédiats.

✚ Interroger le ministère de l'Éducation nationale

Les représentants des personnels enseignants doivent rapidement interroger le MEN : ce dispositif sera-t-il obligatoire à la rentrée prochaine ? Les outils de remédiation seront-ils imposés ?

Comment les enseignants trouveront-ils le temps de procéder à de tels ateliers de remédiation ? Les ateliers RAI vont-ils se dérouler pendant les 24 heures hebdomadaires ?

- Dans ce cas, que feront les autres élèves pendant ce temps ? Vont-ils passer leur temps à un travail solitaire sur fiche ?
- Quel temps restera-t-il au professeur pour enseigner toutes les autres disciplines ? Si l'on ajoute une heure quotidienne de remédiation – 30' en français et 30' en mathématiques – aux enseignements de français et de maths, il ne restera presque plus rien. N'est-ce pas contradictoire avec les objectifs de démocratisation visés ? On sait, par exemple, que les performances des élèves en lecture-compréhension dépendent beaucoup de leurs connaissances encyclopédiques : que deviendront-elles en sciences, en histoire, en géographie... si ces disciplines disparaissent progressivement ?

¹⁶ <http://centerii.org/techassist/solutionfinding/resources/FiveMeaningsOfDI.pdf>

¹⁷ Le bureau de l'Éducation prioritaire à la DGESCO écrit qu'enseigner explicitement « ne saurait être réduit ou assimilé au seul concept "d'instruction directe" venu du continent nord-américain ». https://www.reseau-canope.fr/education-prioritaire/fileadmin/user_upload/user_upload/actualites/enseigner_plus_explicite_cr.pdf

Voir aussi « Pédagogie explicite » de Rayou (2019) : <https://www.cairn.info/revue-recherche-et-formation-2018-1-page-97.htm>

¹⁸ <http://rire.ctreq.qc.ca/2017/11/rai-dt/>

¹⁹ Gauthier, Bissonnette et Richard (2015) *L'enseignement explicite*. De Boeck.

Les ateliers RAI vont-ils se dérouler en plus des 24 heures communes ?

- Dans le cadre des APC ? L'heure hebdomadaire est insuffisante.
- Le MEN restaurera-t-il les deux heures hebdomadaires de l'aide personnalisée d'antan ?
- Redéfinira-t-il le rôle des maitres des réseaux d'aide ?

Tous les élèves dits « en difficulté » seront-ils concernés ?

- Si non, le MEN prévoit-il des dispositifs distincts, inspirés du niveau 3 de la RAI pour les élèves « à besoins » et inspirés du niveau 2 de la RAI pour les élèves « fragiles » ?
- La première solution serait cohérente avec la logique de « la réponse à l'intervention ». On peut cependant en douter lorsqu'on lit sur le site du MEN qu'au-dessous du seuil 1 « une intervention pédagogique doit être mise en œuvre sans délai » alors qu'entre les seuils 1 et 2 seule « une vigilance accrue doit être exercée ».

Pour réfléchir à cette question, nous avons réalisé une simulation basée sur les résultats de 2018-2019.

🚩 Quelle proportion d'élèves relèverait de l'intervention de niveau 2 du dispositif de « Réponse à l'intervention » à la mi-CP et au début du CE1 ?

Si, à la mi-CP, on prenait en compte les élèves déclarés « en difficulté » par la DEPP, la RAI concernerait, en moyenne, environ 20 % des élèves (entre 14 et 28 % selon les items ; cf. tableau 2). Nous l'avons vu plus haut, ce chiffre correspond aux proportions d'élèves généralement pris en charge dans les dispositifs RAI à l'étranger.

Dans l'Éducation prioritaire, malgré les CP dédoublés dont l'efficacité laisse à désirer comme nous le verrons plus loin, les besoins seraient encore plus forts : en moyenne, 30 % des élèves seraient concernés.

Tableau 2 : pourcentage d'élèves jugés « en difficulté » selon le domaine évalué et le secteur de scolarisation, mi-CP

Compétences évaluées	% d'élèves en difficulté	
	Hors Éducation prioritaire	En Éducation prioritaire +
Connaitre le nom des lettres des lettres et le son qu'elles produisent	17	23
Manipuler des phonèmes	13	20
Écrire des mots dictés	23	32
Écrire des syllabes dictées	22	29
Lire à voix haute des mots	28	38
Lire à voix haute des mots inventés	27	35
Lire à voix haute un texte	26	38
Comprendre des phrases à l'oral	14	36

Source : Roland Goigoux à partir des données de la DEPP fournies en annexes de la Note d'information 19.14 d'avril 2019

On notera que le choix ministériel d'un seuil élevé de performance en lecture à haute voix dès le mois de janvier de CP conduit à exiger une intervention dans ce domaine pour plus d'un quart des élèves hors EP et de plus d'un tiers en EP+. Il ne fait aucun doute que les horaires dédiés aux apprentissages fondamentaux au CP vont s'alourdir de fait, si ce n'est dans les décrets.

Tableau 3 : pourcentage d'élèves jugés « en difficulté » selon le domaine évalué et le secteur de scolarisation, au début du CE1

Compétences évaluées	% d'élèves en difficulté	
	Hors Éducation prioritaire	En Éducation prioritaire +
Comprendre des mots entendus	19	49
Comprendre des phrases entendues	20	44
Lire à voix haute des mots	30	45
Lire à voix haute un texte	30	44
Répondre à des questions lues par l'enseignant	17	35
Répondre à des questions lues par l'élève	20	37
Écrire des syllabes dictées	17	28
Écrire des mots dictés	23	35

Source : Roland Goigoux à partir des données de la DEPP fournies en annexes de la Note d'information 19.15 d'avril 2019

Hors Éducation prioritaire, environ 20 % des élèves sont, en moyenne, concernés mais 30 % en lecture à haute voix car le ministère a délibérément choisi un seuil plus exigeant pour des raisons stratégiques : imposer les ateliers de fluence présentés partout comme la nouvelle panacée de la pédagogie de la lecture, au risque de se substituer à tout l'enseignement de la compréhension. Exiger 30 mots à la minute à l'entrée au CE1 impliquera d'intervenir auprès de 30 % des élèves hors EP et **45 % en EP+**. Heureusement, la DEPP n'a pas suivi le guide orange de la DGESCO qui fixait ce seuil à 50 mots par minute : si elle l'avait fait, ce serait 75 % des élèves (tous publics confondus) qui auraient été jugés comme étant en difficulté.

On notera que, dans les deux cas (mi-CP et début CE1), l'écart le plus important entre les élèves de l'Éducation prioritaire et hors Éducation prioritaire se situe dans **la maîtrise du langage oral**. Il va du simple au double ou plus encore. Nous ne le répèterons jamais assez : tout miser sur l'étude du code au cycle 2 risque d'avoir, à long terme, des conséquences négatives sur la compréhension²⁰.

Le ministère ne devrait pas continuer à exclure des évaluations, donc des remédiations et des priorités d'enseignement au cycle 2, les compétences en compréhension et en production du langage. S'il n'avait pas écarté l'évaluation de la compréhension de textes entendus mi-CP et début CE1, il aurait pu conclure à la nécessité de son enseignement au cycle 2, comme nous l'avons fait nous-même au terme de la recherche *LireEcrire*²¹. Les résultats que nous avons obtenus montraient que les écarts entre EP et hors-EP en compréhension de textes entendus²² étaient très grands, du même ordre que ceux mesurés ici en lexique et syntaxe de phrase.

²⁰ Chall, J. S., Jacobs, V. A. & Baldwin, L. E. (2009). *The reading crisis: Why poor children fall behind*. Harvard University Press.

²¹ <http://ife.ens-lyon.fr/ife/recherche/lire-ecrire>

²² Roland Goigoux, Sylvie Cèbe. L'enseignement de la lecture et de l'écriture au cours préparatoire est-il vraiment de moindre qualité en éducation prioritaire ? B. Fouquet-Chauprade, A. Soussi. **Pratiques pédagogiques et éducation prioritaire**, Peter Lang, pp.153-182, 2018. <https://hal.archives-ouvertes.fr/hal-01683287/document>

DEUXIÈME PARTIE : PETITS ARRANGEMENTS AVEC LA RÉALITÉ

La cohérence politique

L'autoritarisme

Nous l'avons montré, le dispositif de traitement des difficultés scolaires charrie avec lui des modalités d'action pédagogique spécifiques (l'instruction directe). Celles-ci peuvent présenter un intérêt lorsqu'elles sont utilisées à bon escient et qu'elles ne restent qu'une technique possible parmi d'autres²³. Notre inquiétude, à la lecture de la circulaire de rentrée²⁴ et à l'écoute des formations nationales dispensées aux cadres, est qu'elles soient en passe de devenir une sorte de pédagogie officielle, du moins en français pour les premiers apprentissages.

Pour les diffuser, le ministère a choisi un mode de management autoritaire. Circulaires et livrets oranges se multiplient (4 nouveaux sont annoncés pour la rentrée prochaine), les ressources de la DEGSCO deviennent des feuilles de route obligatoires, les inspecteurs sont priés d'aller vérifier dans chaque CP dédoublé que les directives sont bien mises en œuvre, les récalcitrants sont menacés²⁵. Cette caporalisation touche aussi les équipes de circonscription : la formation continue ne sert plus qu'à exposer les directives nationales, les diaporamas à projeter sont fournis par la DGESCO, aucune réflexion critique n'est prévue au programme. Le dernier dispositif de formation de cadres départementaux sur la maternelle a été élaboré au plan national et diffusé clés-en-main dans toutes les académies : le choix des vidéos à montrer, le temps alloué aux travaux de groupes et les questions à poser à l'auditoire, tout était pré-défini et devait être appliqué scrupuleusement. Les DASEN, sermonnés par les recteurs, font à leur tour pression sur les IEN qui sont tous convoqués pour écouter la bonne parole délivrée par les affidés du ministre à l'IH2EF à Poitiers. Les intervenants choisis par le MEN sont toujours les mêmes et répètent inlassablement les mêmes choses ; les chercheurs critiques sont black-listés. Cette caporalisation mal vécue sur le terrain²⁶ est complétée par une nouvelle réforme de la formation initiale reprise en main par l'Éducation nationale au détriment de l'Université dont l'autonomie agace. Des maquettes plus homogènes et uniformes au plan national ainsi qu'un contrôle accru, par les rectorats, de la formation dispensée sont prévus. Le ministère employeur veut pouvoir décider ce qu'il est bon d'enseigner et comment. Il y a fort à parier que le petit guide orange (CP) deviendra une référence incontournable l'an prochain dans les nouvelles INSPÉ.

²³ Goigoux, R. (2010). Une pédagogie éclectique au service des élèves qui ont le plus besoin de l'école. *La nouvelle revue de l'adaptation et de la scolarisation*, 4, 21-30. <https://www.cairn.info/revue-la-nouvelle-revue-de-l-adaptation-et-de-la-scolarisation-2010-4-page-21.htm>

²⁴ <http://www.cafepedagogique.net/lexpresso/Documents/docsjoints/circurentree160419.pdf>

²⁵ <http://blog.educpros.fr/claudelelievre/2019/05/03/mieux-vaut-etre-dasen-que-directeur-decole-pour-commenter-le-projet-de-loi-blanquer/>

²⁶ https://blog.francetvinfo.fr/l-institut-humeurs/2019/05/04/portrait-robot-dun-institut-sanctionnable.html?fbclid=IwAR3E8ZCGd_D0dzERJ5V-6Fp8CNzF3WHvWd5H_IJ_tRp-gdO7KeeK3qRtdUA

La caution de La Science

Pour cautionner sa politique, le ministère a constitué un conseil scientifique (CSEN) et prétend que sa politique est rationnelle puisque fondée « sur la preuve ». Mais, lorsqu'on y regarde de plus près, on s'aperçoit que c'est faux, une fois sur deux : bon nombre d'injonctions officielles ne sont justifiées par aucune donnée probante ou sont publiées avant même que les groupes de travail du CSEN rendent leurs conclusions (par exemple, sur les manuels). De plus, le MEN ne reprend les arguments de son propre conseil scientifique que lorsqu'ils l'arrangent (il ignore ceux qui touchent au redoublement, aux rythmes scolaires, à la limitation des effectifs à 24²⁷...). Bref, les scientifiques sont utilisés pour cautionner les grands choix ministériels, parfois à leur insu, parfois en échange de la promotion de leurs travaux.

Tout le monde s'est trompé en croyant qu'avec le CSEN les neurosciences prenaient le pouvoir. Là n'était pas l'essentiel. Qu'ils viennent des neurosciences, de la psychologie, des sciences de l'éducation ou même de la sociologie, les membres du CSEN ont été choisis avant tout parce qu'ils partageaient la même épistémologie : ils défendent les politiques éducatives fondées sur la preuve (« *Evidence Based Policy* ») et veulent que les méthodes pédagogiques soient testées expérimentalement, avec un groupe témoin comparé à un groupe expérimental. Toutes les autres méthodologies de recherche sont dénigrées, leurs financements taris. Comme il n'existe quasiment aucune étude de ce type disponible en France sur les sujets jugés prioritaires, le MEN bluffe, procède à des approximations ou transpose de manière hasardeuse des résultats de recherche produits à l'étranger dans des contextes scolaires très différents.

Incohérences et autres « bobards »

Les bobards du guide orange sur l'apprentissage de la lecture

Nous avons déjà dénoncé les mensonges contenus dans ce guide en mai 2018 dans un article disponible en ligne²⁸ qui se terminait par une analyse du chapitre V consacré aux difficultés d'apprentissage et à l'arrivée prévisible de la « Réponse à l'intervention ». Nous y expliquions qu'imposer une méthode syllabique radicale, exigeant notamment que les mots à lire par les élèves soient 100 % déchiffrables, était infondé sur le plan scientifique. Franck Ramus, membre du CSEN, reconnaissait que le choix de la DGESCO était idéologique et pas scientifique²⁹, mais il lui semblait normal que les politiques fassent des choix qui correspondent à leurs promesses électorales. Nous en convenons mais, dans ce cas, pas au nom de la science !

Un an plus tard, le groupe de CSEN chargé d'étudier les manuels de lecture n'a toujours pas rendu son rapport. Son silence a été exploité par la DGESCO pour organiser une exceptionnelle campagne de propagande du syllabisme radical. Tous les inspecteurs et les maîtres de cycle 2 ont eu droit à la lecture commentée du petit livre orange. Les pressions sur les enseignants se sont multipliées, les « leçons-modèles » du guide ont été érigées en obligations réglementaires. Quelques inspecteurs zélés en ont profité pour imposer à des enseignants vulnérables les manuels syllabiques les plus caricaturaux.

²⁷ Le président Macron vient d'annoncer le plafonnement à 24 élèves par classe : les recherches disponibles laissent penser que cela ne s'accompagnera d'aucun effet sur les apprentissages des élèves. Seules les baisses drastiques (type dédoublement) produisent des effets (modérés comme nous le verrons à la fin de cet article).

²⁸ Article <https://hal.archives-ouvertes.fr/hal-01788869>

²⁹ Frank Ramus sur son compte Facebook, le 27 avril 2018

Pourtant, depuis un an, plusieurs publications scientifiques sont venues conforter nos affirmations³⁰. En juin 2018, par exemple, trois cognitivistes – parmi les meilleures spécialistes de la question – Anne Castles (Sydney), Kathleen Rastle (Londres) et Kate Nation (Oxford) ont publié une synthèse³¹. Le CSEN ne peut prétendre l'ignorer, Stanislas Dehaene vantant sa qualité « magistrale » sur son compte twitter³² : « *a must-read, masterful synthesis of decades of scientific results on reading instruction* ». Pourquoi reste-t-il alors muet sur les erreurs de la DGESCO qui prétend se baser « sur l'état de la recherche » ?

Dans la première partie de la synthèse, les trois auteures rappellent l'importance capitale de l'enseignement explicite des correspondances entre graphèmes et phonèmes dès le début de la scolarité obligatoire. Cependant, **contrairement aux assertions du guide orange**, elles indiquent que plusieurs questions restent en suspens sur le plan scientifique et qu'on ne peut pas imposer des pratiques dont on ne peut établir la légitimité.

1. On ne sait toujours pas, écrivent-elles, si les approches synthétiques (qui vont des unités graphémiques aux mots) sont préférables aux approches analytiques (basées sur la décomposition de mots entiers en syllabes puis en graphèmes). L'obligation d'une approche **exclusivement synthétique** n'a donc aucun fondement scientifique.
2. De la même manière, aucun argument ne justifie l'interdiction de la **mémorisation de mots entiers** que la DGESCO tente d'introduire. Là-encore, contrairement aux auteurs du livre orange, les chercheuses affirment que « *le choix judicieux d'un petit nombre de mots entiers à étudier en détails a sa place dans la salle de classe au côté de l'étude des correspondances graphophonologiques* » (traduit par nous). Elles ajoutent qu'il est important « *d'enseigner la lecture de mots très fréquents même s'ils sont difficiles à décoder* » (idem).
3. Elles signalent aussi que le choix de supports **entièrement décodables** a de nombreux inconvénients « *pour maintenir l'intérêt et la motivation des enfants à lire et atteindre les objectifs de construction du vocabulaire et des connaissances des enfants* » (idem).
4. Les auteurs reviennent également sur les constats de faiblesse en compréhension d'élèves de CM1 qui avaient pourtant bien démarré (au sens où ils étaient de bons décodeurs en CP). Elles insistent sur la nécessité d'enseigner explicitement, au cycle 3, des stratégies de compréhension comme la clarification (reformulation des idées), le résumé, la prédiction et la fabrication de questions. Mais elles rappellent que cela ne sert à rien si les enfants n'ont **pas assez de connaissances** sur ce dont parlent les textes. Choisir, comme on est en train de le faire en France, de consacrer tout le temps d'enseignement du français à la maîtrise du décodage au cycle 2 au détriment de tous les autres enseignements est **une erreur stratégique**.

Les cadres de l'Éducation nationale ne pourraient-ils pas retrouver un peu d'autonomie vis-à-vis de leur hiérarchie pour éviter de propager des affirmations erronées et des recommandations infondées ?

³⁰ Exposées dans une vidéo en ligne. <http://centre-alain-savary.ens-lyon.fr/CAS/education-prioritaire/ressources/theme-1-perspectives-pedagogiques-et-educatives/lire-ecrire-parler-pour-apprendre-dans-toutes-les-disciplines/roland-goigoux-enseigner-lire-ecrire-au-cp>

³¹ Article <https://journals.sagepub.com/doi/full/10.1177/1529100618772271>

³² Le 6 avril 2019. <https://twitter.com/StanDehaene>

Les mauvais résultats masqués des CP dédoublés

Une communication ministérielle habile a piégé les journalistes qui ont relayé, au printemps, le succès des dédoublés de classe alors que les résultats publiés par la DEPP étaient mauvais.

L'effet escompté, de l'aveu même du ministère, était modeste : aux environs de 0,20 (coefficient d de Cohen)³³. Dans le monde scientifique, on parle d'effet « moyen » lorsque $d > 0,50$ et d'un effet « fort » lorsque $d > 0,80$. (Un peu moins lorsque les échantillons sont plus nombreux.) Or, les résultats publiés par la DEPP révèlent que l'effet réel du dispositif CP dédoublé est **très faible** avec un $d = 0,08$ en français³⁴. Cet effet, en revanche, est avéré par un test statistique très significatif. La plupart des journalistes se sont laissés piéger par la communication ministérielle³⁵ qui valorisait la significativité du test et masquait la faiblesse de l'effet. La presse a titré sur « l'effet très significatif du dédoublement » et le tour était joué !

Sur le plan international, les chercheurs ont établi que, dans le meilleur des cas (qui n'est pas le nôtre), le dédoublement des classes est une mesure modérément efficace qui a un coût très élevé³⁶. Si un véritable débat sur l'efficacité des politiques publiques était organisé, on comparerait le rapport coût-efficacité de plusieurs dispositifs innovants. Par exemple, l'impact du dédoublement avec celui du dispositif « Plus de maîtres que de classes » (PMQC) qui, à coût comparable, touchait sept à huit fois plus d'élèves sous le précédent gouvernement. Imaginons que l'effet de ce dispositif ait été lui aussi modeste mais avéré : sachant qu'il bénéficiait à beaucoup plus d'élèves à coût équivalent, qu'en auraient conclu les citoyens ?

Pour éviter d'avoir à affronter cette question, le MEN a censuré la publication des résultats de l'évaluation du dispositif PMQC que la DEPP avait réalisée à la fin de la précédente législature. Un échantillon représentatif d'élèves avait été constitué sous la supervision de Daniel Auverlot, sous-directeur de l'évaluation et de la performance scolaire de la DEPP, devenu recteur de Créteil. Ces élèves avaient été évalués à trois reprises au cours de l'année scolaire. Leurs enseignants avaient répondu à deux questionnaires permettant de caractériser leurs choix organisationnels et pédagogiques afin de pouvoir analyser les progrès des élèves au regard de ces choix. Les premiers résultats de cette étude ont été présentés par la DEPP le 17 avril 2017 à Clermont-Ferrand à l'invitation de madame Campion, rectrice de l'académie et responsable du comité national de suivi du dispositif « Plus de maîtres que de classes »³⁷. Depuis, plus rien : pas question de fragiliser le bilan social du quinquennat.

Aucun bilan de ces quatre années d'expérimentation touchant des milliers d'élèves n'a été publié³⁸.

Un comble pour un ministère qui ne jure que par l'expérimentation et l'évaluation !

Roland Goigoux, le 10 mai 2019

³³ L'effet d'une politique de réduction de la taille des classes est mesuré en pourcentage d'écart-type de score, c'est-à-dire en termes de différence de score moyen entre le groupe témoin et le groupe de référence, rapportée à l'unité du score (son écart-type). Le coefficient d ainsi calculé permet de rendre comparable les différentes mesures de performances.

³⁴ « L'évaluation de l'impact du dédoublement sur les compétences des élèves indique que l'effet est de 8 % d'écart-type en français en faveur des élèves de REP+, par rapport au groupe témoin » écrit la DEPP.

https://cache.media.education.gouv.fr/file/01_-_janvier/00/9/DP-evaluation-janvier-2019_1065009.pdf

³⁵ Ce ne fut pas le cas pour l'Obs et des Echos.

³⁶ <https://educationendowmentfoundation.org.uk/evidence-summaries/>

³⁷ Les missions du comité national de suivi (auquel nous participions) avaient été fixées par lettre de mission du directeur général de l'enseignement scolaire en date du 24 janvier 2014. Elles portaient notamment sur « La définition des modalités d'évaluation de l'efficacité du dispositif ».

³⁸ Embargo total : si aujourd'hui les fonctionnaires de la DEPP laissaient fuiter les résultats, ils risqueraient de perdre leur place.