

PACEM : une expérimentation sur l'utilisation d'évaluations standardisées des acquis des élèves par les enseignants

Jean-François Chesné, Sandrine Prost

Bureau de l'évaluation des actions éducatives et des expérimentations
DEPP, MENJVA

Avec la collaboration d'Arthur Souletie (ENS Cachan)

et de Christelle Raffaeli (Bureau de l'évaluation des actions éducatives et des expérimentations, DEPP, MENJVA)

Le Projet pour l'acquisition de compétences par les élèves en mathématiques (PACEM) est un projet de la direction de l'évaluation, de la prospective et de la performance (DEPP) mené en CM1 et en sixième en mathématiques. Il intègre quatre dimensions principales : l'évaluation des élèves, la formation des enseignants, l'observation de leurs pratiques individuelles dans les classes et l'étude des dynamiques entre les enseignants au sein d'un même établissement. Son objectif principal vise des changements de pratiques induits par une formation, elle-même ancrée sur les résultats des élèves à un test de début d'année scolaire. À l'issue de la première année d'expérimentation, nous pouvons dire que, par comparaison à la progression d'élèves témoins qui n'ont pas suivi le protocole d'expérimentation, la performance des élèves expérimentateurs est significativement meilleure pour la sixième et très significativement meilleure pour le CM1. Cela tendrait à montrer qu'un investissement relativement faible mais spécifique de formation continue et d'accompagnement des enseignants conduit à une amélioration sensible des performances scolaires des élèves.

L'expérimentation PACEM, conçue et développée par Jean-François Chesné, avec l'appui statistique de Sandrine Prost, est un projet mené par le bureau de l'évaluation des actions éducatives et des expérimentations de la DEPP. Celui-ci a pris à sa charge l'intégralité de son développement : conception et mise en œuvre des évaluations des élèves et des formations à destination des enseignants. Si cette singularité caractérise l'étude, elle ne la rend pas moins représentative d'une partie de l'activité des différents bureaux de la DEPP. Innovante et experte, elle s'inscrit dans la dynamique d'autres études comme CEDRE¹ ou Victimation, conduites respectivement par le bureau de l'évaluation des élèves et le bureau des études sur les établissements et l'éducation prioritaire de la DEPP. Soutenue par les membres du groupe de mathématiques de l'inspection générale de l'Éducation nationale (IGEN), elle a bénéficié des appuis académiques et départementaux indispensables à son implantation et à son déploiement dans les établissements scolaires sélectionnés, ainsi que de la

collaboration de chercheurs des universités Denis Diderot et René Descartes de Paris.

Après avoir exposé les fondements et les enjeux de l'expérimentation PACEM, nous en décrivons la mise en œuvre et le développement au cours de l'année 2010-2011 au CM1, puis nous en présenterons les premiers résultats sur ce niveau. Nous énoncerons enfin nos premières interprétations, ainsi que les perspectives envisagées à court et à moyen termes et les nouvelles questions sur les conditions d'intégration des évaluations nationales dans les pratiques des enseignants.

FONDEMENTS ET ENJEUX

Un contexte d'évolution des pratiques

Cette expérimentation s'inscrit d'une part dans un contexte de questionnement sur la réussite des

NOTE

1. CEDRE : cycle d'évaluations disciplinaires réalisées sur échantillon

élèves et sur l'évolution des pratiques enseignantes, d'autre part dans un contexte d'accompagnement de la mise en place de nouveaux objectifs institutionnels.

En effet, un certain nombre d'évaluations nationales ou internationales antérieures conduites par la DEPP a montré des résultats convergents : parmi eux, nous citerons l'augmentation du nombre d'élèves de bas niveau et l'impact de leur environnement socioéconomique sur leur réussite scolaire [1]. Bien que diffusés par le biais de notes d'informations et de dossiers publiés par la DEPP, il est légitime de s'interroger sur leur prise en compte par les enseignants dans leurs pratiques quotidiennes. De façon générale, nous faisons l'hypothèse que, si les évaluations à large échelle (sur échantillons ou exhaustives) ont de plus en plus d'importance dans le pilotage du système éducatif français, la plupart des enseignants semblent considérer qu'elles constituent davantage un outil externe de mesure des acquis des élèves plutôt qu'un ensemble de ressources pour la classe [2].

Dans le même temps, les six dernières années ont vu, avec l'apparition du socle commun de connaissances et de compétences, et la réécriture des programmes de l'école et du collège, des évolutions sensibles dans l'organisation des enseignements et des pratiques pédagogiques souhaitées [3]. L'expérimentation PACEM permet, en elle-même, de porter un regard sur la question de la pénétration de ces nouveaux textes auprès du corps enseignant, de leur interprétation et de l'esprit dans lequel ils peuvent être mis en œuvre.

Enfin, cette expérimentation renvoie à la mise en œuvre des évaluations nationales de CE1 et de CM2 (et bientôt de 5^e) qui posent la ques-

tion, au-delà du retour chiffré de leurs résultats, de leur exploitation locale par les enseignants [4].

Parallèlement à ces questionnements, des dynamiques institutionnelles existent avec une volonté affirmée de faire progresser les acquis des élèves en mathématiques et d'accompagner les enseignants dans l'évolution de leurs pratiques : documents ressources pour les enseignants [5], action nationale en direction des inspecteurs de l'Éducation nationale (IEN) visant à piloter des missions mathématiques dans chaque département.

Un cadre d'expérimentation étayé

Si une expérience de la formation initiale continue et de la liaison école/collège, et une connaissance précise du monde enseignant, ont motivé la nature et les contenus de cette démarche d'expérimentation, il est nécessaire de préciser que celle-ci prend ancrage dans un cadre théorique qui est celui de la double approche, conçu et développé par Robert et Rogalski [6]. Celui-ci permet d'analyser les pratiques des enseignants, relativement aux résultats de leurs élèves en prenant en compte les tâches proposées et les contraintes liées à l'exercice du métier au quotidien, afin de déterminer ce qui peut varier dans ces pratiques [7].

Des hypothèses sur la formation des enseignants mettant en jeu l'efficacité du travail dans une « Zone proximale de développement des pratiques » nous ont conduits ensuite à élaborer des éléments de formation [8].

Des résultats institutionnels, ainsi qu'issus de la recherche, servent également de fondements à ce pro-

jet, notamment les rapports de l'IGEN sur l'école primaire [9], les travaux de Douady et de Perrin-Glorian sur les aires [10], ceux de Butlen sur le calcul mental [11] ou encore ceux de Peltier [12] et de Van Zanten sur l'éducation prioritaire [13].

Enfin, cette étude s'appuie sur la conviction forte que l'on peut « changer les choses » lorsque l'on est enseignant, ce qui n'est en fait qu'une façon de réaffirmer l'impact primordial et identifié par de nombreuses études de l'effet enseignant dans la réussite des élèves [14].

Dès lors, il apparaît très clairement que l'enjeu de cette étude, qui fait s'entrecroiser un travail sur la formation des enseignants et un travail sur les évaluations des élèves, sera d'apporter des réponses sur la question de l'appropriation des résultats d'évaluations standardisées par les enseignants et de s'interroger sur la nature des effets qu'elle peut produire sur leurs pratiques [15].

PRÉSENTATION DU DISPOSITIF

Lancée en septembre 2010, l'expérimentation se terminera en juin 2012. Sa mise en œuvre en 2010-2011 s'est faite dans deux académies différentes et a concerné deux niveaux de classe : le CM1 dans l'académie d'Aix-Marseille, dans des écoles relevant presque toutes de l'éducation prioritaire, et la sixième, dans l'académie de Créteil, sur un échantillon de collèges relevant pour moitié environ de l'éducation prioritaire. En 2011-2012, le dispositif est reconduit en CM1 et en sixième avec le même protocole et étendu aux niveaux CM2 et cinquième en suivant les élèves

testés en 2010-2011. Nous aurons donc ainsi à la fois la possibilité d'étudier son impact sur deux cohortes d'élèves suivies sur deux années et de comparer l'évolution des acquis des élèves. Complétant ces données sur les acquis des élèves, nous disposons également d'éléments d'information sur les pratiques des enseignants : observations et enregistrements audio ou vidéo de séances, réponses à un questionnaire en ligne.

Dans cet article, outre la présentation du dispositif, les résultats exposés ci-après concerneront le premier volet de l'expérimentation menée dans l'académie d'Aix-Marseille en 2010-2011. Après une brève description de la population, nous étudierons plus précisément l'évolution des taux de réussite des élèves aux items d'ancrage d'une part, l'évolution des scores de réussite entre les deux phases d'expérimentation d'autre part, dans le but de voir l'impact de l'expérimentation sur la réussite des élèves de CM1 en mathématiques.

La mise en œuvre de PACEM sur le niveau CM1 en 2010-2011

L'étude s'est concentrée sur un domaine du programme de mathématiques – Grandeurs et mesures – choisi par le groupe de pilotage des Bouches-du-Rhône en mathématiques pour le 1^{er} degré. On peut considérer qu'il s'agit d'un domaine nouveau puisqu'après plus de 30 ans d'absence, il fait sa réapparition en 2002 dans les programmes de mathématiques comme domaine d'étude. L'immense majorité des enseignants actuels n'en possède donc qu'une culture assez restreinte, tant du point de vue du savoir mathématique que des contenus à enseigner.

L'observation des résultats a porté sur trois circonscriptions : deux sont engagées dans l'expérimentation, Marseille 11 et Marseille 12, tandis que l'autre, celle d'Istres, constitue la circonscription témoin.

L'expérimentation PACEM repose sur un certain nombre de principes organisationnels :

- deux tests proposés aux élèves : l'un en début (test 1), l'autre en fin d'année scolaire (test 2) ;
- une formation proposée à certains enseignants ;
- l'accès à un site collaboratif pour les enseignants ;
- un travail de mutualisation des pratiques pour les enseignants des écoles engagées, dans le projet ;
- une collaboration de la DEPP avec les IEN, les conseillers pédagogiques et le groupe de pilotage départemental mathématiques.

La population concernée par le projet

Tous les enseignants ayant une classe de CM1, ou un double niveau en 2010-2011, dans une école des circonscriptions de Marseille 11, Marseille 12 et d'Istres, sont impli-

qués, soit 86 enseignants. Parmi ceux-ci, nous distinguerons trois catégories :

- 25 enseignants expérimentateurs qui ont bénéficié directement d'une formation : ce sont les « correspondants » du projet ;
- 27 autres enseignants expérimentateurs, dits « associés » au projet : ce sont les enseignants des écoles engagées dans le projet qui n'ont pas participé à une formation spécifique liée au projet ;
- 34 enseignants « témoins ».

Au total, l'expérimentation concerne 11 écoles et 24 classes pour Marseille 11, 14 écoles et 28 classes pour Marseille 12, et 15 écoles pour 34 classes dans la circonscription d'Istres.

Le nombre d'élèves ayant participé à l'expérimentation diffère entre les deux tests : 1 568 élèves ont participé à la première phase et 1 527 à la deuxième phase. Afin de comparer les résultats des élèves lors des deux phases d'expérimentation, nos analyses porteront uniquement sur les élèves qui ont été présents aux deux tests, soit exactement 1 400 élèves.

Graphique 1 – Caractéristiques démographiques des élèves de CM1 testés dans l'expérimentation PACEM en 2010-2011 (en %)

Marseille 12	Zone d'éducation	EP 70,6		Non EP 29,4
	Age de scolarité	2,2	À l'heure 80,1	
		Avance	Retard 17,7	
	Sexe	Garçon 47,8	Fille 52,2	
Marseille 11	Zone d'éducation	EP 92,4		Non EP 7,6
	Age de scolarité	1,5	À l'heure 76,8	
		Avance	Retard 21,7	
	Sexe	Garçon 49,5	Fille 50,5	
Istres	Zone d'éducation	10,8	Non EP 89,2	
	Age de scolarité	2,8	À l'heure 86,6	
		Avance	Retard 10,6	
	Sexe	Garçon 53,8	Fille 46,2	

Source : MENJVA-DEPP.

Le groupe des élèves dits « expérimentateurs » représente 58,9 % de l'effectif total réparti sur deux circonscriptions Marseille 11 (N = 406) et Marseille 12 (N = 418). Quant au groupe des élèves dits « témoins », il représente 41,1 % des élèves (N = 576) participant à l'expérimentation.

Globalement, la part des garçons est équivalente à celle des filles (50,8 % vs 49,2 %)². La grande majorité des élèves scolarisée est « à l'heure » (81,9 %), 15,9 % sont en retard et 2,2 % en avance. C'est à Marseille 11 que la part des élèves scolarisés en retard est la plus importante (21,7 % vs 17,7 % à Marseille 12 et 10,6 % à Istres (*graphique 1*), alors qu'Istres enregistre une plus grande proportion d'élèves scolarisés en avance (2,8 % contre 2,2 % à Marseille 12 et 1,5 % à Marseille 11).

Enfin, les élèves scolarisés en éducation prioritaire représentent un peu plus de la moitié de notre population (52,3 %). Ils sont très largement surreprésentés à Marseille 11 et sous-représentés à Istres (92,4 % à Marseille 11 contre 10,8 % à Istres).

Les cahiers d'évaluation

Les cahiers des tests 1 et 2 ont une structure similaire, seul le nombre d'items change. Ils se décomposent en deux parties de 30 minutes de passation : la première, sur les unités de mesures et la seconde sur la résolution de problèmes. Le test 1 comporte 43 items (31 items dans la première partie et 12 dans la seconde) alors que le test de la seconde phase comprend 45 items (30 items dans la première partie et 15 dans la deuxième).

Dans la première partie, les élèves ont, soit à répondre à des questions à choix multiples (QCM), soit à produire des réponses courtes. La deuxième partie ne comporte que des QCM.

Les contenus des deux tests sont évidemment différents, à la fois pour ne pas favoriser une préparation des élèves au test 2 aux dépens de leur apprentissage, mais aussi pour tenir compte de l'année scolaire écoulée. Cependant, 7 items sont communs aux deux tests dans la perspective d'une comparaison temporelle (items d'ancrage).

Les contenus sont des tâches correspondant aux connaissances et aux capacités indiquées dans les programmes de CE2 et de CM1 : la connaissance des unités de mesure et les relations qui les lient, la lecture de l'heure sur une montre à aiguilles ou une horloge et l'utilisation de la règle graduée pour mesurer des longueurs (partie 1 des tests) ; le calcul du périmètre d'un polygone et la résolution de problèmes (parties 2a et 2b). Le rythme des exercices n'était pas imposé. Des indications régulières sur la durée restante étaient données aux élèves (toutes les 10 minutes, et 2 minutes avant la fin des tests).

Des indications pouvaient être fournies aux élèves seulement à partir du moment où elles favorisaient une activité mathématique sans la déformer. Par exemple, les enseignants pouvaient indiquer aux élèves ce qu'est « un écolier » dans l'exercice 1, ou « un alpiniste » ou « une paroi » dans l'exercice 13, mais ne devaient pas répondre aux questions du type « Qu'est-ce que c'est un centilitre ? », « Qu'est-ce que c'est les kilogrammes », « Il y a combien d'heures dans une journée ? », « Qu'est-ce que c'est un segment ? » ou « Qu'est-ce que c'est le périmètre ? ».

Dans l'exercice 10, les enseignants pouvaient, si les élèves le demandaient, dire que « polygone » signifie

« figure ». Chaque question posée par un élève devait être reprise pour toute la classe et les indications données à l'ensemble des élèves.

Les conditions habituelles de travail d'élèves à profil particulier (non lecteurs ou en situation de handicap) ont été respectées : selon les cas, les énoncés des exercices ont été lus individuellement, une durée supplémentaire de passation a été accordée ou l'élève a été aidé par un(e) assistant(e) de vie scolaire (AVS).

Le contenu et la passation des tests sont donc très similaires à d'autres tests nationaux passés par les élèves. Nous sommes toutefois conscients des limites de ce type d'évaluation, en particulier sur le thème des grandeurs et des mesures, puisque les modalités d'évaluation retenues ne renvoient sans doute que partiellement aux compétences potentiellement acquises par les élèves sur l'ensemble d'une année ou d'un cycle scolaire.

La formation

La caractéristique principale de cette expérimentation est de lier la formation des enseignants à l'analyse des items proposés dans les tests et de leurs résultats, et à l'intégration de leurs progressions annuelles dans leur classe. Cette formation est destinée à favoriser, chez les enseignants, l'appropriation d'évaluations standardisées et de les conduire progressivement à modifier leurs pratiques ordinaires individuelles. Elle a également vocation à accentuer la dynamique d'équipe en suggérant des échanges plus réguliers

NOTE

2. Les différences observées au sein de chaque circonscription ne sont pas significatives.

sur l'utilisation d'outils ou la mise en œuvre de pratiques innovantes.

Le projet, perçu dans sa globalité, comporte une dimension à la fois individuelle et collective. Au niveau de la classe, il vise la prise en compte, par chaque enseignant, des réussites des élèves, mais aussi la compréhension de leurs difficultés afin de mettre en œuvre des stratégies pour les traiter. Les apports reçus et les échanges entre collègues lors de la formation ont pour objectif de les amener à adapter leur enseignement quotidien au sein de la classe, à construire de nouvelles séquences pour leur classe, à les penser différemment dans la progression annuelle. Ce questionnement des pratiques personnelles devrait permettre d'enclencher une dynamique à une autre échelle, celle de l'école, puisque ces démarches demandent à être partagées, questionnées, transférées. Ce travail peut même aller jusqu'à la construction d'un projet commun d'enseignement sur le cycle 3.

L'accompagnement des enseignants pendant le projet se décline en trois composantes. La première est une formation minimale en présentiel : une demi-journée de lancement en septembre 2010 pour tous les enseignants expérimentateurs, suivie de deux jours début octobre 2010 réservés aux correspondants. Au cours de l'année, les enseignants expérimentateurs ont accès à une plate-forme collaborative et un suivi local est effectué dans les écoles par les IEN et les conseillers pédagogiques. Le cœur de ce dispositif, le moment de formation en présentiel, s'articule autour de quatre temps :

- une phase de déconstruction : il s'agit d'amener les enseignants à accepter de lâcher prise sur cer-

taines de leurs pratiques et de leurs représentations, de « faire de la place pour du nouveau ». L'utilisation d'éléments statistiques lors de ce premier temps de formation s'avère très utile, permettant d'interroger et donc d'objectiver la difficulté d'une tâche ;

- une phase d'apport de contenus de la part du formateur : contenus mathématiques et éléments de didactique « simples », liens avec le programme, explicitation des capacités attendues ;

- une phase d'appropriation : des tâches alternatives à celles des fichiers et manuels utilisés sont proposées aux enseignants, qu'ils ont à mettre en situation, à travers différents scénarios, potentiellement adaptables à la classe ;

- une phase d'organisation, de structuration et de contextualisation : ce dernier temps de la formation fait écho au premier puisqu'il s'agit de proposer aux enseignants des éléments leur permettant d'une part, de reconstruire des séances et, à plus longue échéance, des séquences, et d'autre part, de réviser la façon dont ils concevaient jusqu'alors leurs progressions.

PREMIERS RÉSULTATS AU CM1 POUR 2010-2011

Résultats par item

Les scores de réussite obtenus au test 1 ont fait l'objet d'une attention particulière puisque les items des cahiers d'évaluation des élèves ont été choisis à la fois comme caractéristiques des attendus (partiels) au regard du programme de CE2 et dans la perspective d'une utilisation

en formation. Une analyse item par item a donc été réalisée, et communiquée aux enseignants expérimentateurs dans un document transmis aux IEN et déposé sur le site collaboratif.

Analyse des items d'ancrage

Les scores obtenus aux sept « items d'ancrage » – dont les énoncés sont présentés dans l'encadré page suivante –, permettent de comparer les résultats obtenus à ces deux tests, et donc de déterminer l'impact de l'expérimentation sur l'évolution des acquis des élèves [16].

Pour chaque item d'ancrage, on observe globalement une très nette amélioration des résultats des élèves entre les phases 1 et 2, allant de + 8,4 points pour l'item 2 à + 32,4 points pour l'item 4 (*graphique 2*). Ce sont les items 4, 5 et 7 qui enregistrent la plus forte progression.

En effet, l'item 4, qui avait été réussi par 32,6 % des élèves lors de la phase 1, l'a été par 64,9 % lors de la phase 2. De même, les items 5 et 7, réussis respectivement par 55,6 % et 12,7 % d'élèves lors de la phase 1, l'ont été par respectivement 77,9 % et 43,4 % des élèves à la phase 2. Notons que l'item 7, seul item d'ancrage portant sur la résolution de problèmes, est nettement moins bien réussi aux deux phases que les autres items d'ancrage qui portaient sur des unités de mesure. Il enregistre néanmoins une progression importante entre les deux phases de 30,7 points.

Cette progression n'est cependant pas homogène entre les trois circonscriptions (*graphique 3*). En effet, c'est à Marseille 11 que les résultats des élèves progressent le plus, et ceci sur chaque item d'an-

Énoncés des items d'ancrage

Item 1 :

Un bébé à sa naissance pèse :

 3 grammes 3 kilogrammes 30 kilogrammes

E4MGU010401

Item 2 :

Le réservoir d'essence de la voiture contient 45

E4MGU020601

Item 3 :

La taille d'un bébé est 53

 mm m cm

E4MGU030401

Item 4

1 minute 30 secondes = secondes

E4MGU050101

Item 5 :

4 semaines = jours ;

E4MGU050401

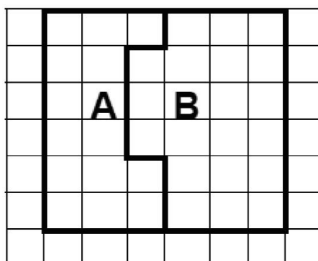
Item 6 :

1 année = mois

E4MGU050501

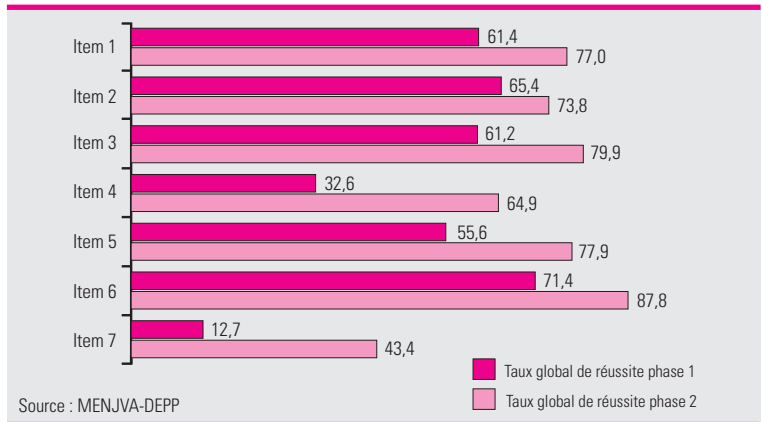
Item 7 :

Un terrain a été partagé comme l'indique la figure ci-dessous.

 Le périmètre de la partie A est le plus grand. Les parties A et B ont le même périmètre. Le périmètre de la partie B est le plus grand.

E4MGP010101

Graphique 2 – Taux de réussite aux items d’ancrage de l’ensemble des élèves.

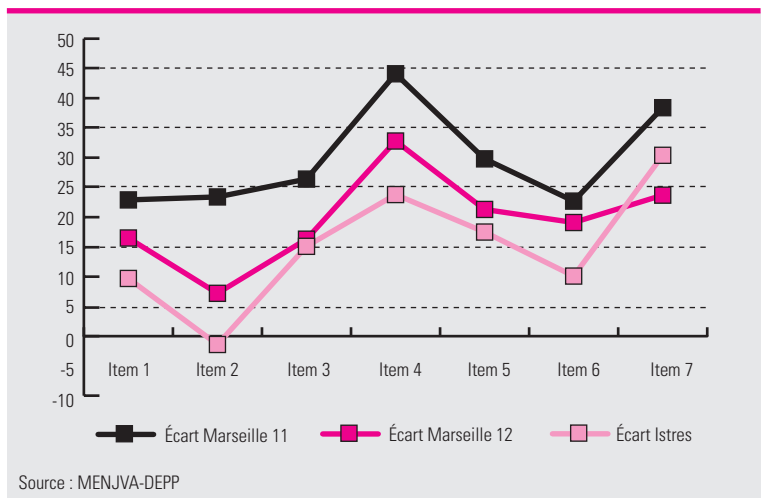


crage, alors qu’au contraire ceux des élèves d’Istres (échantillon témoin) présentent la plus faible progression (mais à partir de scores initiaux plus élevés). Par exemple, sur l’item 4, alors qu’on observe une progression de + 44,1 points pour Marseille 11 ou de + 32,8 points pour Marseille 12, la progression n’est que de + 23,8 points pour Istres. Quant à l’item 2, alors que la progression est égale à + 23,4 points pour Marseille 11, on constate pour Istres un écart de - 1,4 point entre le taux de réussite de la phase 1 et celui de la phase 2.

Analyse des scores de réussite

Lors de la première phase de l’expérimentation, les élèves totalisaient un score moyen de réussite de 27,4 à Istres, 21,4 à Marseille 11 et 23,1 à Marseille 12 sur 43 items. Lors de la deuxième phase, ils avaient en moyenne réussi 30,4 items à Istres, 28,8 items à Marseille 11 et 27,1 items à Marseille 12 sur 45. Les tests étant différents et ne comportant pas le même nombre d’items, nous ne sommes pas en mesure, dans le cadre d’une restitution de premiers résultats, de produire une

Graphique 3 – Évolution des taux de réussite aux items d’ancrage entre la phase 1 et la phase 2 par circonscription.



analyse approfondie de l’évolution des scores entre les deux phases de l’expérimentation. Toutefois, la standardisation des scores de réussite des élèves expérimentateurs sur ceux des témoins, les élèves d’Istres, nous permet de confirmer nos observations sur les taux de réussite aux items d’ancrage entre les trois circonscriptions. Le *graphique 4* montre que les mesures de tendance centrale¹ du score de réussite des élèves de Marseille 11 ont très nettement progressé par rapport à celles d’Istres : lors de la première phase, elles étaient inférieures à celles d’Istres, alors qu’à la deuxième phase, elles sont quasiment au même niveau. Quant à celles de Marseille 12, leur évolution est légèrement supérieure à celle d’Istres.

Pour Marseille 11, l’intervalle interquartile est plus étendu lors de la phase 2 : il augmente de 11 points. De plus, la médiane qui était inférieure à la moyenne lors de la phase 1, lui devient supérieure lors de la phase 2. Ce renversement de tendance indique que la distribution des scores qui était plus allongée vers les grandes valeurs devient plus allongée vers les petites valeurs. Enfin, la distribution des scores est plus étendue lors de la phase 2.

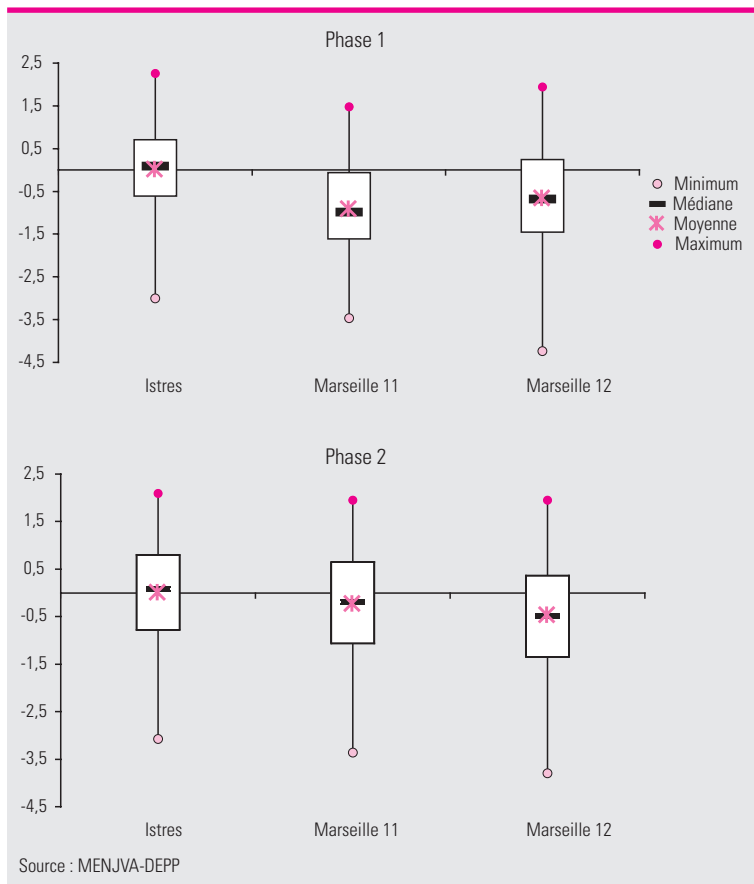
À Marseille 12 en revanche, l’intervalle interquartile reste quasiment le même. La distribution des scores y est symétrique lors des deux phases : la moyenne et la médiane se confondent et l’étendue est stable.

Lors de la première phase, l’écart observé entre le score de réussite de Marseille 11 et celui d’Istres

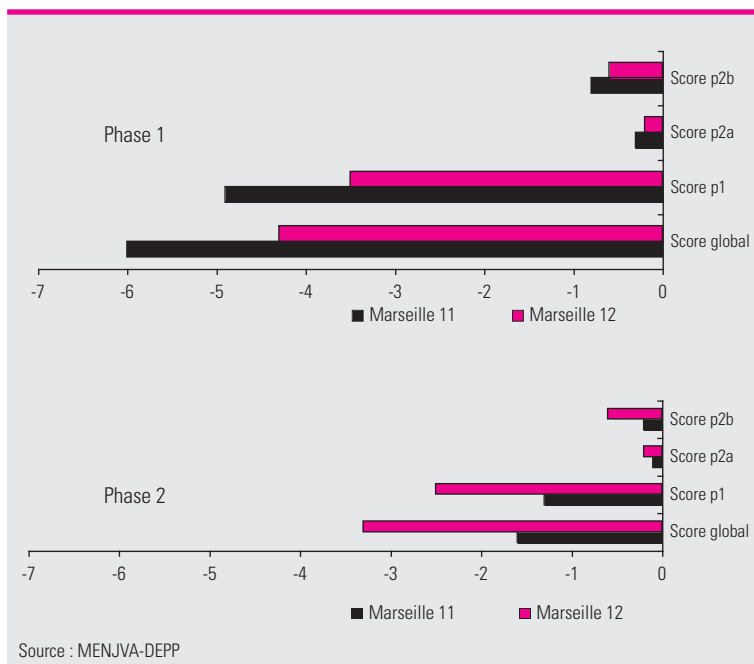
NOTE

3. Il s’agit ici de la moyenne et de la médiane.

Graphique 4 – Scores de réussite globaux par circonscription



Graphique 5 – Écarts entre les scores des élèves expérimentateurs (Marseille 11 et 12) et ceux des élèves témoins (Istres)



est beaucoup plus grand que celui observé entre Marseille 12 et Istres. Cette tendance s'inverse lors de la deuxième phase de l'expérimentation alors même que l'écart Marseille 12 – Istres s'est lui-même réduit (*graphique 5*). Lors de la première phase de l'expérimentation, Marseille 11 enregistre un score inférieur de - 6 points par rapport à Istres, alors que la différence de Marseille 12 avec Istres est égale à - 4,3 points. Lors de la deuxième phase, Marseille 11 n'a plus qu'un écart de - 1,6 point avec Istres, alors que Marseille 12 en a un de - 3,3 points. C'est principalement sur la première partie du test que les écarts se creusent. Sur la deuxième partie du test, les écarts existent, mais sont beaucoup moins marqués.

Si l'on s'intéresse aux groupes d'élèves définis par les quartiles des scores de réussite standardisés (*graphique 6*), on remarque que les scores des élèves de Marseille 11 augmentent plus que ceux des élèves des deux autres circonscriptions pour tous les groupes, du groupe 1, constitué par les élèves qui ont les acquis les plus fragiles, au groupe 4 des élèves qui obtiennent les scores les plus élevés. De plus, alors que les scores des élèves de Marseille 12 se rapprochent eux aussi de ceux d'Istres pour tous les groupes sauf le groupe 3, on peut observer que, si les scores les plus bas étaient ceux de Marseille 11 à la phase 1, ce sont ceux de Marseille 12 qui le sont à la phase 2. On remarque enfin que le bénéfique s'effectue surtout en faveur des élèves dont les résultats sont bas, en particulier à Marseille 11. En revanche, pour les élèves ayant les meilleurs scores à la phase 1 de l'expérimentation, on relève certes une progression, mais moins importante.

Le graphique 7 confirme ces analyses : les élèves dont les résultats étaient les plus bas (groupe 1) améliorent très nettement leur score à la phase 2 contrairement à ceux dont les résultats étaient les plus élevés (groupe 4). Les élèves ayant un score moyen (groupe 2 et 3) suivent la même tendance que les élèves ayant un score élevé, mais de façon moins marquée, excepté à Marseille 11 où ceux du groupe 2 progressent quasiment autant que ceux du groupe 1.

CONCLUSION

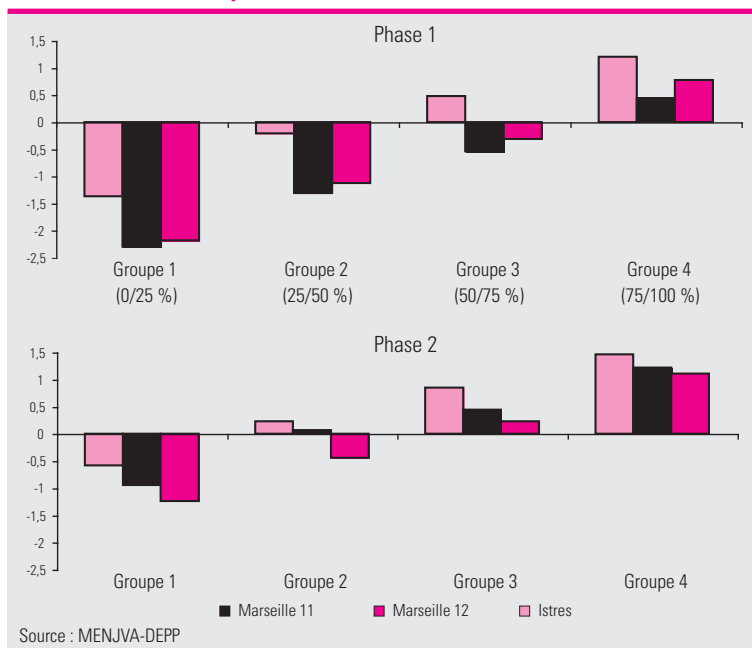
Nous sommes conscients qu'un tel projet, par son approche pluridimensionnelle et les nombreuses variables auxquelles nous avons été confrontés, nécessite une certaine prudence dans la méthodologie à adopter et la présentation de ses résultats.

En effet, la démarche adoptée qui consiste à lier l'évaluation des acquis des élèves aux pratiques des enseignants et à leur formation, tend à en complexifier l'analyse et à en inscrire le traitement dans une temporalité longue. Il est également nécessaire d'avoir à l'esprit que, dans toute étude de ce type, l'effet établissement, classe ou enseignant, sur les résultats des élèves – tout en restant difficile à mesurer – est non négligeable.

Par ailleurs, l'influence d'un tel dispositif dans le temps, sur les apprentissages et sur les pratiques n'est pas sans poser question [17]. Faut-il, du point de vue enseignant, se placer du côté des pratiques efficaces ou des « bonnes pratiques » ? Un tel dispositif est-il efficient à moyen terme ?

L'analyse des taux de réponse aux items d'ancrage et celle des scores

Graphique 6 – Scores moyens de réussite en fonction des quartiles de score de la phase 1, du groupe 1 (élèves aux scores les plus bas) au groupe 4 (élèves aux scores les plus élevés)

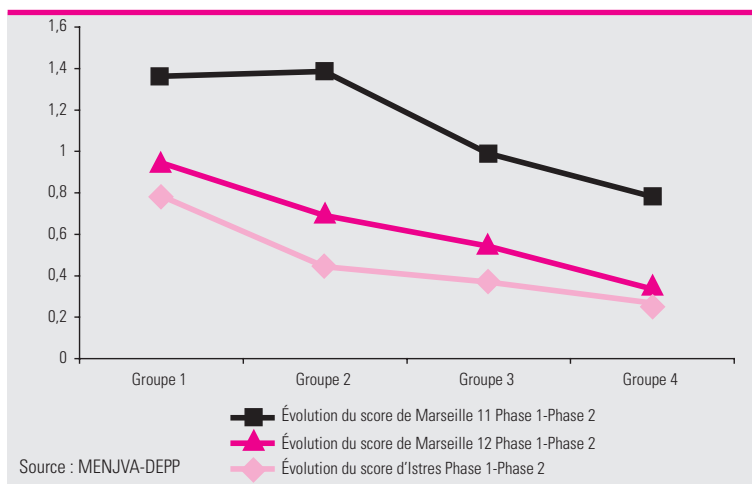


de réussite convergent toutes deux vers les mêmes conclusions. Les résultats des élèves de Marseille 11 et de Marseille 12 ont davantage progressé entre les deux tests que ceux d'Istres, Marseille 11 étant la circonscription où les scores ont le plus augmenté. Notons qu'il s'agit de la circonscription où la part d'élèves relevant de l'éducation prioritaire est la plus importante, avec des élèves

scolarisés en retard plus nombreux et partant dans l'ensemble d'un niveau plus bas.

Sur la base de ces premiers résultats, il semble donc que le projet PACEM ait produit un impact positif sur les acquis des élèves de CM1 ayant bénéficié du dispositif. La progression enregistrée concerne surtout les connaissances sur les unités de mesure des élèves les plus

Graphique 7 – Évolution des scores moyens de réussite entre les deux phases en fonction des quartiles de score de la phase 1



en difficulté, alors que la réussite des élèves dans la résolution de problèmes, peu élevée au départ, a moins progressé. Même si cette différenciation des progrès consta-

tée chez les élèves ne reflète que partiellement des changements de pratiques des enseignants, cela tendrait à montrer qu'un investissement relativement faible mais spécifique

de formation continue et d'accompagnement des enseignants pourrait conduire à une amélioration sensible des performances scolaires. ■

BIBLIOGRAPHIE

- [1] **A. Brun et J.-M. Pastor** (2010), Les compétences en mathématiques des élèves en fin d'école primaire. *Note d'information* n° 10.17, http://media.education.gouv.fr/file/2010/23/5/NIMEN1017_158235.pdf
- [2] **IGEN** (2010), *La mise en œuvre des programmes dans les classes et les évaluations nationales des élèves comme outil de pilotage du premier degré*, Note n° 2010-092, http://media.education.gouv.fr/file/2010/93/9/notereformeprimaire_154939.pdf
- [3] **Programmes de l'école et du collège** (2008) <http://eduscol.education.fr/pid23391/programmes-de-l-ecole-et-du-college.html>
- [4] **Hyer J.-H.**, (2011) *If you build it will they come? Teacher use of student performance data on a web-based tool*. Working Paper 17486, <http://www.nber.org/papers/w17486>
- [5] **Le nombre au cycle 2** (2011), Ressources pour faire la classe, SCEREN http://media.eduscol.education.fr/file/ecole/00/3/Le_nombre_au_cycle_2_153003.pdf
- [6] **Robert A. et Rogalski J.** (2005) *A cross-analysis of the mathematics teacher's activity. An example in a French 10th-grade class*. Educational Studies in Mathematics volume 59, 1 pp 269-298.
- [7] **Vandebrouck F.**, (2007) *La classe de mathématiques : activités des élèves et pratiques des enseignants*. Octares éditions, Toulouse
- [8] **Chesnè J.-F., Pariès M., Robert A.** (2009) *Partir des pratiques en formation professionnelle des enseignants de mathématiques des lycées et collèges*. Petit x n° 80, pp 25-46, Grenoble
- [9] **IGEN** (2006) *L'enseignement des mathématiques au cycle 3 de l'école primaire ; Rapport n° 2006-034 ;* <http://media.education.gouv.fr/file/46/0/3460.pdf>
- [10] **Douady R. et Perrin-Glorian M.-J.** (1983), *Mesure des longueurs et des aires*, brochure n° 48, I.R.E.M. université Paris VII, Paris,
- [11] **Butlen D.** (2007) *Le calcul mental entre sens et technique*. Presses universitaires de Franche-Comté, Besançon
- [12] **Peltier M.-L. et al**, (2004) *Dur, dur d'enseigner en ZEP – Analyse de pratiques des professeurs des écoles*. La Pensée Sauvage, Grenoble
- [13] **Van Zanten A.** (2001) *L'école de la périphérie – Scolarité et ségrégation en banlieue*. PUF, Paris
- [14] **Gauthier C. et Dembélé M.** (2005) *Qualité de l'enseignement et qualité de l'éducation : revue des résultats de recherche*. UNESCO Report, 2005/ED/EFA/MRT/PI/18
- [15] **Nye B., Konstantopoulos S., Hedges L.**, (2004) *How Large Are Teacher Effects?* Educational Evaluation and Policy Analysis, Vol. 26, n° 3, pp. 237-257
- [16] **Laveault D. et Grégoire J.** (2002). *Introduction aux théories des tests en psychologie et en sciences de l'éducation*. De Boeck université, Paris.
- [17] **Konstantopoulos S.** (2007) *How long do teacher effects persist?* Institute for the study of labor (IZA), n° 2 893, Mahwah, Erlbaum, New Jersey