



FEJ
FONDS
D'EXPÉRIMENTATION
POUR LA
JEUNESSE

« PROJET LECTURE »
RAPPORT D'ÉVALUATION
« DEPP, LABORATOIRE EMC, IREDU »

Ministère des Sports, de la Jeunesse, de l'Éducation populaire et de la Vie associative
Fonds d'Expérimentation pour la Jeunesse
95 avenue de France – 75650 Paris cedex 13
www.experimentation.jeunes.gouv.fr

Rapport d'évaluation HAP11

Page 1 sur 86



Cette évaluation a été financée par le Fonds d'Expérimentation pour la Jeunesse dans le cadre de l'appel à projets lancé en 2011 par le Ministère chargé de la jeunesse.

Le fonds d'expérimentation est destiné à favoriser la réussite scolaire des élèves et améliorer l'insertion sociale et professionnelle des jeunes de moins de vingt-cinq ans. Il a pour ambition de tester de nouvelles politiques de jeunesse grâce à la méthodologie de l'expérimentation sociale. A cette fin, il impulse et soutient des initiatives innovantes, sur différents territoires et selon des modalités variables et rigoureusement évaluées. Les conclusions des évaluations externes guideront les réflexions nationales et locales sur de possibles généralisations ou extensions de dispositifs à d'autres territoires.

Les résultats de cette étude n'engagent que leurs auteurs, et ne sauraient en aucun cas engager le Ministère.

Ministère des Sports, de la Jeunesse, de l'Éducation populaire et de la Vie associative
Direction de la Jeunesse, de l'Éducation populaire et de la Vie associative

Mission d'animation du Fonds d'Expérimentation pour la Jeunesse
95, avenue de France 75 650 Paris Cedex 13
Téléphone : 01 40 45 93 22
<http://www.experimentation.jeunes.gouv.fr>

Pour plus d'informations sur le déroulement du projet, vous pouvez consulter sur le site <http://www.experimentation.jeunes.gouv.fr> la note de restitution finale soumise au FEJ par le porteur de projet.



Intitulé du projet

« Projet Lecture »

Structure porteuse du projet : Agir pour l'école
Françoise Colloc'h, présidente

Structures porteuses de l'évaluation : DEPP, Laboratoire
EMC- Lyon 2, SODAR-IREDU

Nom de la structure : Direction de l'Évaluation de la Prospective et de la
Performance

Responsable : directrice, Catherine Moisan

Contact projet : Thierry Rocher

Nom de la structure : Laboratoire EMC – université Lyon 2

Responsable du projet : Jean Ecalle

Nom de la structure : SODAR-IREDU

Responsable : directeur, Thierry Chevallier

Contact projet : Bruno Suchaut

Durée d'expérimentation :

2 ans

Date de remise du rapport d'évaluation : février 2014



RÉSUMÉ

Le projet Lecture est un projet de suivi de cohorte, tout au long du cycle d'apprentissage de la lecture : de la grande section de maternelle au CE1. Son objectif est de réduire de moitié le nombre d'élèves en grande difficulté à l'issue du CE1.

Le projet consiste à systématiser le recours des enseignants aux pratiques pédagogiques efficaces préconisées par la recherche, notamment l'entraînement structuré des compétences fondamentales de la lecture, en petits groupes, et à focaliser leur attention sur les élèves les plus faibles. La valeur ajoutée réside dans le recours aux enseignements de la recherche, en psychologie cognitive, notamment, au service de l'objectif de réduction des inégalités d'accès à l'apprentissage de la lecture constatées. Le protocole préconise donc la mise en œuvre d'interventions précoces dans la scolarité des élèves. Cette expérimentation se situe dans la continuité du programme PARLER qui a été conduit de 2005 à 2008 dans plusieurs écoles de REP de l'agglomération grenobloise et qui avait montré des effets positifs. Le protocole expérimental propose un guide à l'enseignant pour mettre en œuvre en grande section et au CP un enseignement structuré des compétences identifiées comme nécessaires à l'apprentissage de la lecture : conscience phonologique, compréhension orale et code alphabétique. Les conseillers pédagogiques de circonscription assurent le suivi, la formation et l'accompagnement des enseignants engagés dans le dispositif. La méthode promeut le travail des élèves en petits groupes de niveau homogène, chaque séance d'entraînement durant 30 minutes. La fréquence des entraînements est augmentée pour les élèves les plus faibles.

L'évaluation du dispositif expérimental est sous la responsabilité de la DEPP, du laboratoire EMC de l'Université de Lyon 2 et de l'IREDU. Sur les deux années de l'expérimentation, quatre évaluations, deux en grande section et deux en CP, ont été mises en œuvre.

Sur les deux années, l'analyse des effets fait ressortir des résultats positifs sur les dimensions travaillées spécifiquement dans le dispositif, ces effets étant plus importants pour les élèves les plus faibles.

Si la littérature scientifique américaine abonde en recherches concordantes sur les effets positifs d'une intervention précoce en phonologie, il n'existe pas en France d'études empiriques d'envergure sur le sujet. De ce point de vue, l'expérimentation fournit des éléments utiles en termes de recherche et d'action pédagogique. Il y aura en effet un enjeu à observer si les résultats positifs observés en cycle 2 (GS- CP- CE1) sont durables et transposables dans l'apprentissage de la lecture chez ces élèves.



NOTE DE SYNTHÈSE

La lutte contre les difficultés scolaires constitue un important enjeu social. Environ 20 % des élèves éprouvent des difficultés face à l'écrit et ce taux est en augmentation depuis une dizaine d'années (Daussin et al., 2011). L'école est face à une diversité de profils de lecteurs en difficultés et l'enjeu est d'y faire face dans les meilleures conditions. L'une d'elles est de connaître de façon la plus précise possible les processus cognitifs et langagiers déficitaires et menant à des difficultés lors de l'apprentissage de la lecture.

Le projet « Lecture » est un projet de suivi de cohorte, tout au long du cycle d'apprentissage de la lecture : de la grande section de maternelle au CE1. Son objectif est de réduire de moitié le nombre d'élèves en grande difficulté à l'issue du CE1.

Le protocole de l'expérimentation « Lecture » est conçu à partir du principe d'enseignement explicite. Il préconise la mise en œuvre d'interventions précoces dans la scolarité des élèves afin de lutter efficacement contre l'échec scolaire. Les grands principes de cette expérimentation sont des adaptations au contexte français d'un rapport de 2000 rédigé par le National Reading Panel qui identifiait les interventions efficaces favorisant l'apprentissage de la lecture : connaissance des lettres, phonologie, connaissance du principe alphabétique. En outre, l'expérimentation « Lecture » se situe dans la continuité du programme PARLER (M. Bianco, P. Bressoux, C. Lequette, G. Pouget, Laboratoire des Sciences de l'Éducation, Université Pierre-Mendès-France de Grenoble, étude menée par Michel Zorman) qui a été conduit de 2005 à 2008 dans plusieurs écoles de REP de Grenoble et de l'agglomération grenobloise et qui a montré des effets positifs.

Le projet « Lecture » consiste ainsi à systématiser le recours des enseignants à des pratiques pédagogiques reposant sur l'entraînement structuré des compétences fondamentales de la lecture et à focaliser leur attention sur les élèves les plus faibles. La valeur ajoutée réside dans l'application des enseignements de la recherche, en psychologie cognitive, notamment, au service de l'objectif de réduction des inégalités d'accès à l'apprentissage de la lecture constatées.

Le protocole expérimental propose un guide à l'enseignant pour mettre en œuvre en grande section et au CP un enseignement structuré des compétences identifiées comme nécessaires à l'apprentissage de la lecture : conscience phonologique, compréhension orale et code alphabétique. Les conseillers pédagogiques de circonscription et des maîtres formateurs assurent le suivi, la formation et l'accompagnement des enseignants engagés dans le dispositif. La méthode promeut le travail des élèves en petits groupes de niveau homogène, chaque séance d'entraînement durant 30 minutes. La fréquence des entraînements est augmentée pour les élèves les plus faibles. Des documents d'accompagnement pour la mise en œuvre précisent les exigences du protocole quant au cadre des séances, à la conduite de chaque séance et à la progression annuelle des séances à respecter. Le matériel pédagogique est fourni aux enseignants via une plateforme d'échanges.

Si la littérature scientifique américaine abonde en recherches concordantes sur les effets positifs d'une intervention précoce en phonologie, il n'existe pas en France d'études empiriques d'envergure sur le sujet. De ce point de vue, l'expérimentation fournit des éléments utiles en termes de recherche et d'action pédagogique.

L'expérimentation du projet « Lecture » s'est tenue dans les départements du Nord, du Pas-de-Calais, des Hauts-de-Seine et de la Seine-Saint-Denis lors de l'année scolaire 2011-2012, auprès d'élèves de grande section (GS), et lors de l'année scolaire 2012-2013 auprès d'élèves de cours préparatoire (CP). Dans les circonscriptions identifiées, des classes participent à l'expérimentation sur la base du volontariat.

L'évaluation du dispositif expérimental est sous la responsabilité de la DEPP, du laboratoire EMC de l'Université de Lyon 2 et de l'IREDU. Sur les deux années de l'expérimentation, quatre évaluations des acquis des élèves, deux en GS (T1 et T2) et deux en CP (T3 et T4), ont été mises en œuvre. Afin d'estimer les effets du dispositif, la progression des élèves du groupe expérimental est comparée à celle des élèves d'un échantillon témoin, sélectionné aléatoirement en respectant certaines caractéristiques du groupe expérimental.

Pour l'année de GS, 59 écoles ont pris part à l'expérimentation. Cette première année, sur environ 6000 élèves visés - 3000 issus de classes expérimentales et environ 3000 de classes témoins tirées au sort - 3569 élèves ont participé aux deux moments d'évaluation (début et fin d'année scolaire). La non-réponse a plus particulièrement affecté l'échantillon témoin. Cependant, cette attrition différentielle ne semble pas avoir déformé l'échantillon selon les caractéristiques observables disponibles.

Pour l'année de CP, 28 écoles ont été concernées par l'expérimentation, 1 113 élèves ont participé aux évaluations de début et de fin de CP. Parmi eux, 728 élèves ont suivi le protocole expérimental en GS et en CP, et 369 élèves l'ont suivi en CP mais pas en GS. Parmi les 91 écoles tirées pour l'échantillon témoin, 78 écoles ont finalement participé aux deux moments de mesure, soit 3 534 élèves.

Afin de comparer les progressions des élèves selon les différentes épreuves, les scores ont été standardisés afin de tenir compte de leur dispersion, mesurée par l'écart-type. Ainsi, les progrès sont mesurés en les rapportant à l'écart-type des scores des élèves du groupe témoin.

Sur la première année de l'expérimentation qui a porté sur l'année de grande section de maternelle, deux modélisations ont été engagées pour estimer les effets moyens du dispositif : la première procède par régression linéaire, la seconde par matching.

Le modèle de régression porte sur chacun des scores observés au deuxième temps de mesure, en fonction des scores observés au premier temps ainsi que du sexe et de l'âge des élèves. Les élèves des classes expérimentales progressent davantage sur les dimensions Lettres (+ 17 %), Phonologie (+ 22 %), Lecture voie non lexicale (+ 41 %) et Mathématiques (+ 23 %). Pour ces domaines, les progrès sont statistiquement significatifs. En revanche, pour les dimensions Vocabulaire et Lecture voie lexicale, les effets ne sont pas significatifs. En compréhension, l'effet est légèrement significatif.

Étant donné les différences de performance observées entre les deux groupes lors du premier temps de mesure, une approche procédant par matching (appariement) a été engagée pour comparer avec l'approche par régression. La méthode employée ici et consistant à appairer sur les scores de la première session d'évaluation donne des résultats très proches de ceux de la régression : Lettres (+ 16 %), Phonologie (+ 25 %), Lecture voie non lexicale (+ 41 %) et Mathématiques (+ 31 %). En outre, la Compréhension ressort de manière significative dans ce cas (+ 15 %).

Une analyse complémentaire, selon les déciles et non plus en moyenne, montre des effets différents selon le niveau du score observé au deuxième temps de mesure. En phonologie, ces effets sont plus importants pour les élèves les plus faibles et en compréhension, ils ne sont significatifs que pour les élèves les plus faibles.

Sur la deuxième année de l'expérimentation, au CP, trois groupes d'élèves sont distingués : les élèves ayant expérimenté en GS et en CP, ceux ayant expérimenté en CP uniquement et les témoins. Deux modélisations ont été engagées pour estimer les effets moyens du dispositif (régression linéaire, la seconde par matching).

Le modèle de régression porte sur chacun des scores observés au T4, en fonction des scores observés au T3 sur les dimensions non travaillées spécifiquement dans le protocole expérimental, ainsi que du sexe et de l'âge des élèves. Rapportés à l'écart-type des scores de l'échantillon témoin, les scores des élèves ayant participé à l'expérimentation en GS et en CP augmentent davantage en phonologie (+ 29 %). Les élèves ayant participé à l'expérimentation en CP uniquement progressent davantage en phonologie également (+ 25 %) et en lecture voie non lexicale (+ 15 %). En revanche ils auraient moins progressé en compréhension orale (- 11 %). Les différences sur les autres dimensions ne sont pas significatives.

Pour compléter ces résultats, une approche procédant par matching est utilisée, en appariant les élèves selon les scores du T3 des dimensions non travaillées spécifiquement dans le protocole expérimental, ainsi que le sexe et l'âge. Selon les résultats obtenus, les élèves ayant participé à l'expérimentation en GS et en CP progressent davantage en phonologie (+ 34 %), mais auraient moins progressé en compréhension orale (- 14 %). Pour les élèves ayant participé à l'expérimentation en CP uniquement les résultats confirment qu'ils progressent davantage en phonologie (+ 27 %) et en lecture voie non lexicale (+ 17 %).

L'analyse complémentaire, selon les déciles et non plus en moyenne, précise des effets différents selon le niveau du score observé en fin de CP et notamment, en phonologie et en lecture voie non lexicale où les effets sont plus importants pour les élèves les plus faibles tandis qu'en compréhension orale, les effets négatifs concernent plus particulièrement les élèves de niveau moyen.

Enfin, une comparaison des scores moyens des élèves des groupes témoins du temps 2 et du temps 3, calculés sur les items communs à ces deux temps, montre que les élèves du groupe témoin de début de CP ont des scores supérieurs aux élèves du groupe témoin de fin de GS, de + 10 % en vocabulaire à + 84 % en phonologie. Cette différence pourrait s'expliquer par le fait que, l'évaluation ayant eu lieu au mois de novembre, toutes ces dimensions aient donc été travaillées par les élèves de CP : cette évolution correspondrait à une progression ordinaire entre fin de GS et début de CP.

En résumé, selon l'hypothèse que les deux échantillons témoins (de GS et de CP) sont comparables et qu'on n'observe pas d'évolution différentielle de la dispersion, l'effet global du dispositif sur deux années serait caractérisé par une progression significative en phonologie (+ 51 %), en lecture voie non lexicale (+ 48 %) avec des effets importants pour les élèves les plus faibles, et une absence d'effet sur les dimensions compréhension et lecture voie lexicale.

Des analyses complémentaires, statistiques et pédagogiques, viendront compléter ces résultats. En particulier, une évaluation menée en CE1 durant l'année 2013-2014 devrait permettre de déterminer si les effets observés en grande section et en CP sont transférés ou non en fin de CE1, sur des compétences de plus haut niveau, et en particulier sur les acquis en maîtrise de la langue, tel qu'attendus par les programmes scolaires et par le socle commun de connaissances et de compétences.



PLAN DU RAPPORT

INTRODUCTION GENERALE	10
1. L'EXPÉRIMENTATION	11
1.1. OBJECTIFS DU PROJET ET NATURE DE L'ACTION EXPÉRIMENTÉE	11
1.1.1. <i>Objectifs du projet</i>	11
1.1.2. <i>L'action expérimentée</i>	11
1.2. OBJECTIFS ET MODALITÉS DE L'ÉVALUATION	12
1.2.1. <i>Problématique et méthodologie mise en œuvre</i>	12
1.2.2. <i>Validité interne</i>	15
2. ENSEIGNEMENTS DE POLITIQUE PUBLIQUE	21
2.1. RÉSULTATS DE L'ÉVALUATION	21
2.1.1. <i>Les publics touchés par l'expérimentation</i>	21
2.1.2. <i>Les effets du dispositif expérimenté</i>	22
2.2. VALIDITÉ EXTERNE DES RÉSULTATS ET PERSPECTIVES DE GÉNÉRALISATION/ESSAIMAGE : CARACTÈRE TRANSFÉRABLE DU DISPOSITIF ET CHANGEMENT D'ÉCHELLE.	42
CONCLUSION GENERALE	45
BIBLIOGRAPHIE CONSTRUCTION DES EPREUVES T1	47
OUVRAGES	47
EPREUVE LETTRES	47
EPREUVE HABILITÉS PHONOLOGIQUES	47
EPREUVE COMPRÉHENSION DE RÉCIT	47
EPREUVE VOCABULAIRE	47
BIBLIOGRAPHIE GÉNÉRALE.....	48
ANNEXE 1 : PROTOCOLE D'ENQUÊTE.....	50
ANNEXE 2 : SYNTHÈSE D'ENQUÊTE	56
ANNEXE 3 : ÉCHANTILLONNAGE - STRATIFICATION	60
ANNEXE 4 : ANALYSE DES ITEMS T1T2.....	62
ANNEXE 5 : DISTRIBUTION DES SCORES T1T2	66
ANNEXE 6 : ANALYSE DES ITEMS T3T4.....	71
ANNEXE 7 : DISTRIBUTION DES SCORES T3T4	75



INTRODUCTION GENERALE

Dans ses conclusions du 11 mai 2010 sur la dimension sociale de l'éducation et de la formation, le Conseil de l'Union européenne a invité les États membres notamment en ce qui concerne l'enseignement préscolaire et scolaire :

- à garantir un accès plus large à des services d'enseignement et d'accueil préscolaires de qualité élevée, afin de permettre à tous les enfants, en particulier ceux issus de milieux défavorisés ou ayant des besoins pédagogiques spécifiques, de prendre un bon départ et de les motiver à apprendre;
- à améliorer la qualité des services offerts dans les écoles en vue de faire échec aux risques de marginalisation socioéconomique ou culturelle;
- à accorder toute l'attention voulue à l'acquisition des compétences de base, en particulier la lecture et le calcul.

Ces orientations ont été mises en œuvre en France au travers notamment de la priorité donnée à la maîtrise des savoirs fondamentaux. Au-delà de l'effort fait pour la lecture et l'écriture (au cœur du plan de prévention de l'illettrisme), l'attention est portée sur la grammaire, l'orthographe et le vocabulaire ainsi que l'acquisition des automatismes en mathématiques.

Les programmes entrés en vigueur en septembre 2008 insistent sur ces éléments des apprentissages.

La maîtrise des éléments du socle à la fin de la scolarité obligatoire impose d'adapter les approches pédagogiques les plus efficaces en réponse aux besoins spécifiques de chaque élève et particulièrement ceux en difficulté scolaire.

Le recours à des programmes structurés d'apprentissage s'appuyant sur les connaissances des sciences cognitives a constitué une orientation du ministère qui souhaitait dans l'ensemble des écoles la mise en place, pour le cycle 2, d'actions favorisant la progressivité des enseignements et l'acquisition des compétences fondamentales pour la lecture (notamment la phonologie et la compréhension.)

Complémentaire de ces initiatives qu'a engagées le ministère, l'extension des champs d'expérimentation des principes du programme PARLER permettrait de tirer les fruits des enseignements de recherche, d'expérimentation et d'évaluation de l'expérience conduite dans l'Académie de Lyon.

En regard des travaux réalisés dans l'académie de Lyon, l'expérimentation a été l'occasion de compléter les analyses et notamment en grande section.

Dans ce contexte, en 2011, la DGESEO a donc proposé au Fonds d'expérimentations pour la jeunesse de financer un Programme d'Extension territoriale et pédagogique (intégration de la grande section) du travail de mise en place d'outils pédagogiques simples à mettre en

œuvre pour l'apprentissage de la lecture dans un petit groupe d'élèves. Ce projet est porté par l'association « Agir pour l'école ».

Intitulé « **Lecture** », ce projet a pour objectif de mesurer l'efficacité de la diffusion de pratiques pédagogiques préconisées par la recherche pour trouver des solutions contre la difficulté scolaire et en particulier l'illettrisme.

L'évaluation du dispositif expérimental est sous la responsabilité de la DEPP, du laboratoire EMC de l'Université de Lyon 2 (Jean Ecalte) et de l'IREDU (Bruno Suchaut). Elle fait l'objet d'une convention quadripartite entre la DEPP, l'université de Lyon 2, l'IREDU et l'association Agir pour l'école.

1. L'EXPÉRIMENTATION

1.1. Objectifs du projet et nature de l'action expérimentée

1.1.1. Objectifs du projet

Le protocole de l'expérimentation est conçu à partir du principe d'enseignement explicite. Dans l'objectif de lutter contre l'échec scolaire, le protocole préconise la mise en œuvre d'interventions précoces dans la scolarité des élèves. Les grands principes de cette expérimentation sont des adaptations au contexte français d'un rapport de 2000 rédigé par le *National Reading Panel* qui identifiait les interventions efficaces favorisant l'apprentissage de la lecture, à partir de méta-analyses à grande échelle.

L'expérimentation Lecture se situe dans la continuité du programme PARLER (M. Bianco, P. Bressoux, C. Lequette, G. Pouget, Laboratoire des Sciences de l'Éducation, Université Pierre-Mendès-France de Grenoble, étude menée par Michel Zorman) qui a été conduit de 2005 à 2008 dans plusieurs écoles de REP de Grenoble et de l'agglomération grenobloise et qui a montré des effets positifs.

1.1.2. L'action expérimentée

Dans les circonscriptions identifiées, des classes participent à l'expérimentation sur la base du volontariat ou sur décision des directions académiques et des inspections de l'éducation nationale. Le protocole propose un guide à l'enseignant pour mettre en œuvre, de la grande section au CE1, un enseignement structuré des compétences qui sont identifiées comme nécessaires à l'apprentissage de la lecture : la conscience phonologique, la compréhension orale, le code alphabétique, et la fluence de lecture en particulier. Des documents de formation précisent les exigences du protocole quant au cadre des séances, à la conduite de chaque séance et à la progression annuelle à respecter. Le matériel pédagogique est mis à disposition des enseignants sur une plateforme en ligne. Les conseillers pédagogiques de circonscription assurent le suivi, la formation et l'accompagnement des enseignants engagés dans le dispositif. La méthode promeut le travail des élèves en petits groupes de niveau homogène, chaque séance d'entraînement durant 30 minutes. La fréquence des entraînements est augmentée pour les élèves les plus faibles.

Les pratiques pédagogiques habituelles en classes de grande section, de CP et de CE1 sont peu connues, notamment en ce qui concerne le temps et les modalités du travail de la lecture avec les élèves les plus faibles. Il est donc malaisé de définir l'apport du dispositif par rapport au droit commun, caractérisé par son hétérogénéité. L'ambition du projet est d'aligner les pratiques vers les préconisations issues de projets de recherche expérimentale précurseurs en orientant l'enseignement vers les compétences les plus déterminantes de l'apprentissage de la lecture et en guidant les enseignants vers une prise en compte systématique des élèves faibles en les prenant en charge pendant un volume horaire et selon des modalités bien précises.

1.2. Objectifs et modalités de l'évaluation

1.2.1. Problématique et méthodologie mise en œuvre

a. Objectifs de l'évaluation

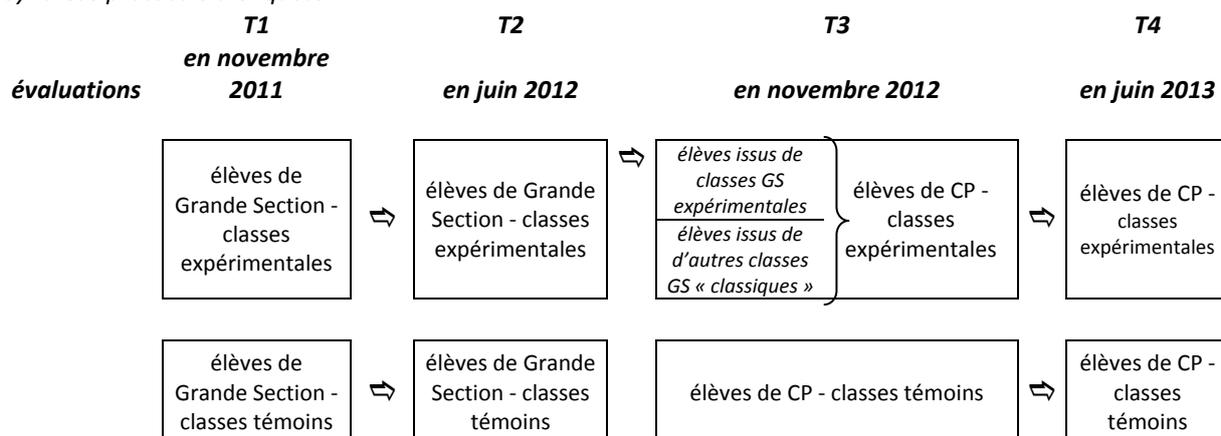
Le dispositif d'évaluation doit permettre de répondre à la question suivante : Le programme permet-il réellement de réduire la grande difficulté scolaire ? En outre, il doit permettre de contrôler l'absence d'impacts négatifs sur les autres disciplines (en particulier les mathématiques).

Le recueil de l'information s'est fait via des tests de type papier-crayon en passation collective ou semi-collective. Les domaines retenus sont le vocabulaire, la reconnaissance des lettres, la compréhension de récit et les habiletés phonologiques (tâche de détection d'intrus, tâche de suppression de syllabes), ainsi que la lecture de mots (voie lexicale), la lecture de pseudo-mots (voie non lexicale) et des éléments de mathématiques.

Pour l'année 2011-2012, deux moments de mesure des compétences et des connaissances des élèves de grande section de maternelle (GS) ont été organisés, le premier en novembre et le second en juin. La progression sur l'année des élèves du groupe expérimental est comparée à celle des élèves du groupe témoin, afin d'estimer les effets du dispositif.

Comme l'année précédente, pour l'année 2012-2013, deux moments de mesure des compétences et connaissances des élèves au CP ont été programmés, en début et en fin d'année scolaire.

Synthèse protocole d'enquête :



Le rôle du laboratoire de Lyon2 a consisté à produire des épreuves originales issues des travaux de recherche fondamentale conduits au sein de l'équipe Apprentissage, Développement et Troubles du Langage du laboratoire EMC. Ces épreuves sont basées sur les prédicteurs de réussite en lecture dans les deux composantes, identification de mots écrits (IME) et compréhension (CL). Compte tenu du temps limité disponible pour l'évaluation dans les classes, le choix des domaines investigués concernaient pour l'IME, les habiletés phonologiques, les connaissances sur les lettres, le décodage et la lecture de mots; pour la CL, le vocabulaire et la compréhension orale. Les épreuves et les tâches évoluaient avec l'avancée dans l'étude au cours des passations successives.

b. Instruments construits pour l'évaluation du projet Lecture en novembre 2011 en grande section de maternelle (T1 GS).

Quatre domaines sont investigués, la connaissance des lettres, les habiletés phonologiques, la compréhension orale de récit et le vocabulaire. Ils constituent de solides prédicteurs des capacités de lecture-écriture tant en lecture de mots qu'en compréhension de lecture (voir Ecalle & Magnan, 2010).

c. Instruments utilisés pour l'évaluation du projet Lecture en en juin 2012 en grande section de maternelle (T2 GS).

La lecture est une activité complexe dont la maîtrise implique la mise en œuvre simultanée de deux principales activités (Gough & Turner, 1986¹) : l'identification des mots écrits et la compréhension de la signification de ces mots. Aussi, plusieurs compétences sous-tendent l'efficacité de la mise en place de ces deux activités. D'une part, la connaissance de lettres, les habiletés phonologiques sont liées au développement des processus d'identification et de production de mots écrits (e.g., Puolakanaho et al., 2007²). D'autre part, le niveau de vocabulaire et la compréhension orale de textes sont reliés à la compréhension en lecture

¹ Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6-10.

² Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Lepänen, P.H.T., Poikkeus, A.M., Tolvanen, A., Torppa, M., & Lyytinen, H. (2007). Very early phonological and language skills: Estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry*, 48(9), 923-931.

(e.g., Cain & Oakill, 2007³). Précisément, ces compétences sont reconnues comme de solides prédicteurs de la réussite lors de l'apprentissage de la lecture. Par conséquent, un faible niveau dans ces compétences prédictives suggère des risques éventuels de difficultés ultérieures dans le développement des capacités de lecture.

Pour ce deuxième moment de mesure, six domaines sont investigués, la connaissance des lettres, les habiletés phonologiques, la compréhension orale de récit et le vocabulaire, la lecture de mots et de pseudo-mots et des éléments de mathématiques (à but de contrôle des éventuels effets négatifs).

d. Instruments utilisés pour l'évaluation du projet Lecture en novembre 2012 en classe de Cours Préparatoire (T3 CP).

Pour ce troisième moment de mesure, six domaines sont évalués, la connaissance des lettres, les habiletés phonologiques, la compréhension orale de récit et le vocabulaire, la lecture de mots et de pseudo-mots et des éléments de mathématiques. Afin d'assurer un ancrage, certaines épreuves constituent des reprises du temps de mesure précédent.

Communications scientifiques

- Note d'étape annuelle sur l'évaluation – DEPP, Lyon 2, IREDU. Cette note présente une synthèse des résultats obtenus sur la première cohorte en GS. Elle est disponible sur le site www.experimentation.jeunes.gouv.fr/1048-%E2%9C%B1-apprentissage-de-la-lecture-en-cycle-2
- Revue française de pédagogie : article accepté, à paraître en avril 2014. Cet article écrit par les évaluateurs du projet fait une analyse des évaluations initiales de la première cohorte.
Labat, H., Farhat, S.-L., Andreu, S., Rocher, T., Cros, L., Magnan, A. & Ecalle, J. (2014). *Évaluation des connaissances précoces prédictives de l'apprentissage de la lecture en grande section de maternelle. Revue Française de Pédagogie.*
- Labat, H., Farhat, S.-L., Potocki, A., Ecalle, J., & Magnan, A. (2012). Évaluation des connaissances précoces en lecture en début de Grande Section de Maternelle. Communication orale présentée au 54ème Congrès de la SFP. Montpellier, 3-5 septembre.
- Revue ANAE (Approche neuropsychologique des apprentissages chez l'enfant) Synthèse des recherches qui inspirent l'association et des résultats sur la première cohorte en GS.
La Pédagogie préventive à l'école, les avancées actuelles et à venir, A. Bougnères, L. Cros

³ Cain, K., & Oakill, J. (2007). *Children's comprehension problems in oral and written language. A cognitive perspective.* New York : Guilford press.

1.2.2. Validité interne

a. Comparabilité du groupe bénéficiaire et du groupe témoin

Pour l'année 2011-2012, l'échantillon se compose des élèves des classes engagées dans l'expérimentation, et des élèves de 59 écoles maternelles témoins tirées au sort selon la même répartition que les 59 écoles engagées dans l'expérimentation (département, éducation prioritaire, taille de l'école).

Pour l'année 2012-2013 les élèves du dispositif expérimental sont en CP et 28 écoles sont engagées dans l'expérimentation. 90 écoles témoins ont été tirées au sort selon la même répartition que les écoles engagées (département, éducation prioritaire, résultats aux évaluations CE1). La population visée était de 6 500 élèves au total.

Dans les écoles engagées dans l'expérimentation, deux types de population sont identifiés : les élèves ayant bénéficié l'année précédente du dispositif, et ceux entrant nouvellement dans le dispositif (élèves n'ayant pas bénéficié du dispositif en GS l'année précédente). En revanche, il n'y a pas de suivi des élèves du groupe témoin entre la GS et le CP.

Tirage de l'échantillon témoin T1T2

La base de sondage comporte l'ensemble des écoles publiques des départements 59, 62, 92 et 93, dont l'effectif de GS en 2010 est supérieur ou égal au minimum observé en 2009 dans les écoles du groupe expérimental, soit 20 élèves.

On n'a conservé que les circonscriptions dans lesquelles il n'y a pas d'écoles du groupe expérimental.

L'échantillon est tiré selon un plan de sondage stratifié. On utilise de l'information auxiliaire pour découper la population en sous-ensembles homogènes (les strates) et réaliser un sondage dans chacune d'elles. Cela permet de mieux répartir l'échantillon dans ces sous-population et d'augmenter la précision des estimateurs.

La stratification est faite selon 3 dimensions :

- éducation prioritaire ou non ;
- département géographique (4 départements) ;
- taille de l'école.

On a tiré 59 écoles témoins selon la même répartition que les 59 expérimentales selon le croisement entre trois caractéristiques : le département, la zone d'éducation (prioritaire ou non) et la taille (nombre d'élèves de GS inférieur à 40, compris entre 40 et 60, supérieur à 60).

Tirage de l'échantillon témoin T3T4

La base de sondage comporte l'ensemble des écoles publiques des départements 59, 62, 92 et 93, dont l'effectif de CP en 2011 est supérieur ou égal au minimum observé en 2011 dans les écoles du groupe expérimental, soit 20 élèves.

On n'a conservé que les circonscriptions dans lesquelles il n'y a pas d'écoles du groupe expérimental. Cela représente 1 656 écoles.

Ayant souhaité utiliser les scores des évaluations CE1 2011 pour la stratification, on a supprimé également les écoles pour lesquelles nous n'avions pas ces données (soit 105 écoles, il reste donc 1 551 écoles).

La stratification est faite selon 3 dimensions :

- éducation prioritaire ou non ;
- département géographique (4 départements) ;
- score total moyen aux évaluations CE1 2011 (4 groupes constitués selon les quartiles observés dans la base obtenue).

On obtient donc 32 strates. [cf. annexe 3. tableau 2]

On ne conserve dans la base de sondage que les strates pour lesquelles il y a au moins une école du groupe expérimental, soit 15 strates.

Finalement, la base de sondage comporte 751 écoles. [cf. annexe 3 tableau 3]

On tire dans cette base un échantillon témoin de 90 écoles, réparties dans les strates proportionnellement au groupe expérimental. Ces 90 écoles représentaient 4 531 élèves en 2011. [cf. annexe 3 tableau 4]

b. Données T1T2

Participation T1T2

Après nettoyage et concaténation des données, il ressort que 4 468 élèves de GS ont participé au temps 1 et 3 885 élèves au temps 2. Au final, **3 569** élèves ont participé aux deux sessions d'évaluation.

	Nb élèves
T1 seul	899
T1 et T2	3 569
T2 seul	316
total	4 784
total T1	4 468
total T2	3 885

Attrition T1T2

On s'intéresse ici aux 4 468 élèves présents au temps 1.

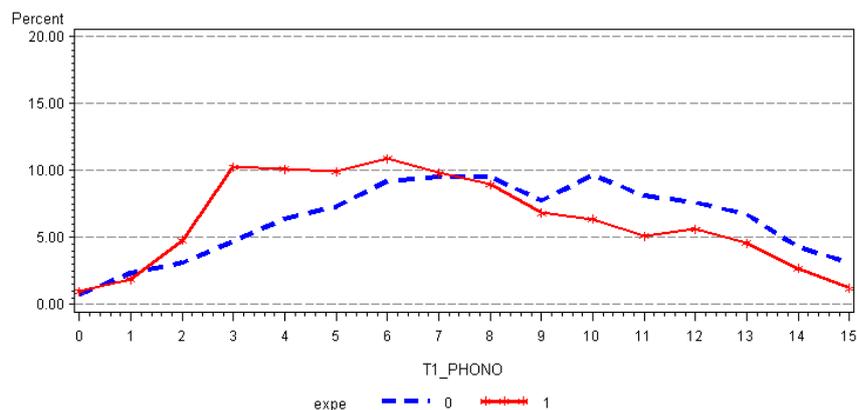
Parmi ceux-ci, 3 569 ont également participé au temps 2 et 899 n'y ont pas participé.

Au temps 1, les élèves du groupe expérimental présentent des scores moyens inférieurs à ceux du groupe témoin (cf. l'exemple de la phonologie avec le graphique ci-dessous). Cela est vrai aussi bien pour les élèves qui ont ensuite participé au temps 2 que pour ceux qui n'y ont pas participé. Les élèves témoins qui n'ont pas participé au temps 2 ont des scores moyens légèrement inférieurs à ceux des élèves témoins qui ont participé aux deux temps de mesure, mais cette différence n'est pas statistiquement significative.

Au vu des caractéristiques démographiques des différents groupes, expérimental ou témoin, ayant participé au deux temps de mesure ou seulement au premier, on constate qu'il y a beaucoup de valeurs manquantes concernant le sexe et l'âge des élèves n'ayant participé qu'au temps 1 (aussi bien chez les élèves du groupe expérimental que chez les élèves témoins). La répartition selon ces variables démographiques des élèves ayant participé aux deux temps de mesure est proche de celle observée sur l'ensemble des élèves ayant participé au temps 1, avec une proportion un peu plus importante de garçons, et une proportion équivalente d'élèves nés avant juin 2006 ou nés à partir de juin 2006.

En conséquence, l'attrition ne semble pas avoir déformé l'échantillon selon ces caractéristiques observables qui seront néanmoins intégrées comme variables de contrôle dans les analyses, et ne semble pas non plus différentielle selon le groupe, expérimental ou témoin.

Distribution des scores en phonologie à T1 en fonction du groupe



Note de lecture : en rouge figure la distribution des scores en phonologie observée à T1 pour le groupe expérimental. En bleu, il s'agit de la même distribution pour le groupe témoin

Comparaison des caractéristiques démographiques T1T2

	GROUPE EXPE		GROUPE TEMOIN		ENSEMBLE	
	élèves	écoles	élèves	écoles	élèves	écoles
échantillon prévu	environ 3 000	59	environ 3 000	59	environ 6 000	118
participation temps1	2 564	55	1 904	38	4 468	93
<i>sexe - % garçons</i>	47.7 %		48.7 %		48.1 %	
<i>sexe - % filles</i>	43.6 %		44.9 %		44.1 %	
<i>sexe - % NC</i>	8.8 %		6.4 %		7.8 %	
<i>âge - % né avant juin 2006</i>	39.9 %		42.6 %		41.1 %	
<i>âge - % né à partir de juin 2006</i>	39.5 %		42.9 %		41.0 %	
<i>âge - % NC</i>	20.5 %		14.5 %		17.9 %	
	GROUPE EXPE		GROUPE TEMOIN		ENSEMBLE	
	élèves	écoles	élèves	écoles	élèves	écoles
dont temps1 et temps2	2 067	48	1 502	32	3 569	80
<i>sexe - % garçons</i>	50.9 %		50.5 %		50.8 %	
<i>sexe - % filles</i>	45.7 %		45.5 %		45.6 %	
<i>sexe - % NC</i>	3.4 %		4.0 %		3.6 %	
<i>âge - % né avant juin 2006</i>	44.5 %		44.2 %		44.4 %	
<i>âge - % né à partir de juin 2006</i>	44.7 %		44.0 %		44.4 %	
<i>âge - % NC</i>	10.9 %		11.8 %		11.3 %	
dont temps1 uniquement	497	42	402	31	899	73
<i>sexe - % garçons</i>	34.0 %		42.0 %		37.6 %	
<i>sexe - % filles</i>	34.8 %		42.5 %		38.3 %	
<i>sexe - % NC</i>	31.2 %		15.4 %		24.1 %	
<i>âge - % né avant juin 2006</i>	21.1 %		36.8 %		28.1 %	
<i>âge - % né à partir de juin 2006</i>	18.3 %		38.6 %		27.4 %	
<i>âge - % NC</i>	60.6 %		24.6 %		44.5 %	

Note de lecture :

2 564 élèves du groupe expérimental ont participé au temps 1, dont 2 067 élèves qui ont participé aux temps1 et temps2, et 497 qui n'ont participé qu'au temps1.

Parmi les 2 067 élèves du groupe expérimental qui ont participé aux temps1 et temps2, ils sont 44.5% à être nés avant juin 2006 (5 ans révolus au moment de la passation du temps2), 44.7% à être nés à partir de juin 2006 (pas encore 5 ans au moment de la passation du temps2), et 10.9% pour lesquels l'âge n'est pas connu.

Comparaison des scores du temps 1

Les scores sont standardisés, de moyenne 0 et d'écart-type 1 pour l'ensemble des élèves du groupe témoin ayant participé au temps 1.

Groupe expé	Ensemble des élèves présents au Temps 1			Elèves présents aux temps 1 et 2			Elèves présents au temps 1 uniquement		
	Effectif	moyenne	Ecart-type	Effectif	moyenne	Ecart-type	Effectif	moyenne	Ecart-type
Vocabulaire	2 507	-0.27	1.18	2 021	-0.18	1.13	486	-0.65	1.30
Lettres	2 512	-0.18	1.14	2 028	-0.19	1.15	484	-0.16	1.12
Compréhension	2 495	-0.07	1.01	2 011	-0.03	1.01	484	-0.26	0.97
Phonologie	2 451	-0.33	0.95	1 978	-0.28	0.97	473	-0.54	0.85
Groupe témoin	Effectif	moyenne	Ecart-type	Effectif	moyenne	Ecart-type	Effectif	moyenne	Ecart-type
Vocabulaire	1 839	0.00	1.00	1 457	0.04	0.99	382	-0.15	1.03
Lettres	1 819	0.00	1.00	1 439	0.04	0.97	380	-0.16	1.10
Compréhension	1 786	0.00	1.00	1 415	0.02	1.01	371	-0.07	0.97
Phonologie	1 689	0.00	1.00	1 357	0.05	1.00	332	-0.20	0.96

Effet de l'interaction groupe-participation T1T2

Effet de l'interaction groupe (expé/témoin) – participation (t1t2 ou t1).

Référence = témoin t1t2

	Vocabulaire			Lettres		
	Coefficient	Erreur standard	probabilité	Coefficient	Erreur standard	probabilité
Intercept	0.04	0.07	0.55	0.04	0.05	0.35
Témoin t1	-0.19	0.11	0.10	-0.20	0.13	0.13
Expe t1	-0.68	0.18	0.00	-0.20	0.12	0.10
Expe t1t2	-0.22	0.10	0.04	-0.23	0.07	0.00
	Compréhension			Phonologie		
	Coefficient	Erreur standard	probabilité	Coefficient	Erreur standard	probabilité
Intercept	0.02	0.07	0.78	0.05	0.07	0.51
Témoin t1	-0.09	0.12	0.46	-0.25	0.14	0.07
Expe t1	-0.28	0.13	0.03	-0.59	0.12	0.00
Expe t1t2	-0.05	0.10	0.61	-0.33	0.10	0.00

NB : seuls les chiffres en gras sont statistiquement significatifs

Scores (calculs et distributions) T1T2

Les scores calculés sont la somme des indicatrices de réussites aux items retenus pour l'analyse.

Les distributions des scores bruts montrent qu'au temps 1 les séquences de compréhension orale et de phonologie sont bien adaptées au niveau des élèves évalués, tandis que les séquences de vocabulaire et surtout de connaissance des lettres ont tendance à saturer. Pour la plupart des séquences, les distributions des deux groupes sont assez proches, bien que celle du groupe expérimental soit un peu moins bonne que celle du groupe témoin, en particulier en phonologie.

Au temps 2 les épreuves de lecture par voie lexicale et non lexicale sont adaptées au niveau des élèves, voire un peu difficiles, ce qui est nécessaire dans la mesure où ce sont des items qui seront amenés à être repris lors des sessions d'évaluation suivantes en début et fin de CP.

c. Données T3T4

Participation T3T4

	T3 total	T3 uniquement	T3 et T4	T4 uniquement	T4 total
Témoins	4 051	517	3 534	84	3 618
Expé GS et CP	769	41	728	6	734
Expé CP uniquement	392	23	369	31	400
Expé GS uniquement	3	0	3	0	3
Expé ni GS ni CP	13	0	13	0	13
Total	5 228	581	4 647	121	4 768

Au sein des écoles expé, 1 113 élèves ont participé aux évaluations des temps 3 (début de CP) et 4 (fin de CP).

Parmi eux, 728 ont suivi le protocole expérimental en GS et en CP, 369 l'ont suivi en CP mais pas en GS. 16 élèves auraient du le suivre en CP mais leur classe ne l'a finalement pas suivi, dont 3 élèves qui l'avaient suivi en GS et 13 élèves qui ne l'ont finalement jamais suivi.

Analyse des items T3T4 (cf. Annexe 6)

Des analyses psychométriques classiques ont permis de détecter des items présentant des dysfonctionnements. Ces items ont été éliminés avant de procéder au calcul des scores, notamment ceux évaluant la fluence dont l'épreuve a été écartée car ils ne présentaient pas les caractéristiques nécessaires.

Scores (calculs et distributions) T3T4

Les scores calculés sont la somme des indicatrices de réussites aux items retenus pour l'analyse.

Distribution des scores bruts T3T4 (cf. Annexe 7)

Les distributions des scores bruts montrent que la plupart des séquences sont plutôt bien réussies dès le début de CP, mises à part les séquences relatives à la lecture voie lexicale et la phonologie qui sont plus adaptées au niveau des élèves.

2. ENSEIGNEMENTS DE POLITIQUE PUBLIQUE

2.1. Résultats de l'évaluation

2.1.1. Les publics touchés par l'expérimentation

a. Caractéristiques des bénéficiaires de l'expérimentation

Les bénéficiaires du projet sont les élèves des classes sélectionnées ou volontaires de grande section de maternelle et de CP. Il s'agit donc du public habituel de l'éducation nationale, avec toutefois une surreprésentation des écoles situées dans des zones d'éducation prioritaire. Les élèves concernés sont en grande majorité dans des quartiers défavorisés. Sur les 59 écoles concernées par le projet en 2011-2012,

- 24 sont hors éducation prioritaire (à Boulogne-Billancourt, Lille et au Puy-en-Velay),
- 13 en RRS (réseau de réussite scolaire) et 22 en RAR (réseau ambition réussite), en Seine-Saint-Denis (Clichy-sous-Bois, Montfermeil, la Courneuve), à Lille et à Calais, et à Villeneuve-la-Garenne.

Les années suivantes, cette proportion est restée stable. Le public visé est donc moins performant en lecture et dans l'ensemble des compétences scolaires et cognitives visées que la moyenne nationale, et le nombre d'élèves en difficulté y est plus élevé.

Les enseignants concernés par le projet semblent présenter des caractéristiques spécifiques. Un travail de recherches est en cours, il est mené par un doctorant de l'École d'Économie de Paris (Adrien Bouguen), bénéficiant d'une convention avec le porteur du projet et les évaluateurs. Le projet de recherche porte précisément sur ces questions d'effets enseignants.

b. Adhésion du public ciblé au dispositif

Le niveau d'adhésion du public au dispositif doit s'apprécier au regard de certaines caractéristiques, en particulier celles liées au concept de liberté pédagogique. Cette notion définie en 2005 dans la loi Fillon – « *la liberté pédagogique de l'enseignant s'exerce dans le respect des programmes et des instructions [...] et dans le cadre du projet d'école ou d'établissement avec conseil et sous contrôle des membres des corps d'inspection* » Art. L. 912-1-1 – garantit à chaque enseignant le droit de choisir les pratiques pédagogiques qui lui semblent les plus appropriées mais dont la portée varie selon les circonscriptions et les enseignants et l'implication variable des inspections de circonscriptions et directions académiques dans le projet. Ces éléments ont conduit à une appropriation très variable du projet selon les circonscriptions, sans que le porteur de projet n'ait ni le souhait, ni le rôle ni la possibilité de susciter l'adhésion des enseignants.

C'est ainsi que l'effectif a considérablement chuté entre la première et la deuxième année, alors même que les enseignants engagés en grande section furent (et restent encore aujourd'hui) nombreux à continuer à mettre en œuvre le projet. 66 classes de CP seulement furent engagées, au lieu des 174 visées.

Les facteurs de refus du protocole par les enseignants ont été identifiés par le porteur de projet par sondage auprès des enseignants (volontaires et non volontaires). Il s'agit principalement

- de la « lourdeur du dispositif » (lequel exige en effet que l'enseignant consacre au moins 90 minutes par jour à un protocole très formalisé) ;
- et de « l'aspect mécanique » du protocole qui vise en effet l'automatisation de compétences cognitives par la répétition d'entraînements parfois rébarbatifs aux yeux des enseignants.

Ces deux caractéristiques du programme ne sont toutefois pas modifiables dans la perspective d'envisager une diffusion plus large ; elles sont en effet dictées par la recherche expérimentale dans le domaine et sont vraisemblablement parmi les conditions d'efficacité du programme.

Parmi les enseignants qui se sont engagés, une certaine variabilité a par ailleurs été observée par le porteur de projet qui a porté une attention particulière à la mesure du degré d'appropriation du programme par les enseignants.

2.1.2. Les effets du dispositif expérimenté

a. Effets de l'expérimentation sur le public bénéficiaire en GS (T1-T2)

On s'intéresse aux 3 569 élèves présents aux deux temps de mesure.

Statistiques descriptives des scores

Les scores sont standardisés, de moyenne 0 et d'écart-type 1 pour le groupe témoin.

Temps1	ensemble			groupe expé			groupe témoin		
	Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
Vocabulaire	3 478	-0.13	1.09	2 021	-0.22	1.14	1 457	0.00	1.00
Lettres	3 467	-0.14	1.12	2 028	-0.24	1.19	1 439	0.00	1.00
Compréhension	3 426	-0.03	1.00	2 011	-0.05	1.01	1 415	0.00	1.00
Phonologie	3 335	-0.20	0.99	1 978	-0.33	0.97	1 357	0.00	1.00

Temps2	ensemble			groupe expé			groupe témoin		
	Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
Vocabulaire	3 443	-0.13	1.13	2 008	-0.22	1.21	1 447	0.00	1.00
Lettres	3 408	0.03	0.98	1 971	0.04	0.96	1 449	0.00	1.00
Compréhension	3 365	-0.01	0.97	1 944	-0.01	0.95	1 431	0.00	1.00
Phonologie	3 319	0.04	0.99	1 932	0.06	0.98	1 395	0.00	1.00
Lecture voie lexicale	3 320	-0.04	0.98	1 907	-0.07	0.97	1 422	0.00	1.00
Lecture voie non lexicale	3 317	0.15	1.01	1 913	0.26	1.00	1 412	0.00	1.00
Mathématiques	3 382	0.05	1.00	1 981	0.09	0.99	1 410	0.00	1.00

Valeurs manquantes

Sur les 3 569 élèves présents aux temps 1 et 2, 2 709 élèves n'ont aucune valeur manquante sur l'ensemble des séquences.

Les valeurs manquantes à l'un des scores du temps 1 sont imputées de manière à préserver les distributions des variables ainsi que leurs relations entre elles.

Régression

Effet du groupe (expé/témoin) sur les scores du temps 2 en contrôlant le niveau au temps 1, l'âge et le sexe (avec cluster école).

Les valeurs manquantes à l'une des séquences du temps 1 (N = 338) ont été imputées. Les modélisations avec et sans imputations donnent des résultats très proches. Au temps 2, on observe des valeurs manquantes aux différents scores analysés (N = de 100 à 226).

	Vocabulaire T2			Lettres T2		
	Coefficient	Erreur standard	Probabilité	Coefficient	Erreur standard	Probabilité
Constante	0.06	0.06	0.33	0.01	0.04	0.79
Groupe expérimental	-0.07	0.08	0.42	0.17	0.04	0.00
Vocabulaire T1	0.47	0.03	0.00	0.07	0.03	0.02
Lettres T1	0.05	0.02	0.01	0.49	0.03	0.00
Compréhension T1	0.12	0.02	0.00	-0.02	0.02	0.41
Phonologie T1	0.11	0.02	0.00	0.02	0.02	0.18

	Compréhension T2			Phonologie T2		
	Coefficient	Erreur standard	Probabilité	Coefficient	Erreur standard	Probabilité
Constante	0.06	0.06	0.33	0.03	0.06	0.58
Groupe expérimental	0.10	0.06	0.11	0.22	0.07	0.00
Vocabulaire T1	0.19	0.02	0.00	0.15	0.02	0.00
Lettres T1	0.03	0.02	0.16	0.14	0.02	0.00
Compréhension T1	0.33	0.02	0.00	0.13	0.02	0.00
Phonologie T1	0.15	0.02	0.00	0.30	0.02	0.00

	Lecture voie lexicale T2			Lecture voie non lexicale T2		
	Coefficient	Erreur standard	Probabilité	Coefficient	Erreur standard	Probabilité
Constante	0.09	0.06	0.14	0.04	0.06	0.50
Groupe expérimental	0.04	0.06	0.53	0.41	0.07	0.00
Vocabulaire T1	0.02	0.02	0.43	0.05	0.02	0.06
Lettres T1	0.12	0.02	0.00	0.28	0.02	0.00
Compréhension T1	0.02	0.03	0.47	0.03	0.02	0.16
Phonologie T1	0.25	0.03	0.00	0.26	0.02	0.00

	Mathématiques T2		
		Erreur	
	Coefficient	standard	Probabilité
Constante	0.13	0.06	0.03
Groupe expérimental	0.23	0.07	0.00
Vocabulaire T1	0.12	0.02	0.00
Lettres T1	0.21	0.02	0.00
Compréhension T1	0.12	0.02	0.00
Phonologie T1	0.22	0.02	0.00

NB : seuls les chiffres en gras sont statistiquement significatifs

Note de lecture :

La régression du score de phonologie au temps 2 en fonction du groupe (expé/témoin), du niveau au temps 1 en vocabulaire, lettres, compréhension et phonologie, ainsi que de l'âge et du sexe (coefficients non présentés dans le tableau) estime une différence de +0.22 (soit 22% d'écart-type) en moyenne pour les élèves du groupe expérimental par rapport aux élèves du groupe témoins, à niveau au temps 1, sexe et âge équivalent. Ce coefficient est statistiquement significatif (probabilité inférieure à 0.05).

Matching

Le matching permet de construire des estimateurs pour évaluer l'effet de l'expérimentation en comparant les résultats des élèves du groupe expérimental avec ceux d'élèves similaires appartenant au groupe témoin. Pour assurer la correspondance entre les élèves du groupe expérimental et ceux du groupe témoin, on utilise le score de propension, qui indique la probabilité d'appartenir au groupe expérimental conditionnellement aux scores obtenus au temps 1.

Les valeurs manquantes à l'une des séquences du temps 1 (N = 338) ont été éliminées. *Genetic Matching* sur score de propension (pas de cluster école).

Temps 2	ATE	Erreur standard	T	Probabilité	Ecart-type groupe témoin	Effet du traitement standardisé
Vocabulaire	0.00	0.09	0.05	0.96	2.16	0.00
Lettres	0.31	0.08	4.05	0.00	1.99	0.16
Compréhension	0.37	0.10	3.63	0.00	2.52	0.15
Phonologie	0.82	0.12	6.65	0.00	3.33	0.25
Lecture voie lexicale	0.03	0.07	0.43	0.66	1.72	0.02
Lecture voie non lexicale	1.10	0.10	10.80	0.00	2.68	0.41
Mathématiques	0.69	0.09	7.77	0.00	2.22	0.31

ATE (average treatment effect) : Effet moyen du traitement

Effet du traitement standardisé : Les effets sont rapportés à l'écart-type du groupe témoin.

NB : seuls les chiffres en gras sont statistiquement significatifs

Note de lecture :

Le matching estime un effet moyen de l'expérimentation (ATE) de 0.82 en phonologie. Ce coefficient, rapporté à l'écart-type du score de phonologie du groupe témoin au temps 2, correspond à un effet de +25% d'écart-type. Il est statistiquement significatif (probabilité inférieure à 0.05).

Synthèse

La régression contrôle de l'âge, du sexe et des scores aux temps 1 avec une estimation tenant compte du niveau école. Le *matching* repose sur les seuls scores aux temps 1 et l'estimation de l'effet n'intègre pas le niveau école. En toute rigueur, il resterait à développer un *matching* tenant compte du niveau école.

Les résultats des deux premières analyses (régression et *matching*) sont proches et montrent un effet significatif dans les dimensions lettres, phonologie, lecture voie non lexicale et mathématiques.

	Effet moyen de l'expérimentation estimé par régression	Effet moyen de l'expérimentation estimé par <i>Matching</i>
Vocabulaire	-0.07	0.00
Lettres	0.17	0.16
Compréhension	0.10	0.15
Phonologie	0.22	0.25
Lecture voie lexicale	0.04	0.02
Lecture voie non lexicale	0.41	0.41
Mathématiques	0.23	0.31

Les effets sont rapportés à l'écart-type du groupe témoin.

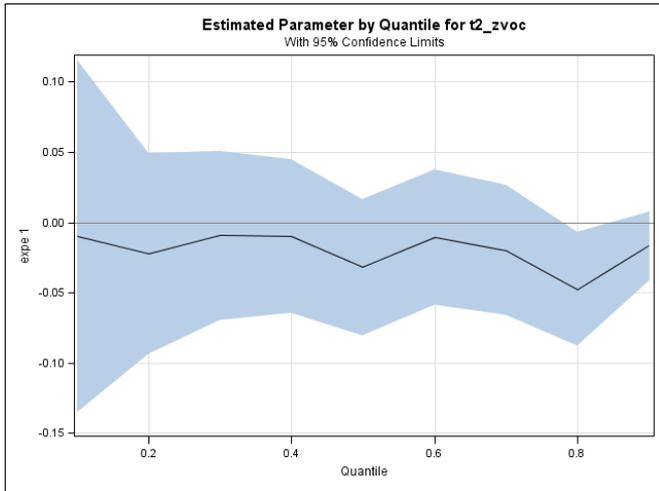
NB : seuls les chiffres en gras sont statistiquement significatifs

Effets sur les quantiles

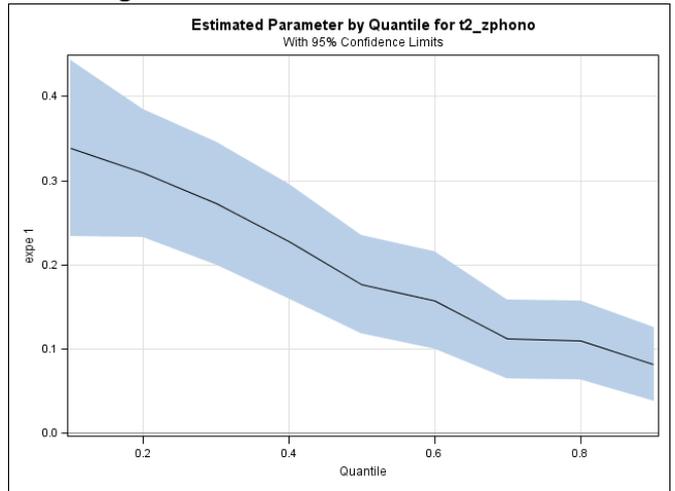
Les analyses précédentes ont porté sur l'estimation d'effets moyens : on a cherché à estimer le gain moyen – en termes de scores au deuxième temps de mesure – prodigué par le dispositif expérimental. L'effet est-il le même quel que soit le niveau des élèves ?

La régression quantile permet d'apporter des éléments de réponse à cette question. Les graphiques ci-dessous donnent les estimations des effets du dispositif sur chacun des scores observés au deuxième temps, non plus en moyenne mais selon les déciles de scores. Par exemple, s'agissant de la phonologie, on retrouve pour la médiane (quantile=0.5) un effet proche de celui observé en moyenne (+ 20 %) et le graphique montre que cet effet est plus important pour les premiers déciles que pour les derniers déciles de la distribution. Autre résultat intéressant : l'effet moyen non significatif observé sur la compréhension est à nuancer, il est significatif pour les faibles niveaux de la distribution. Le graphique concernant les lettres est absent car les scores sur cette épreuve saturent ; il est ainsi difficile de séparer les déciles.

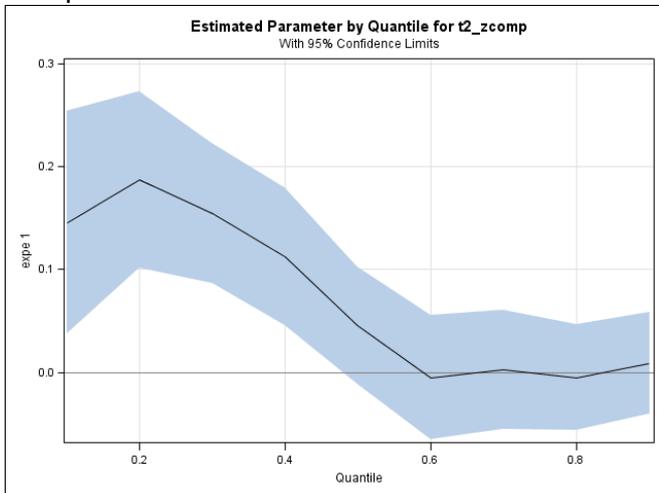
Vocabulaire



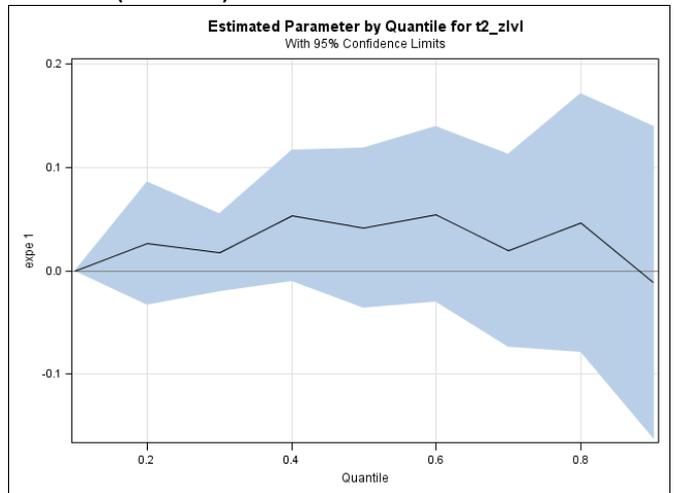
Phonologie



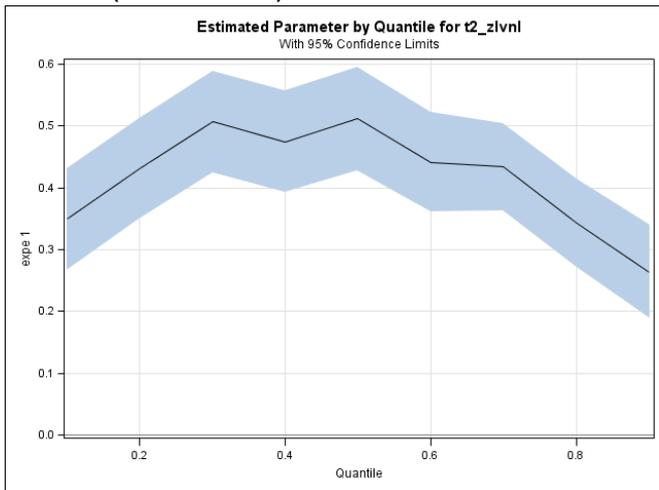
Compréhension



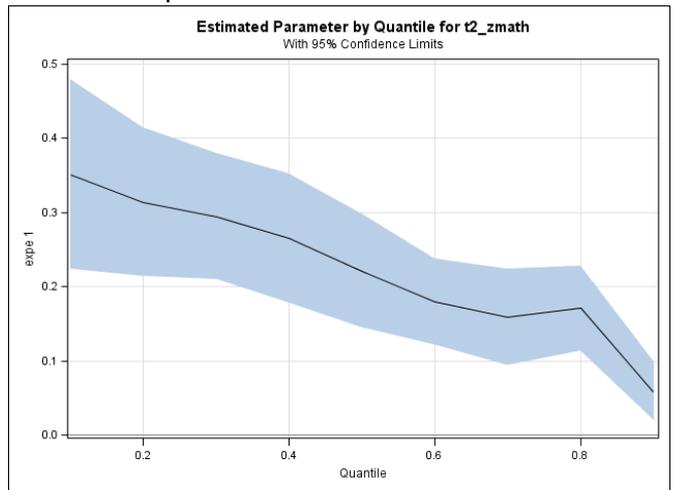
Lecture (lexicale)



Lecture (non lexicale)



Mathématiques



Note : régressions quantiles sur chacun des scores observés au deuxième temps de mesure en fonction des scores observés au premier temps (avec imputations des valeurs manquantes), du sexe et de l'âge des élèves. Les régressions sont pondérées en tenant compte du score de propension calculé avec les mêmes variables explicatives (démarche de type *matching*).

Régression selon le niveau au 1er temps de mesure

Les régressions quantiles permettent d'estimer l'effet de l'expérimentation selon le niveau observé au 2^{ème} temps de mesure. Pour compléter ces résultats, les élèves ont été répartis en quatre groupes de niveau selon les quartiles de score de phonologie observés cette fois au 1^{er} temps de mesure.

L'effet moyen de l'expérimentation est à nouveau estimé par une régression linéaire, mais indépendamment pour chacun des groupes ainsi obtenus.

Les résultats rejoignent ceux des régressions quantiles.

En compréhension, l'effet est significatif uniquement pour le groupe le plus faible. En phonologie, l'effet est significatif pour tous les groupes, mais avec un coefficient d'autant plus élevé que le niveau du groupe est faible.

	Expé			Témoïn			Ensemble		
	Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
groupe 1	581	-1.42	0.30	269	-1.48	0.34	850	-1.44	0.32
groupe 2	628	-0.58	0.22	391	-0.56	0.22	1 019	-0.57	0.22
groupe 3	325	0.09	0.14	265	0.09	0.14	590	0.09	0.14
groupe 4	519	1.02	0.41	565	1.04	0.43	1 084	1.03	0.42

	Vocabulaire T2			Lettres T2		
	Erreur			Erreur		
	Coefficient	standard	Probabilité	Coefficient	standard	Probabilité
expe groupe 1	-0.16	0.11	0.15	0.32	0.09	0.00
expe groupe 2	0.01	0.11	0.94	0.21	0.07	0.00
expe groupe 3	-0.07	0.11	0.54	0.09	0.05	0.09
expe groupe 4	-0.04	0.08	0.67	0.08	0.03	0.01

	Compréhension T2			Phonologie T2		
	Erreur			Erreur		
	Coefficient	standard	Probabilité	Coefficient	standard	Probabilité
expe groupe 1	0.26	0.09	0.01	0.35	0.10	0.00
expe groupe 2	0.10	0.07	0.13	0.23	0.10	0.02
expe groupe 3	0.12	0.09	0.22	0.26	0.10	0.01
expe groupe 4	-0.01	0.06	0.92	0.16	0.06	0.01

	Lecture voie lexicale T2			Lecture voie non lexicale T2		
	Erreur			Erreur		
	Coefficient	standard	Probabilité	Coefficient	standard	Probabilité
expe groupe 1	0.07	0.07	0.27	0.38	0.09	0.00
expe groupe 2	0.00	0.07	0.98	0.29	0.08	0.00
expe groupe 3	0.13	0.11	0.22	0.53	0.11	0.00
expe groupe 4	-0.02	0.10	0.82	0.45	0.08	0.00

	Mathématiques T2		
		Erreur	
	Coefficient	standard	Probabilité
expe groupe 1	0.36	0.10	0.00
expe groupe 2	0.21	0.08	0.01
expe groupe 3	0.24	0.09	0.01
expe groupe 4	0.18	0.09	0.04

NB : seuls les chiffres en gras sont statistiquement significatifs

Note de lecture :

Pour les 25% d'élèves ayant obtenu les moins bons scores en phonologie au temps 1, la régression du score de phonologie au temps 2, en contrôlant les scores du temps 1, l'âge et le sexe, estime une différence de +0.35 (soit 35% d'écart-type) en moyenne pour les élèves du groupe expérimental par rapport aux élèves du groupe témoins. Ce coefficient est statistiquement significatif (probabilité inférieure à 0.05). Pour les 25% d'élèves ayant obtenu les meilleurs scores, le coefficient estimé est de 0.16 (+16% d'écart-type). Il est également significatif.

b. Effets de l'expérimentation sur le public bénéficiaire en CP (T3-T4)

Statistiques descriptives des scores

On s'intéresse aux 3 534 témoins, **728** expé GS+CP et **369** expé CP uniquement (soit 4 631 élèves au total) qui ont participé aux T3 et T4.

Les scores sont standardisés de façon à être centrés-réduits pour le groupe témoin (de moyenne 0 et d'écart-type 1).

session	Scores standardisés	TEMOINS			EXPE CP (ensemble)		
		Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
T3	écriture	3 459	0.00	1.00	1 080	0.11	0.87
T3	compréhension orale	3 456	0.00	1.00	1 081	0.14	0.99
T3	vocabulaire	3 455	0.00	1.00	1 075	0.02	0.96
T3	lecture voie lexicale	3 447	0.00	1.00	1 076	0.06	1.03
T3	lecture voie non lexicale	3 445	0.00	1.00	1 076	0.23	0.91
T3	habiletés phonologiques	3 437	0.00	1.00	1 066	0.41	0.96
T3	sons des lettres	3 436	0.00	1.00	1 069	0.22	0.78
T3	mathématiques	3 445	0.00	1.00	1 065	0.03	1.04
T4	compréhension orale	3 427	0.00	1.00	1 077	-0.02	1.02
T4	vocabulaire	3 423	0.00	1.00	1 078	-0.01	0.98
T4	orthographe	3 422	0.00	1.00	1 074	0.11	0.98
T4	lecture voie lexicale	3 427	0.00	1.00	1 065	0.03	0.98
T4	lecture voie non lexicale	3 430	0.00	1.00	1 067	0.14	0.86
T4	habiletés phonologiques	3 406	0.00	1.00	1 074	0.30	0.86
T4	compréhension en lecture	3 339	0.00	1.00	1 031	0.07	1.00
T4	mathématiques	3 452	0.00	1.00	1 070	-0.01	1.05
T4	fluence nb de mot total	2 479	0.00	1.00	905	-0.08	0.89
T4	fluence nb de mot corrects	2 479	0.00	1.00	905	-0.01	0.91
T4	fluence tx de réussite	2 479	0.00	1.00	905	0.11	0.96

Au sein du groupe expérimental, on peut distinguer les élèves ayant bénéficié du dispositif en grande section (expe GS+CP) de ceux n'en ayant pas bénéficié (expé CP uniquement).

session	Scores standardisés	EXPE GS+CP			EXPE CP UNIQUEMENT		
		Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
T3	écriture	718	0.15	0.83	362	0.03	0.95
T3	compréhension orale	721	0.19	0.97	360	0.05	1.03
T3	vocabulaire	715	0.06	0.89	360	-0.06	1.09
T3	lecture voie lexicale	712	0.10	1.03	364	0.00	1.03
T3	lecture voie non lexicale	712	0.28	0.89	364	0.13	0.93
T3	habiletés phonologiques	711	0.45	0.97	355	0.32	0.95
T3	sons des lettres	710	0.25	0.74	359	0.16	0.84
T3	mathématiques	708	0.04	1.05	357	0.02	1.02
T4	compréhension orale	714	0.03	1.02	363	-0.10	1.02
T4	vocabulaire	714	-0.01	1.00	364	0.00	0.95
T4	orthographe	713	0.17	0.97	361	0.01	0.98
T4	lecture voie lexicale	706	0.04	1.01	359	0.02	0.93
T4	lecture voie non lexicale	707	0.14	0.86	360	0.13	0.87
T4	habiletés phonologiques	713	0.34	0.84	361	0.23	0.88
T4	compréhension en lecture	681	0.06	1.01	350	0.07	0.97
T4	mathématiques	711	-0.01	1.08	359	-0.02	0.97
T4	fluence nb de mot total	604	-0.09	0.91	301	-0.05	0.85
T4	fluence nb de mot corrects	604	0.00	0.94	301	-0.01	0.85
T4	fluence tx de réussite	604	0.12	0.98	301	0.11	0.91

Imputation des valeurs manquantes au temps 3

Les valeurs manquantes à l'une des séquences du temps 3 ont été imputées.

(N = 92 à 128 selon les séquences, 880 valeurs au total soit 2.4 % des valeurs)

session	Scores imputés standardisés	TEMOINS			EXPE CP		
		Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
T3 MI	écriture	3 534	0.00	1.00	1 097	0.11	0.87
T3 MI	compréhension orale	3 534	0.00	1.00	1 097	0.15	0.99
T3 MI	vocabulaire	3 534	0.00	1.00	1 097	0.02	0.96
T3 MI	lecture voie lexicale	3 534	0.00	1.00	1 097	0.06	1.03
T3 MI	lecture voie non lexicale	3 534	0.00	1.00	1 097	0.23	0.91
T3 MI	habiletés phonologiques	3 534	0.00	1.00	1 097	0.39	0.97
T3 MI	sons des lettres	3 534	0.00	1.00	1 097	0.22	0.78
T3 MI	mathématiques	3 534	0.00	1.00	1 097	0.03	1.05

session	Scores imputés standardisés	EXPE GS+CP			EXPE CP UNIQUEMENT		
		Effectif	Moyenne	Ecart-type	Effectif	Moyenne	Ecart-type
T3 MI	écriture	728	0.15	0.83	369	0.04	0.94
T3 MI	compréhension orale	728	0.20	0.96	369	0.05	1.02
T3 MI	vocabulaire	728	0.06	0.90	369	-0.06	1.08
T3 MI	lecture voie lexicale	728	0.09	1.03	369	0.01	1.04
T3 MI	lecture voie non lexicale	728	0.28	0.89	369	0.14	0.93
T3 MI	habiletés phonologiques	728	0.44	0.97	369	0.29	0.95
T3 MI	sons des lettres	728	0.24	0.75	369	0.16	0.85
T3 MI	mathématiques	728	0.04	1.06	369	0.02	1.03

Régressions

On distingue trois groupes d'élèves : expé GS+CP, expé CP uniquement, et témoins.

Les régressions (avec cluster école) permettent de mesurer l'effet du groupe sur les scores du temps 4, en contrôlant le niveau au temps 3, le sexe et l'âge (en 6 modalités : avance, retard, et trimestre de naissance pour les élèves à l'heure). Les scores au T3 des dimensions travaillées dans l'expérimentation ne sont pas pris en compte comme variables explicatives (Phonologie, lecture voie non lexicale et connaissance du son des lettres).

Compréhension orale				Vocabulaire		
	Coefficient	Erreur standard	Probabilité	Coefficient	Erreur standard	Probabilité
<u>Avant imputation</u>						
expeGS+CP	-0.09	0.05	0.10	-0.09	0.06	0.17
expeCP	-0.11**	0.06	0.05	0.00	0.06	0.96
<u>Après imputation</u>						
expeGS+CP	-0.07	0.05	0.12	-0.07	0.05	0.14
expeCP	-0.11**	0.05	0.03	0.01	0.06	0.85

	Production orthographique			Lecture voie lexicale		
	Coefficient	Erreur standard	Coefficient	Erreur standard	Coefficient	Erreur standard
<u>Avant imputation</u>						
expeGS+CP	0.10	0.08	0.10	0.08	0.10	0.08
expeCP	0.05	0.09	0.05	0.09	0.05	0.09
<u>Après imputation</u>						
expeGS+CP	0.07	0.07	0.07	0.07	0.07	0.07
expeCP	0.05	0.08	0.05	0.08	0.05	0.08

	Lecture voie non lexicale			Habilités phonologiques		
	Coefficient	Erreur standard	Coefficient	Erreur standard	Coefficient	Erreur standard
<u>Avant imputation</u>						
expeGS+CP	0.09*	0.05	0.09*	0.05	0.09*	0.05
expeCP	0.16**	0.06	0.16**	0.06	0.16**	0.06
<u>Après imputation</u>						
expeGS+CP	0.07	0.04	0.07	0.04	0.07	0.04
expeCP	0.15**	0.06	0.15**	0.06	0.15**	0.06

	Compréhension en lecture			Mathématiques		
	Coefficient	Erreur standard	Coefficient	Erreur standard	Coefficient	Erreur standard
<u>Avant imputation</u>						
expeGS+CP	-0.02	0.08	-0.02	0.08	-0.02	0.08
expeCP	0.08	0.08	0.08	0.08	0.08	0.08
<u>Après imputation</u>						
expeGS+CP	-0.04	0.07	-0.04	0.07	-0.04	0.07
expeCP	0.09	0.08	0.09	0.08	0.09	0.08

** Significatif au seuil de 5 %, * au seuil de 10 %

	Fluence (nb mots corrects)		
<u>Avant imputation</u>	Coefficient	Coefficient	Coefficient
expeGS+CP	-0.05	-0.05	-0.05
expeCP	0.00	0.00	0.00
<u>Après imputation</u>			
expeGS+CP	-0.08	-0.08	-0.08
expeCP	-0.02	-0.02	-0.02

**** Significatif au seuil de 5 %, * au seuil de 10 %**

Note de lecture :

La régression du score en habiletés phonologiques au temps 4 en fonction du groupe (expé/témoin), du niveau au temps 3 (hors dimensions travaillées dans l'expérimentation), de l'âge et du sexe, estime une différence de +0.30 en moyenne pour les élèves du groupe expé GS+CP, et de +0.25 pour les élèves du groupe expé CP uniquement, par rapport aux élèves du groupe témoin, à niveau au temps 3, sexe et âge équivalent. Ces coefficients sont statistiquement significatifs (probabilité inférieure à 0.05).

Matching

Le matching est réalisé sur le score de propension calculé à partir des scores du T3 imputés, ainsi que le sexe et l'âge pour l'ensemble des élèves n'ayant aucune valeur manquante aux scores du T4. Pour le calcul du score de propension, les scores des dimensions travaillées dans le dispositif expérimental ne sont pas pris en compte (habiletés phonologiques, lecture voie non lexicale et connaissance du son des lettres).

Un premier matching est réalisé entre les élèves expé GS+CP et les élèves témoins. Puis un deuxième matching est réalisé entre les élèves expé CP uniquement et les élèves témoins.

Expé GS+CP

3534 Témoins + 728 expé GS+CP

Temps 4	ATE	Erreur standard	T	Probabilité	Ecart-type groupe témoin	Effet du traitement standardisé
Compréhension orale	-0.14	0.05	-2.65	0.01	0.96	-0.14
Vocabulaire	-0.08	0.05	-1.60	0.11	0.94	-0.09
Production orthographique	0.09	0.05	1.76	0.08	0.96	0.09
Lecture voie lexicale	-0.01	0.05	-0.17	0.87	0.94	-0.01
Lecture voie non lexicale	0.04	0.04	0.95	0.34	0.87	0.05
Habiletés phonologiques	0.34	0.05	6.69	0.00	0.98	0.34
Compréhension en lecture	0.02	0.05	0.38	0.70	0.97	0.02
Mathématiques	-0.10	0.05	-2.16	0.03	0.87	-0.12
Fluence - nb de mots corrects	-0.11	0.05	-2.04	0.04	1.00	-0.11

ATE (average treatment effect) : Effet moyen du traitement

Effet du traitement standardisé : Les effets sont rapportés à l'écart-type du groupe témoin (il s'agit des scores standardisés, mais l'écart-type du groupe témoin est légèrement différent de 1 car on a supprimé les individus avec valeurs manquantes).

Expé CP uniquement

3534 Témoins + 369 expé CP

	ATE	Erreur standard	T	Probabilité	Ecart-type groupe témoin	Effet du traitement standardisé
Compréhension orale	-0.06	0.07	-0.91	0.36	0.96	-0.07
Vocabulaire	0.02	0.07	0.35	0.73	0.94	0.03
Production orthographique	0.04	0.07	0.60	0.55	0.96	0.04
Lecture voie lexicale	0.04	0.07	0.53	0.60	0.94	0.04
Lecture voie non lexicale	0.15	0.06	2.39	0.02	0.87	0.17
Habilités phonologiques	0.26	0.07	3.79	0.00	0.98	0.27
Compréhension en lecture	0.13	0.07	1.87	0.06	0.97	0.13
Mathématiques	-0.02	0.06	-0.24	0.81	0.87	-0.02
Fluence - nb de mots corrects	-0.01	0.07	-0.13	0.89	1.00	-0.01

ATE (average treatment effect) : Effet moyen du traitement

Effet du traitement standardisé : Les effets sont rapportés à l'écart-type du groupe témoin (il s'agit des scores standardisés, mais l'écart-type du groupe témoin est légèrement différent de 1 car on a supprimé les individus avec valeurs manquantes).

NB : seuls les chiffres en gras sont statistiquement significatifs

Note de lecture :

Le matching des élèves du groupe expé GS+CP avec les élèves du groupe témoin estime un effet moyen de l'expérimentation (ATE) de 0.34 en habiletés phonologiques, soit un effet de +34% d'écart-type. Il est statistiquement significatif (probabilité inférieure à 0.05).

Le matching des élèves du groupe expé CP uniquement avec les élèves du groupe témoin estime un effet de +27% d'écart-type, également significatif.

Synthèse T3T4

	Expé GS+CP		Expé CP uniquement	
	Effet moyen de l'expérimentation estimé par régression	Effet moyen de l'expérimentation estimé par <i>Matching</i>	Effet moyen de l'expérimentation estimé par régression	Effet moyen de l'expérimentation estimé par <i>Matching</i>
Temps 4				
Compréhension orale	-0.07	-0.14	-0.11	-0.07
Vocabulaire	-0.07	-0.09	0.01	0.03
Production orthographique	0.07	0.09	0.05	0.04
Lecture voie lexicale	-0.04	-0.01	0.02	0.04
Lecture voie non lexicale	0.07	0.05	0.15	0.17
Habilités phonologiques	0.29	0.34	0.25	0.27
Compréhension en lecture	-0.04	0.02	0.09	0.13
Mathématiques	-0.10	-0.12	0.00	-0.02
Fluence - nb de mots corrects	-0.08	-0.11	-0.02	-0.01

NB : seuls les chiffres en gras sont statistiquement significatifs

La régression contrôle l'effet de l'âge, du sexe et des scores aux temps 3 avec une estimation tenant compte du niveau école. Le *matching* repose sur les seuls scores aux temps 3 et l'estimation de l'effet n'intègre pas le niveau école. En toute rigueur, il resterait à développer un *matching* tenant compte du niveau école.

Les résultats des deux premières analyses (régression et *matching*) sont proches et montrent un effet positif et significatif dans la dimension habiletés phonologiques, aussi bien pour les élèves ayant expérimenté en GS et en CP que pour ceux ayant expérimenté en CP uniquement. En lecture voie non lexicale, on observe également un effet positif et significatif mais uniquement pour les élèves n'ayant expérimenté qu'en CP. Ce résultat ne s'observe pas pour les élèves ayant expérimenté en GS et en CP, mais cette absence d'effet significatif en moyenne est à compléter par une analyse selon le niveau des élèves puisque le dispositif cible les plus faibles.

Régressions quantiles

On estime les effets du dispositif sur chacun des scores observés au T4 non plus en moyenne mais selon les déciles de scores, en fonction des scores observés au T3 (avec imputations des valeurs manquantes), du sexe et de l'âge des élèves. Les scores au T3 des dimensions travaillées dans l'expérimentation ne sont pas pris en compte comme variables explicatives (Phonologie, lecture voie non lexicale et connaissance du son des lettres).

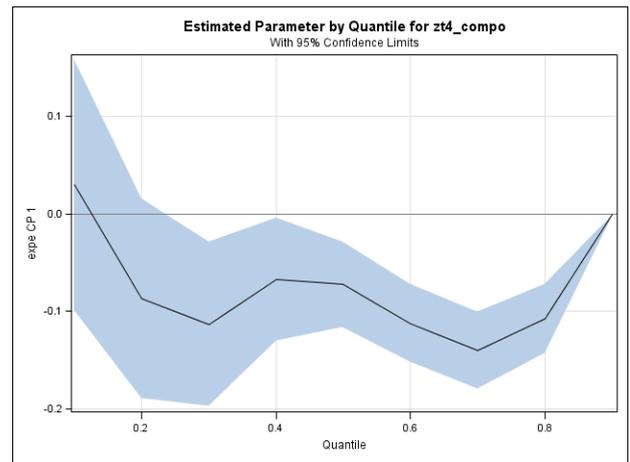
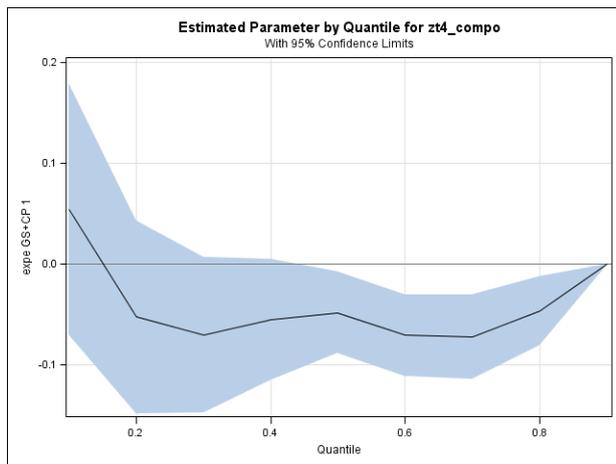
Des régressions séparées sont effectuées pour les élèves ayant suivi le dispositif en GS et CP d'une part, et pour les élèves ayant suivi le dispositif uniquement en CP d'autre part (comparé au groupe témoin à chaque fois).

Les régressions sont pondérées en tenant compte du score de propension calculé avec les mêmes variables explicatives.

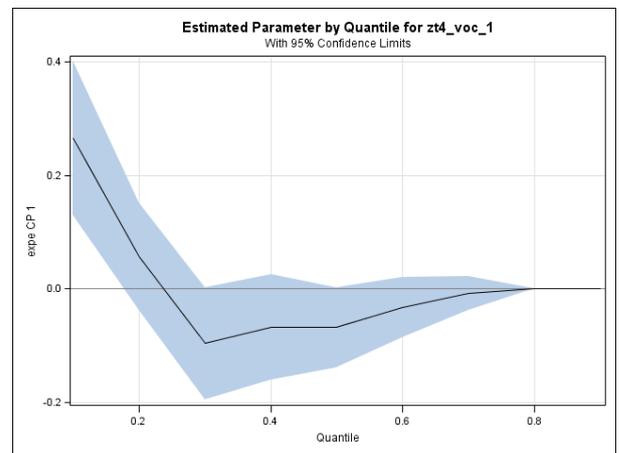
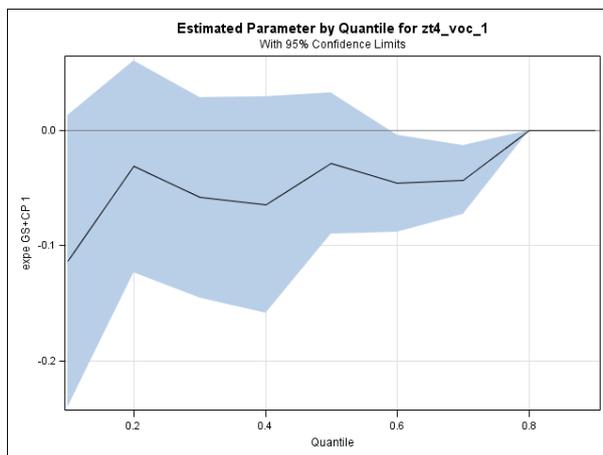
En phonologie, les effets du dispositif (proche de + 30 % en moyenne) sont particulièrement importants pour les élèves les plus faibles : + 40 % pour les 10 % d'élèves les plus faibles ayant expérimenté en CP uniquement, et + 45 % pour les 10 % d'élèves les plus faibles ayant expérimenté en GS et en CP. En revanche l'effet est très faible voire inexistant pour les élèves les plus performants, qui ont donc progressé de façon similaire aux élèves du groupe témoin.

En lecture voie non lexicale, on observe également un effet plus fort pour les élèves les plus faibles, proche de 30 % pour les 10 % d'élèves les plus faibles ayant expérimenté en CP uniquement et de 25 % pour les 10 % d'élèves les plus faibles ayant expérimenté en GS et en CP. On note que pour les élèves ayant expérimenté en GS et en CP, l'effet n'est significatif que pour les 20 % d'élèves les plus faibles, ce qui peut expliquer qu'il ne sortait pas significatif en moyenne pour l'ensemble du groupe. En revanche, pour les élèves ayant expérimenté en CP uniquement, l'effet est moins fort mais toujours significatif pour les élèves « moyens » du 3^e au 6^e décile.

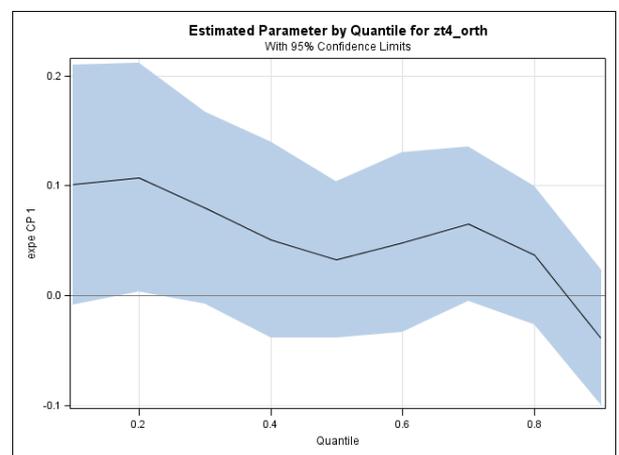
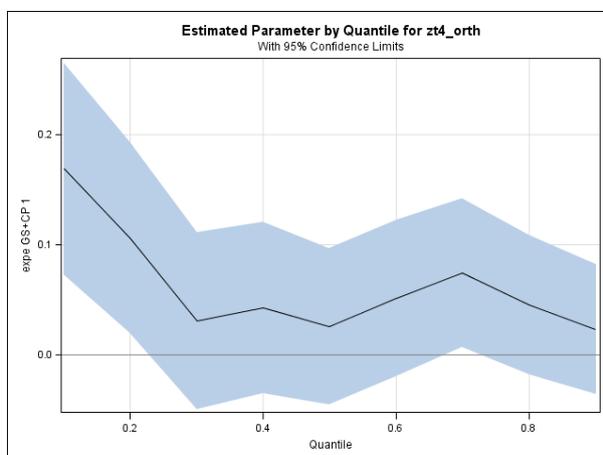
Compréhension orale



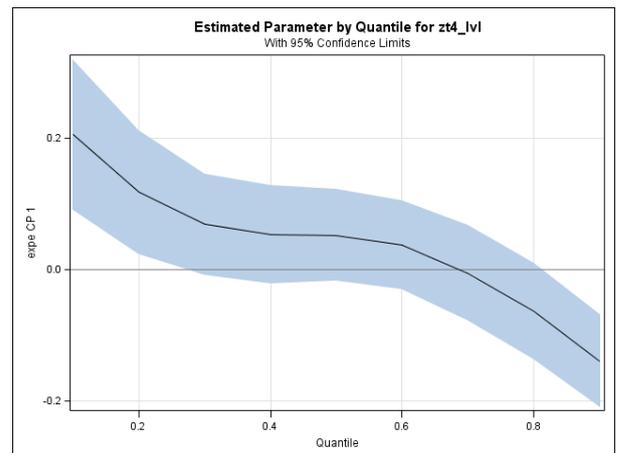
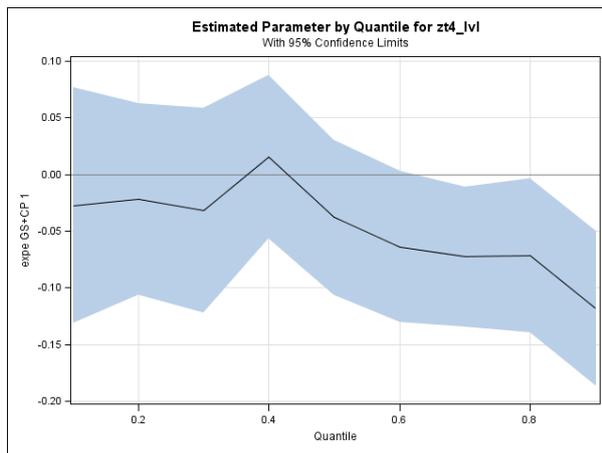
Vocabulaire



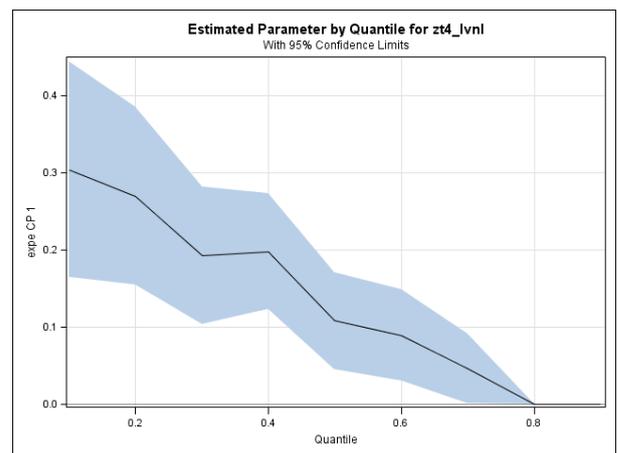
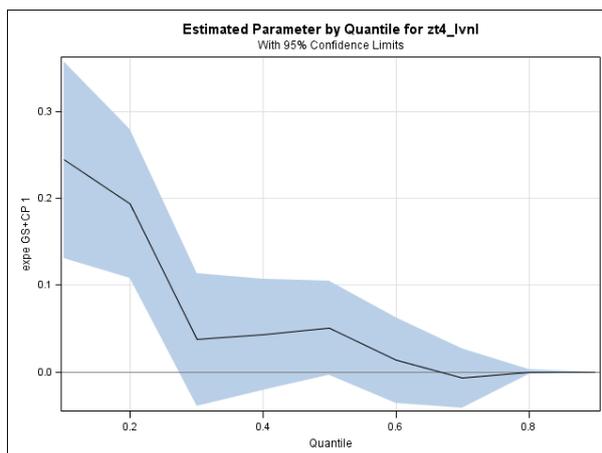
Production orthographique



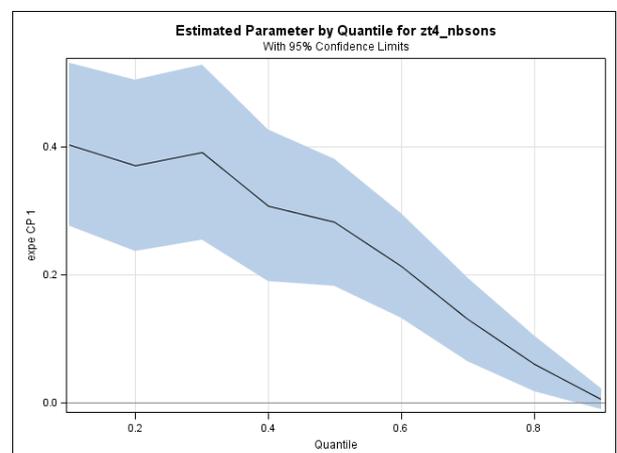
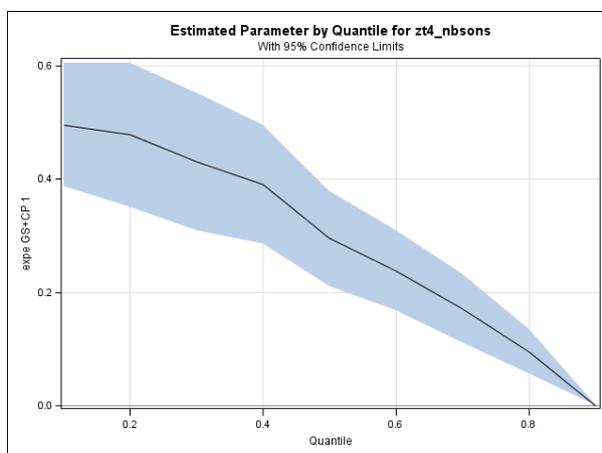
Lecture voie lexicale



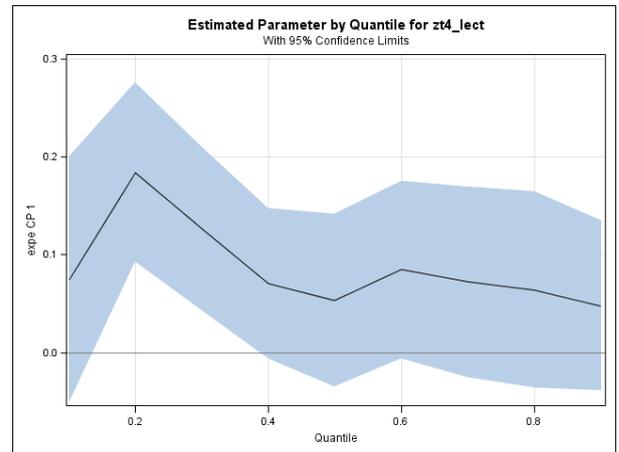
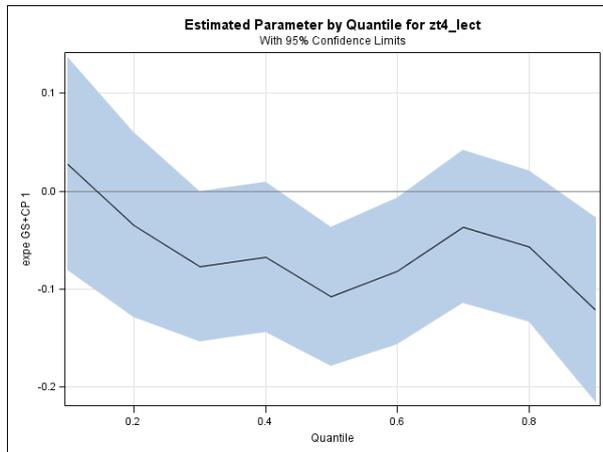
Lecture voie non lexicale



Habilités phonologiques

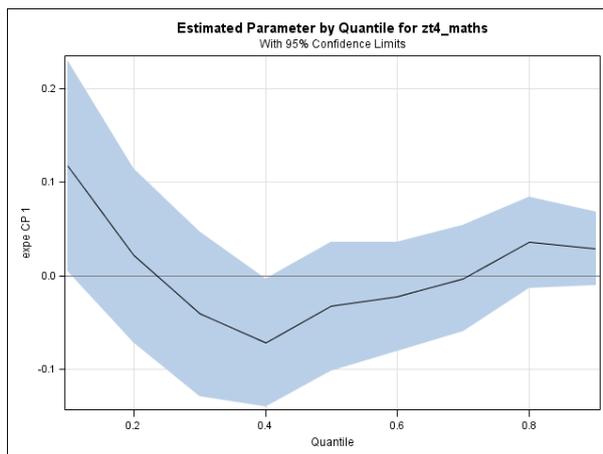


Compréhension en lecture

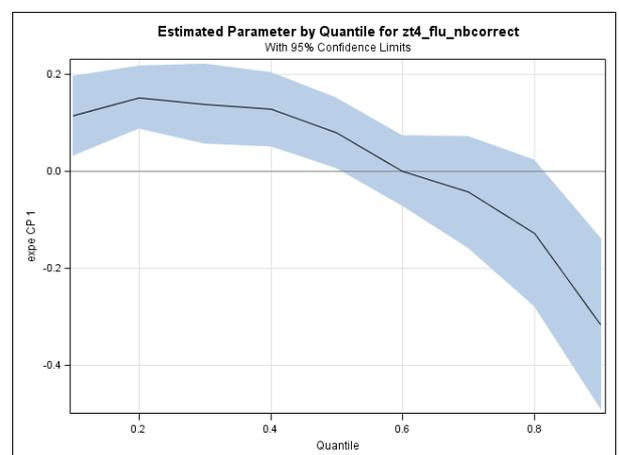
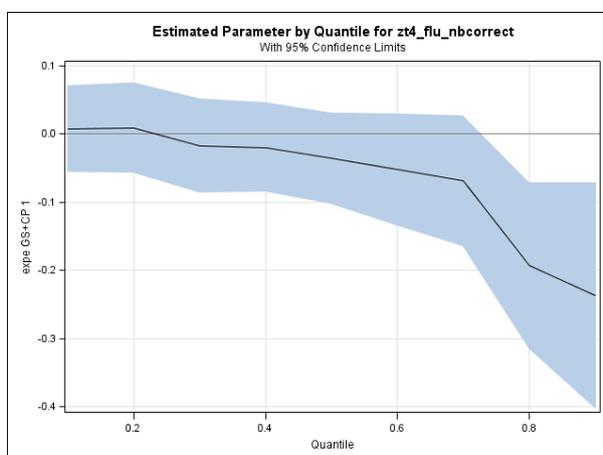


Mathématiques

Pour le groupe expé GS+CP, la régression n'a pas pu être estimée à cause de problème de convergence.



Fluence – nombre de mots corrects



c. Effet global de l'expérimentation sur le public bénéficiaire

Pour la grande section, l'analyse des effets fait ressortir les résultats suivants :

- Les élèves des classes expérimentales progressent davantage sur les dimensions reconnaissance de lettres, phonologie et lecture déchiffrement, qui font l'objet d'un travail spécifique dans le dispositif expérimental.
- En phonologie, ces effets sont plus importants pour les élèves les plus faibles.
- En compréhension, l'effet est significatif pour les élèves les plus faibles

Plus précisément,

- Les effets du dispositif expérimental sont positifs et statistiquement significatifs sur les dimensions : reconnaissance des lettres, habiletés phonologiques et lecture par voie non lexicale. Les élèves des classes expérimentales progressent davantage sur ces dimensions, qui font l'objet d'un travail spécifique dans le dispositif expérimental. De façon plus détaillée, les valeurs des effets – habituellement rapportées à l'écart-type des scores de l'échantillon témoin – sont les suivantes : + 17 % en reconnaissance des lettres, + 22 % en habiletés phonologiques et + 41 % en lecture par voie non lexicale.
- En phonologie, ces effets sont plus importants pour les élèves les plus faibles. Afin d'illustrer ces résultats, si l'on considère les élèves en difficulté en termes d'habiletés phonologiques en fin de GS comme étant les 15 % les moins performants au test, le dispositif permet de réduire ce taux d'échec à 10 %.
- En compréhension, l'effet est significatif pour les élèves les plus faibles mais non significatif pour les autres. En revanche, pour les dimensions de vocabulaire et de lecture par la voie lexicale, les effets ne sont pas significatifs.
- Enfin, les effets sont assez variables selon les circonscriptions. Ces variabilités sont à mettre en lien avec le degré d'implication des écoles dans le dispositif expérimental.

Concernant l'année de CP, en début d'année, les élèves ayant suivi le protocole expérimental en GS ont des résultats supérieurs à ceux du groupe témoin, principalement dans les dimensions travaillées dans l'expérimentation - habiletés phonologiques, lecture voie non lexicale, connaissance des sons des lettres – et, dans une moindre mesure, en écriture et en compréhension orale. Les élèves du groupe expérimental n'ayant pas expérimenté en GS ont également des résultats supérieurs à ceux du groupe témoin dans les dimensions travaillées dans l'expérimentation, ce qui peut s'expliquer par le fait que l'évaluation a eu lieu tardivement, au mois de novembre.

L'analyse des effets fait ressortir les résultats suivants :

- Les élèves ayant participé à l'expérimentation en GS et en CP progressent davantage en phonologie (+ 29 %).
- Les élèves ayant participé à l'expérimentation en CP uniquement progressent davantage en phonologie également (+ 25 %) et en lecture voie non lexicale (+ 15 %).
- En revanche, pour les deux groupes, les élèves auraient moins progressé, de façon statistiquement significative, en compréhension orale (- 14 %), et en mathématiques (- 12 %) notamment.
- Les différences sur les autres dimensions ne sont pas significatives.

Une analyse complémentaire, (selon les déciles et non plus en moyenne), précise des effets différents selon le niveau du score observé au T4 et notamment :

- En phonologie et en lecture voie non lexicale, les effets sont plus importants pour les élèves les plus faibles.
- En compréhension orale, les effets négatifs concernent plus particulièrement les élèves de niveau moyen.

Enfin, une comparaison des scores moyens des élèves des groupes témoins du temps 2 et du temps 3, calculés sur les items communs à ces deux temps, montre que les élèves du groupe témoin du T3 (début de CP) ont des scores supérieurs aux élèves du groupe témoin du T2 (fin de GS), de + 10 % en vocabulaire, à + 84 % en phonologie, ce qui pourrait s'expliquer par le fait que, l'évaluation ayant eu lieu au mois de novembre, toutes ces dimensions ont donc été travaillées par les élèves de CP et cette évolution correspondrait à une progression ordinaire entre fin de GS et début de CP.

De fait, selon la double hypothèse crédible (mais qui reste à préciser par des analyses complémentaires en cours de réalisation) que les deux échantillons témoins (GS et CP) sont comparables et qu'on n'observe pas d'évolution différentielle de la dispersion, l'effet global du dispositif sur deux années serait caractérisé par :

- une progression significative en phonologie (+ 51 %), en lecture voie non lexicale (+ 48 %) avec des effets importants pour les élèves les plus faibles ;
- une absence d'effet sur les dimensions compréhension et lecture voie lexicale.

Les résultats des deux analyses procédant par régression et par matching sont proches.

Régression	Effet T1-T2	effet T3-T4 (expé GS+CP)	EFFET GLOBAL
Vocabulaire	-0.07	-0.07	-0.14
Compréhension	0.10	-0.07	0.03
Phonologie	0.22	0.29	0.51
Lecture voie lexicale	0.04	-0.04	0.00
Lecture voie non lexicale	0.41	0.07	0.48
Mathématiques	0.23	-0.10	0.13

Matching	Effet T1-T2	effet T3-T4 (expé GS+CP)	EFFET GLOBAL
Vocabulaire	0.00	-0.09	-0.09
Compréhension	0.15	-0.14	0.01
Phonologie	0.25	0.34	0.59
Lecture voie lexicale	0.02	-0.01	0.01
Lecture voie non lexicale	0.41	0.05	0.46
Mathématiques	0.31	-0.12	0.19

NB : seuls les chiffres en gras sont statistiquement significatifs

La significativité des effets observés répond indirectement à la question de l'effet minimum détectable.

d. Effets attendus et effets induits/inattendus

Concernant ces effets, trois points méritent précisions :

- 1) Les effets sont différents selon les dimensions cognitives, allant du non significatif au très significatif, ce qui montre que l'effet de l'expérimentation ne peut être imputable au seul effet maître lié au volontariat (i.e. expérimentation non randomisée).
- 2) Les effets sont importants sur les dimensions travaillées (phonologie, reconnaissance non lexicale), ce qui était attendu. De même, attendus également, les effets sont plus importants pour les élèves les plus faibles, en particulier en phonologie.
- 3) L'effet très positif observé en mathématiques en GS est inattendu et mériterait de plus amples investigations. Il est ainsi possible que le travail réalisé en phonologie ait pu bénéficier au développement de compétences calculatoires.

Il reste à savoir si ces effets vont concrètement se transférer à des compétences de plus haut niveau évaluées en fin de CE1 (évaluation en cours de réalisation).

e. Gouvernance et coordinations d'acteurs

La mise en place du projet dans les écoles a exigé un travail d'équipe entre les enseignants pour l'organisation et la préparation des séances (appropriation du matériel, préparation des séances, création de ressources complémentaires aux supports fournis notamment), associant souvent la direction de l'école, et parfois les écoles voisines entre elles. Une dynamique de circonscription (rassemblant donc plusieurs écoles) a même été observée sur certains terrains (dans le département du Pas-de-Calais notamment).

Le souci du porteur de projet d'associer les collectivités à la démarche des écoles a également permis, dans certaines circonscriptions, une coordination inédite entre les services communaux et les équipes des écoles, notamment en ce qui concerne le matériel pédagogique.

Une coordination spécifique a également résulté de la mise en place du projet entre les équipes de différentes circonscriptions engagées, alors qu'elles sont habituellement assez cloisonnées. Cela a permis l'organisation de formation et de la mise à disposition de ressources pédagogiques au niveau départemental dans les Hauts-de-Seine par exemple.

Globalement, cette coordination s'est effectuée spontanément, sans qu'un effort particulier ne soit nécessaire.

Les acteurs concernés sont représentatifs de la réalité nationale avec une surreprésentation des quartiers de l'éducation prioritaire tout de même.

f. Conclusion de la partie et mise en perspective

Sur la grande section (Résultats T1T2)

Cette première partie présente les résultats de l'expérimentation « Lecture » qui s'est tenue dans les départements du Nord, du Pas-de-Calais, des Hauts-de-Seine et de la Seine-Saint-Denis lors de l'année scolaire 2011-2012, auprès d'élèves de Grande Section de maternelle.

Sur environ 6 000 élèves visés, dont environ 3 000 issus de classes expérimentales et environ 3 000 de classes témoin, 3 569 élèves ont participé aux deux temps de mesure (début et fin d'année scolaire). La non-réponse a plus particulièrement affecté l'échantillon témoin. Cependant, l'attrition n'apparaît pas différentielle selon les deux groupes, expérimental et témoin.

De manière descriptive, en début d'année, les élèves du groupe expérimental ont des résultats significativement inférieurs à ceux des élèves du groupe témoin. Ce constat peut s'expliquer par la non-réponse qui affecte différemment les deux groupes et qui concerne en général plus souvent des classes/élèves de niveaux de performances moins élevés. Toujours de manière descriptive, il ressort que lors du deuxième temps de mesure, les élèves des classes expérimentales rattrapent leur retard sur plusieurs dimensions, dont la phonologie.

Deux modélisations ont été engagées pour estimer les effets moyens du dispositif : la première procède par régression linéaire, la seconde par *matching*. Le modèle de régression porte sur chacun des scores observés au deuxième temps de mesure, en fonction des scores observés au premier temps ainsi que du sexe et de l'âge des élèves. En rapportant leurs scores à l'écart-type des scores de l'échantillon témoin, les élèves des classes expérimentales progressent davantage sur les dimensions Lettres (+ 17 %), Phonologie (+ 22 %), Lecture voie non lexicale (+ 41 %) et Mathématiques (+ 23 %). Pour les dimensions Vocabulaire et Lecture voie lexicale, les effets ne sont pas significatifs. En compréhension, l'effet est légèrement significatif. Étant donné les différences de performance observées entre les deux groupes lors du premier temps de mesure, une approche procédant par *matching* (appariement) est théoriquement plus adaptée. La méthode employée ici et consistant à apparier sur les scores de la première session d'évaluation donne des résultats très proches de ceux de la régression : Lettres (+ 16 %), Phonologie (+ 25 %), Lecture voie non lexicale (+ 41 %) et Mathématiques (+ 31 %). En outre, la Compréhension ressort de manière significative dans ce cas (+ 15 %)⁴.

Une analyse complémentaire, selon les déciles et non plus en moyenne, montre des effets différents selon le niveau du score observé au deuxième temps de mesure. En phonologie, ces effets sont plus importants pour les élèves les plus faibles et en compréhension, il ressort que les effets du dispositif sont significatifs pour les élèves les plus faibles et non significatifs pour les autres.

⁴ En complément, et pour conforter ces premiers résultats, un développement méthodologique est en cours : une approche de type *matching* qui intègre le niveau classe (sinon école) ainsi que les valeurs manquantes observées au premier temps sur certaines dimensions (comme cela est déjà fait ici, par imputation, dans le cadre des modèles de régression).

Enfin, il ressort que les effets sont assez différents selon les circonscriptions. Ces variabilités sont à mettre en lien avec le degré d'implication des écoles dans le dispositif expérimental.

Sur le CP (Résultats T3T4)

Plus de 4 000 élèves ont participé aux deux temps de mesure en début et fin de CP. Au sein des 28 écoles concernées par l'expérimentation, 1 113 élèves ont participé aux évaluations des temps 3 (début de CP) et 4 (fin de CP). Parmi eux, 728 élèves ont suivi le protocole expérimental en GS et en CP, et 369 élèves l'ont suivi en CP mais pas en GS. Ont également été évalués 16 élèves qui auraient dû suivre le protocole expérimental en CP mais dont la classe n'a finalement pas participé, dont 3 élèves qui l'avaient suivi en GS et 13 qui ne l'ont finalement jamais suivi. Parmi les 91 écoles tirées pour l'échantillon témoin, 85 avaient participé au T3, et 78 ont finalement participé aux T3 et T4, soit 3 534 élèves.

En début d'année, les 728 élèves ayant suivi le protocole expérimental en GS ont des résultats supérieurs à ceux du groupe témoin principalement dans les dimensions travaillées dans l'expérimentation - habiletés phonologiques, lecture voie non lexicale, connaissance des sons des lettres - et dans une moindre mesure en écriture et en compréhension orale. Les 369 élèves du groupe expérimental n'ayant pas expérimenté en GS ont également des résultats supérieurs à ceux du groupe témoin dans les dimensions travaillées dans l'expérimentation, ce qui peut s'expliquer par le fait que l'évaluation a eu lieu tardivement, au mois de novembre.

Deux modélisations ont été engagées pour estimer les effets moyens du dispositif : la première procède par régression linéaire, la seconde par *matching*. Le modèle de régression porte sur chacun des scores observés au T4, en fonction des scores observés au T3 sur les dimensions non travaillées spécifiquement dans le protocole expérimental, ainsi que du sexe et de l'âge des élèves. Rapportés à l'écart-type des scores de l'échantillon témoin, les élèves ayant participé à l'expérimentation en GS et en CP progressent davantage en phonologie (+ 29 %). Les élèves ayant participé à l'expérimentation en CP uniquement progressent davantage en phonologie également (+ 25 %) et en lecture voie non lexicale (+ 15 %). En revanche ils auraient moins progressé en compréhension orale (- 11 %). Les différences sur les autres dimensions ne sont pas significatives. Pour compléter ces résultats une approche procédant par *matching* est utilisée, en appariant les élèves selon les scores du T3 des dimensions non travaillées spécifiquement dans le protocole expérimental, ainsi que le sexe et l'âge. Selon les résultats obtenus, les élèves ayant participé à l'expérimentation en GS et en CP progressent davantage en phonologie (+ 34 %), mais auraient moins progressé, de façon statistiquement significative, en compréhension orale (- 14 %), et en mathématiques (- 12%) notamment. Pour les élèves ayant participé à l'expérimentation en CP uniquement les résultats confirment qu'ils progressent davantage en phonologie (+ 27 %) et en lecture voie non lexicale (+ 17 %).

Une analyse complémentaire, selon les déciles et non plus en moyenne, montre des effets différents selon le niveau du score observé au T4. En phonologie et en lecture voie non lexicale, les effets sont plus importants pour les élèves les plus faibles. En compréhension orale, les effets négatifs concernent plus particulièrement les élèves de niveau moyen.

Les résultats moyens par circonscriptions montrent des effets assez variables, liés vraisemblablement au degré d'implication des écoles dans le dispositif expérimental.

Enfin, une comparaison des scores moyens des élèves des groupes témoins du temps 2 et du temps 3, calculés sur les items communs à ces deux temps, montre que les élèves du groupe témoin du T3 (début de CP) ont des scores supérieurs aux élèves du groupe témoin du T2 (fin de GS), de + 10 % en vocabulaire, à + 84 % en phonologie, ce qui pourrait s'expliquer par le fait que, l'évaluation ayant eu lieu au mois de novembre, et que toutes ces dimensions ont donc été travaillées par les élèves.

Sur les deux années d'expérimentation (Résultats T1T4)

Si la littérature scientifique américaine abonde de recherches concordantes sur les effets positifs d'une intervention précoce en phonologie, il n'existe pas en France d'études empiriques d'envergure sur le sujet. De ce point de vue, l'expérimentation fournit des éléments utiles en termes de recherche et d'action pédagogique.

Il y a en effet un enjeu à observer si les résultats positifs observés en GS sont durables, pour ne pas dire cumulatifs, dans l'apprentissage de la lecture chez ces élèves. Il s'agira d'analyser l'effet de l'expérimentation sur les trois années du cycle des apprentissages fondamentaux. De ce point de vue l'évaluation menée en CE1, nous dira s'il y a transfert sur des compétences de plus haut niveau. Une attention particulière sera portée au contrôle d'absence d'impact négatif sur les autres disciplines (en particulier les mathématiques). En tout état de cause, il s'agit de chercher à tirer les enseignements les plus solides de cette expérimentation avant d'en inférer quelque implication et recommandation que ce soit en termes d'inflexion de politique éducative.

2.2. Validité externe des résultats et perspectives de généralisation/essaimage : caractère transférable du dispositif et changement d'échelle.

2.2.1 Représentativité du terrain, des acteurs et du public bénéficiaire

A la rentrée 2011, le projet « Lecture » a été déployé dans environ 200 classes de grande section de maternelle de 6 académies : Lille, Créteil, Versailles, Paris, Orléans-Tours et Clermont-Ferrand. Trois entités ont participé à l'évaluation externe du programme qui portait sur trois académies : Lille, Créteil et Versailles. La Direction de l'évaluation, de la performance et de la prospective du Ministère de l'éducation nationale (DEPP), l'Institut de recherche sur l'éducation (IREDU) et le laboratoire d'étude des mécanismes cognitifs (EMC) de l'Université Lyon 2 ont élaboré un protocole d'évaluation mesurant l'impact du projet sur les compétences des élèves.

Le laboratoire d'étude des mécanismes cognitifs de l'Université Lyon 2 et la DEPP ont construit les tests d'évaluation du niveau des élèves entrant en grande section de maternelle. L'évaluation élaborée mesure quatre compétences prédictives du niveau de

lecture : la connaissance des lettres, le vocabulaire, la conscience phonologique et la compréhension orale en début et en fin de grande section, ainsi que la lecture phonologique, la lecture lexicale et les mathématiques en fin de grande section.

La méthode d'évaluation qui a été retenue consiste à comparer les résultats des élèves qui bénéficient du programme, le « groupe expérimental », à ceux d'élèves qui n'en bénéficient pas, le « groupe témoin ». Les élèves du « groupe témoin » ont été choisis pour avoir, en moyenne, les mêmes caractéristiques que ceux du « groupe expérimental » (catégorie socioprofessionnelle des parents, taille des écoles et proportion d'écoles situées en zones d'éducation prioritaire). L'ensemble des élèves a été évalué en novembre 2011, pour disposer d'un point de départ permettant de comparer les progressions de chaque groupe. 3 673 élèves ont participé aux deux sessions d'évaluation, groupe test et groupe témoin confondus.

Des résultats positifs, variables selon le niveau initial des élèves et selon le degré de mise en œuvre du dispositif en classe.

Les analyses sur les données disponibles montrent un résultat positif sur les compétences travaillées pendant l'année. Ainsi, en une année, suivant les compétences, les élèves faibles du groupe expérimental rattrapent en moyenne 20 % à 40 % de leur retard sur un élève moyen du groupe témoin. La proportion d'élèves en grande difficulté en phonologie est passée de 15 % à 10 %. Les élèves ayant bénéficié du dispositif ont également beaucoup progressé en lecture (voie non lexicale) par rapport au groupe témoin.

Les élèves les plus faibles progressent plus que les autres en phonologie. En compréhension, les effets ne sont positifs que pour les élèves faibles.

La mise en œuvre du dispositif a été variable selon les classes. Là où le protocole a été mis en œuvre le plus rigoureusement, en termes de volume et d'intensité des séances notamment, le dispositif a eu pour impact de réduire la proportion d'élèves en difficulté de 15 % à 7 %. L'effet est de 80 % d'écart-type là où les pratiques enseignantes observées étaient conformes au protocole et de 30 % là où l'application était moyennement conforme.

2.2.2 Contexte expérimental

Quelques éléments permettent d'éclairer le contexte. L'évaluateur n'est intervenu qu'à deux reprises, au début et à la fin de l'année scolaire. L'implication des équipes pédagogiques (expérimentales comme témoins) a été réduite au minimum pour la passation des tests. Les corrections des réponses des élèves ont été centralisées, garantissant un niveau maximum de standardisation. Les résultats n'ont été communiqués aux écoles qu'après l'expérimentation, limitant l'influence de l'évaluation sur les résultats de l'expérimentation.

Le fait que les effets observés dépendent des dimensions cognitives évaluées est le signe de l'absence d'un artefact global lié à l'implication des acteurs, et manifeste que ces effets sont bien le résultat d'une action spécifique.

2.2.3 Transférabilité du dispositif

Dès la maternelle, les enseignants sont exposés aux premières difficultés d'apprentissage des enfants, essentiellement sur des compétences liées au langage. La maîtrise différente de ces compétences trahit très tôt d'énormes disparités selon les enfants.

Par la mise à disposition de ressources pédagogiques adaptées aux besoins des plus faibles, le projet Lecture est conçu pour aider les enseignants à garantir l'accès aux apprentissages à un maximum d'élèves. La démarche d'accompagnement des enseignants repose sur deux axes : l'aide à l'évaluation et l'assistance dans la mise en œuvre de séances qui correspondent aux besoins différenciés ainsi identifiés.

a. La première étape : l'évaluation.

Dès la grande section de maternelle, on peut identifier la population la plus fragile en termes d'apprentissage futur de la lecture. Le niveau de langage des enfants à cinq ans impacte leur niveau de lecture et leur destin scolaire. Le niveau à l'entrée au CP influe d'autant plus fortement sur la carrière scolaire des élèves qu'ils ont des acquis initiaux plus faibles. L'évaluation permet à l'enseignant de concentrer son attention sur les progrès de certains élèves en particulier, capables comme les autres d'entrer dans la lecture, mais déjà lésés par un niveau de langage trop insuffisant.

b. La deuxième étape : l'entraînement ciblé des compétences orales préalables à la lecture

La démarche de l'enseignant est simplifiée : ce sont les compétences les plus prédictives de la réussite future, la conscience phonologique et la connaissance des lettres, qui sont entraînées en classe.

L'entraînement systématique de la conscience phonologique constitue une étape indispensable vers l'apprentissage de la lecture pour les élèves les plus faibles en conscience phonologique. Si la plupart des enfants accèdent à la lecture sans qu'un travail explicite et intensif sur la conscience phonologique ne soit nécessaire, pour les plus faibles, l'intervention spécifique à l'école est nécessaire, avant de débiter l'apprentissage de la lecture (décodage), c'est-à-dire avant le CP.

Une douzaine d'heures d'entraînement en petits groupes de niveau homogène sont nécessaires. Le protocole est utilisable en classe, soit en grande section de maternelle, soit en tout début de CP tout en étant toutefois relativement exigeant. La fréquence quotidienne des séances durant quelques semaines et le temps global d'implication des élèves dans les tâches proposées sont des conditions incontournables d'efficacité du dispositif.

Les enseignements délivrés à l'ensemble de la classe, alors même que les disparités de niveaux, notamment en vocabulaire, langage, et compréhension orale, privent les élèves qui en auraient le plus besoin de l'accès au discours de l'enseignant. L'enseignement structuré, qui repose sur une connaissance fine des besoins de chaque élève ou groupe d'élèves, permet d'ajuster son intervention aux capacités de chacun, et ainsi d'éviter le décrochage des plus faibles tout en optimisant les apprentissages des plus avancés.

2.2.4 Limites de transférabilité

Les enseignants qui ont manifesté leur refus d'adhésion au programme ont invoqué deux facteurs principaux :

- la « lourdeur du dispositif » qui exige que l'enseignant consacre au moins 90 minutes par jour à un protocole très formalisé ;
- « l'aspect mécanique » du protocole qui vise en effet l'automatisation de compétences cognitives par la répétition d'entraînements.

Ces deux caractéristiques du programme (temps et répétition des entraînements), qui ne sont pas modifiables, signent les limites de transférabilité du dispositif dans la perspective d'envisager une diffusion plus large.

CONCLUSION GENERALE

Sur les dimensions spécifiquement travaillées dans le cadre du dispositif, les élèves ayant participé à l'expérimentation du projet ont globalement plus progressé que les élèves du groupe témoin. Les progrès sont plus importants pour les élèves les plus faibles.

Sur la première année de l'expérimentation qui a porté sur l'année de grande section de maternelle, les élèves des classes expérimentales progressent davantage sur les dimensions Lettres (+ 17 %), Phonologie (+ 22 %), Lecture voie non lexicale (+ 41 %) et Mathématiques (+ 23 %). Pour ces domaines, les progrès sont importants et statistiquement significatifs. En revanche, pour les dimensions Vocabulaire et Lecture voie lexicale, les effets ne sont pas significatifs. En compréhension, l'effet est légèrement significatif. L'analyse complémentaire, selon les déciles et non plus en moyenne, montre des effets différents selon le niveau du score observé au deuxième temps de mesure. En phonologie, ces effets sont plus importants pour les élèves les plus faibles et en compréhension, ils sont significatifs uniquement pour les élèves les plus faibles.

Au CP, les résultats des analyses montrent un effet positif et significatif dans la dimension habiletés phonologiques, aussi bien pour les élèves ayant expérimenté en GS et en CP que pour ceux ayant expérimenté en CP uniquement. En lecture voie non lexicale, on observe également un effet positif et significatif mais uniquement pour les élèves n'ayant expérimenté qu'en CP. Ce résultat ne s'observe pas pour les élèves ayant expérimenté en GS et en CP. Cette absence d'effet significatif en moyenne est complétée par une analyse selon le niveau des élèves puisque le dispositif cible les plus faibles. L'analyse complémentaire précise des effets différents selon le niveau du score observé en fin de CP et notamment, en phonologie et en lecture voie non lexicale, les effets sont plus importants pour les élèves les plus faibles tandis qu'en compréhension orale, les effets négatifs concernent plus particulièrement les élèves de niveau moyen.

De fait, selon l'hypothèse que les deux échantillons témoins (GS et CP) sont comparables et qu'on n'observe pas d'évolution différentielle de la dispersion, l'effet global du dispositif sur deux années serait caractérisé par une progression significative en phonologie (+ 51 %), en lecture voie non lexicale (+ 48 %) avec des effets importants pour les élèves les plus faibles, une absence d'effet sur les dimensions compréhension et lecture voie lexicale.

Les recherches concordantes sur les effets positifs d'une intervention précoce en phonologie proviennent principalement d'études américaines, or, en France il existe peu d'études empiriques d'envergure sur le sujet. De ce point de vue, l'expérimentation fournit des éléments utiles en termes de recherche et d'action pédagogique.

Des analyses complémentaires viendront compléter ces résultats. Une évaluation menée en CE1 durant l'année 2013-2014 devrait permettre de déterminer si les effets observés en grande section et en CP sont transférés ou non en CE1. À plus long terme, une évaluation des mêmes élèves en fin de cours moyen deuxième année pourrait être menée afin de mesurer le transfert sur des compétences de plus haut niveau.

Il y a en effet un enjeu à observer si les résultats positifs observés en début de cycle 2 sont durables pour ne pas dire cumulatifs dans l'apprentissage de la lecture chez ces élèves. En tout état de cause, les évaluateurs proposent de chercher à tirer les enseignements les plus solides de cette expérimentation avant d'en inférer quelques implications et recommandations en termes d'inflexion de politique éducative.



BIBLIOGRAPHIE

BIBLIOGRAPHIE CONSTRUCTION DES EPREUVES T1

Ouvrages

Ecalle, J., & Magnan, A. (2002b). *L'apprentissage de la lecture. Fonctionnement et développement cognitifs*. Paris: A. Colin.

Ecalle, J., & Magnan, A. (2010). *L'apprentissage de la lecture et ses difficultés*. Paris: Dunod.

Epreuve Lettres

Biot-Chevrier, C., Ecalle, J., & Magnan, A. (2009). Poids des connaissances alphabétiques précoces dans l'acquisition de l'écrit : une étude longitudinale auprès d'enfants de 5 à 10 ans. *Archives de Psychologie*, 74, 204-226.

Biot-Chevrier, C., Ecalle, J., & Magnan, A. (2008). Pourquoi la connaissance du nom des lettres est-elle si importante dans l'apprentissage de la langue écrite ? *Revue Française de Pédagogie*, 162, 15-27.

Ecalle, J. (2004). Les connaissances des lettres et l'écriture du prénom chez l'enfant français avant l'enseignement formel de la lecture-écriture. *Psychologie Canadienne*, 45(1), 111-118.

Ecalle, J., Magnan, A., & Biot-Chevrier, C. (2008). Alphabet knowledge and early literacy skills in French beginning readers. *European Journal of Developmental Psychology*, 5(3), 303-325.

Labat, H., Ecalle, J., & Magnan, A. (2010). Effet d'entraînements bi-modaux à la connaissance des lettres. Etude transversale chez des enfants de 3 et 5 ans. *Psychologie Française*, 55, 113-127.

Labat, H., Magnan, A., & Ecalle, J. (in press). Effet d'une exploration multisensorielle séquentielle orientée sur le développement de la lecture et de l'écriture chez les enfants de 5 ans faibles connaisseurs de lettres. *L'Année Psychologique*.

Epreuve Habiletés phonologiques

Ecalle, J., & Magnan, A. (2007). Development of phonological skills and learning to read in French. *European Journal of Psychology of Education*, 22(2), 153-167.

Sanchez, M., Ecalle, J., Magnan, A. (2010). Sensibilité précoce à la structure du mot oral et écrit: quels liens prédictifs avec les performances ultérieures en lecture? *ANAE*, 107-108, 111-118.

Sanchez, M., Magnan, A. & Ecalle, J. (2007). Habiletés phonologiques chez des enfants dysphasiques de GS et CP: étude comparative avec des enfants au développement langagier normal. *Psychologie Française*, 52, 41-54.

Sanchez, M., Magnan, A., Ecalle, J., & (in press). Knowledge about word structure in beginning readers: What specific links are there with word reading and spelling? *European Journal of Psychology of Education*, DOI: 10.1007/s10212-011-0071-8

Epreuve Compréhension de récit

Ecalle, J., Bouchafa, H., Potocki, A., & Magnan, A. (in press). Comprehension of written sentences as a core component of children's reading comprehension. *Journal of Research in Reading* DOI: 10.1111/j.1467-9817.2011.01491.x

Cain, K., & Oakhill, J.V. (1999). Inference making ability and its relation to comprehension failure. *Reading and Writing*, 11, 489-503.

Epreuve Vocabulaire

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28(4), 661-677.

BIBLIOGRAPHIE GÉNÉRALE

- Anthony J.L. & Francis D.J. (2005). "Development of phonological awareness". *Current Directions in Psychological Science*, vol. 14 n°5, p. 255-259.
- ANTHONY J.L., LONIGAN C.J., BURGESS S.R., DRISCOLL K., PHILLIPS B.M., CANTOR B.G. (2002). « Structure of Preschool Phonological Sensitivity: Overlapping Sensitivity to Rhyme, Words, Syllables, and Phonemes ». *Journal of Experimental Child Psychology*, vol 82, n°1, p. 65–92.
- ANTHONY J.L., LONIGAN C.J., DRISCOLL K., PHILLIPS B.M. & BURGESS S.R. (2003). « Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations ». *Reading Research Quarterly*, vol. 38, p. 470–487
- BIANCO M., BRESSOUX P., DOYEN A.L., LAMBERT E., LIMA L., PELLENQ C., & ZORMAN M. (2010). « Early training in oral comprehension and phonological skills: Results of a three-year longitudinal study ». *Scientific Studies of Reading*, vol. 14, n° 3, p. 211-246.
- BIANCO M., PELLENQ C., LAMBERT E., BRESSOUX P., LIMA, L. & DOYEN A.L. (2011). « Impact of early-code skill and oral comprehension training on reading achievement in first grade ». *Journal of Research of Reading*.
- BILLARD C., FLUSS J., DUCOT B., WARSZAWSKI J., ECALLE J., MAGNAN A., RICHARD G. & ZIEGLER J. (2008). « Étude des facteurs liés aux difficultés d'apprentissage de la lecture. À partir d'un échantillon de 1062 enfants de seconde année d'école élémentaire ». *Archives de Pédiatrie*, Vol.15, n°6, p.1058-1067.
- Biot-Chevrier C., Ecalle, J. & Magnan, A. (2009). Poids des connaissances alphabétiques précoces dans l'acquisition de l'écrit : une étude longitudinale auprès d'enfants de 5 à 10 ans. *Archives de Psychologie*, 74, 204-226.
- BONJOUR E. & GOMBERT J. E. (2004). « Profils de lecteurs à l'entrée en sixième », *L'orientation scolaire et professionnelle*, vol. 33, n°1, p. 69-101.
- BRIGAUDIOT M. (2000). *Apprentissage progressif de l'écrit à l'école maternelle*. Paris : Hachette.
- Cain K. & Oakhill J.V. (1999). « Inference making and its relation to comprehension failure ». *Reading and Writing*, vol. 11, p. 489–503.
- Cain K. & Oakhill J.V. (2007). *Introduction to comprehension development*. New York, NY, US: Guilford Press, New York, NY.
- CALVO M.G. (2005). "Relative contribution of vocabulary knowledge and working memory span to elaborative inferences in reading ». *Learning and Individual Differences*, vol. 25, n°1, p.53-65.
- CASALIS S., & LOUIS-ALEXANDRE, M. F. (2000). « Morphological analysis, phonological analysis and learning to read French: A longitudinal study. *Reading and Writing*, vol. 12, p. 303–335.
- Castles A. & Coltheart M. (2004). "Is there a causal link from phonological awareness to success in learning to read?" *Cognition*, vol. 91 n°1, p. 77-111.
- Daussin, J.-M., Keskpaiik, S. & Rocher, T. (2011). L'évolution du nombre d'élèves en difficulté face à l'écrit depuis une dizaine d'années, *France Portrait Social*, pp 137-152.
- DUNCAN L.G., SEYMOUR P.H.K., COLE P. & MAGNAN A. (2006). « Differing sequences of metaphonological development in French and English ». *Journal of Child Language*, vol. 33, n°2, p. 369-399.
- ECALLE J. (2007). « THaPho. Test des Habiletés Phonologiques ». Paris : Editions Mot-à-Mot.
- ECALLE J. & MAGNAN A. (2010). *L'apprentissage de la lecture et ses difficultés*. Paris : Dunod.
- ECALLE J., MAGNAN A. & CALMUS, C. (2009). « How computer-assisted learning using ortho-phonological units could improve literacy skills in low-progress readers ». *Computers & Education*, vol. 52, n°3, p.554-561.
- Ecalle J., Magnan A. & Gibert, F. (2006). Class size effects on literacy skills and literacy interest in first grade: A large-scale investigation. *Journal of School Psychology*, 44, 191-209.
- EHRI L.C., NUNES S. R., WILLOWS D. M., SCHUSTER B.V., YAGHOUB-ZADEH Z. & SHANAHAN, T. (2001). « Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis ». *Reading Research Quarterly*, vol. 36, p.250-287.
- EVANS M.A., BELL M., SHAW D., MORETTI S. & PAGE J. (2006). « Letter names, letter sounds and phonological awareness: an examination of kindergarten children across letters and of letters across children ». *Reading and writing*, vol. 19, n°9, p.959-989.
- FLORIT E., ROCH M., ALTOE G. & LEVORATO M.C. (2009). « Listening comprehension in preschoolers: The role of memory. » *British Journal of Developmental Psychology*, vol 27, n°4, p. 935-951.

- FOULIN J.N. (2007). « La connaissance des lettres chez les prélecteurs : aspects pronostiques, fonctionnels et diagnostiques Letter knowledge in pre-readers: prognostic, functional, and diagnostic aspects ». *Psychologie Française*, vol.52, n°4, p. 431-444.
- GOUGH P.B. & TUNMER W.E. (1986). « Decoding, Reading, and reading disability ». *Remedial and Special Education*, vol.7, p. 6-10.
- KEARNS G. & BIEMILLER A. (2011). « Two-questions vocabulary assessment: Developing a new method for group testing in kindergarten through second Grade. » *Journal of Education*, vol, 190 n°1/2, p. 31-41.
- Khoms A. (2001). *Evaluation du langage oral*, ECPS, Paris. Distribué par TEMA.
- Labat H., Magnan A. & Ecalte J. (2011). « Effet d'une exploration "multisensorielle séquentielle orientée" sur le développement de la compréhension du principe alphabétique chez les enfants de 5 ans faibles connaisseurs de lettres. » *L'Année Psychologique*, vol 111, n°4, p. 641-671.
- Lecocq P. (1996). *Epreuve de Compréhension Syntaxico-Sémantique*, Presses universitaires du Septentrion.
- LONIGAN C.J., BURGESS S.R. & Anthony, J.L. (2000). « Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study ». *Developmental Psychology*, vol.36, p. 596-613.
- MELBY-LERVAG M., LYSTER S. & HULME C. (2012). « Phonological skills and their role in learning to read : A meta-analysis review ». *Psychological Bulletin*, vol. 138, p. 322-352.
- MIALARET G. (1954). « Pratique journalière et recherche scientifique en pédagogie ». *Bulletin de Psychologie*, vol. 5, n°, p. 309-313.
- MJENR. (1999). « Etude spécifique relative aux élèves en difficulté en lecture à l'entrée en 6e (septembre 1997) ». *Dossiers*, n°112.
- MORAIS J., BERTELSON P., CARY L. & ALEGRIA J. (1986). « Literacy training and speech segmentation ». *Cognition*, vol. 24, n°1/2, p.45-64.
- New B., Pallier C., Brysbaert M. & Ferrand L. (2004). "Lexique 2: A new french lexical database. *Behavior Research Methods, Instruments & Computers*", vol. 36, n°3, p. 516-524.
- OAKHILL J.V. & CAIN K. (2012). "The precursors of reading ability in young readers : evidence from a four-year longitudinal study ». *Scientific studies of reading*, vol. 16, n°2, p. 91-121.
- OCDE. (2009). *Résultats du PISA 2009 : Synthèse*. PISA, Editions :OCDE
- ONL. (1996). *Regard sur la lecture et les apprentissages*. Editions :ONL
- RAMUS F. (2008). « Quelles évaluations pour les applications des sciences et technologies cognitives? Prospective Interdisciplinaire en réseau pour les sciences et technologies cognitives ». En ligne : <<http://pirstec.risc.cnrs.fr/contribs/afficher/29>>
- Sanchez M., Magnan A. & Ecalte J. (2007). « Habiletés phonologiques chez des enfants dysphasiques de GS et CP: Étude comparative avec des enfants au développement langagier normal. » *Psychologie Française*, vol 52 n°1, p. 41-54.
- SCARBOROUGH . (1998). « Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. » *Annals of Dyslexia*, vol 48, p. 115-136.
- Schatschneider C., Francis D.J., FOORMAN B.R., Fletcher J.M., METHA P. (1999). « The dimensionality of phonological awareness: An application of item response theory ». *Journal of Educational Psychology*, vol 91, n°3, p. 439-449.
- SENECHAL M., OUELETTE G., & RODNEY,D. (2006). « The misunderstood giant: On the predictive role of early vocabulary to future reading. In S. B. Neuman & D. Dickinson (Eds.), *Handbook of early literacy research: Vol. 2*. New York: Guilford Press, p. 173–182.
- Treiman R. (2006). "Knowledge about letters as a foundation for reading and spelling". In R.M. Joshi & P.G. Aaron, *Handbook of orthography and literacy*. Ed. Mahwah, NJ: Erlbaum, p. 581-599.
- VLOEDGRAVEN J. & VERHOEVEN L. (2009). "The nature of phonological awareness throughout the elementary grades: An item response theory perspective ». *Learning and Individual Differences*, vol.19, n° 2, p. 161–169.



ANNEXE 1 : PROTOCOLE D'ENQUÊTE

Instruments construits pour l'évaluation du projet Lecture en novembre 2011 en grande section de maternelle (T1 GS).

Procédure

Les épreuves administrées semi-collectivement (groupes de 4 à 8 enfants) nécessitent une passation qui respecte strictement les consignes. Les 4 épreuves ne devaient pas être administrées lors d'une seule passation (effet de fatigue).

Descriptif des épreuves

Lettres

La tâche de reconnaissance de lettres consiste à entourer la lettre dénommée parmi 7 lettres. Les 26 lettres sont présentées dans un ordre aléatoire. (26 items)

Habilités phonologiques

Deux tâches sont proposées. La tâche de détection d'intrus consiste à barrer l'image (mot dénommé) qui ne possède pas d'unité commune avec deux autres mots. L'unité commune est une syllabe ou un phonème en position initiale ou finale. Douze items (6 dans la condition "syllabe" et 6 condition "phonème") sont traités (max: 12 points).

Dans la tâche de suppression de syllabe à choix forcé, l'enfant doit à partir d'un mot dont la première syllabe est supprimée retrouver un autre mot (*pinceau - seau*). Quatre items sont proposés, l'item cible (*seau*) et trois items distracteurs, l'un phonologique (*landau*), un autre sémantique (*crayon*) et un intrus (*cochon*). (6 items).

Compréhension de récit

Après la lecture d'une courte histoire (texte de 11 phrases, 136 mots; Brigaudiot, 2000), 12 questions à choix forcé sont posées à l'enfant. Quatre types de questions sont proposées, littérales, inférentielles (Cain & Oakhill, 1999). Pour les inférences, on distingue les inférences de cohésion locale, de cohésion globale nécessaires et non nécessaires (4*3 items). (12 items).

Vocabulaire

Dans une épreuve classique de vocabulaire en réception, l'enfant doit entourer un mot parmi 4 représentés par des images. Le choix des mots cibles a été effectué en fonction de leur fréquence lexicale à l'oral (Lexique 3; New et al, 2007), des mots fréquents aux mots plus rares. Ils sont présentés dans un ordre aléatoire. Les items tests, outre l'item cible sont composés de 3 intrus dont un distracteur sémantique (item appartenant au même champ sémantique que l'item cible). (24 items)

Instruments construits pour l'évaluation du projet Lecture en juin 2012 en grande section de maternelle (T2 GS).

Procédure

Les épreuves administrées semi-collectivement (groupes de 4 à 8 enfants) nécessitent une passation qui respecte strictement les consignes. Les épreuves ne devaient pas être administrées lors d'une seule passation (effet de fatigue).

Descriptif des épreuves

Afin d'assurer un ancrage, certaines épreuves constituent des reprises du temps de mesure précédent.

Vocabulaire

Reprise partielle de l'épreuve T1 (14 items).

Lettres

Reprise partielle de l'épreuve T1 (15 items).

Compréhension de récit

Reprise de l'épreuve T1 (12 items).

Habilités phonologiques

Les deux tâches sont proposées au T1 : tâche de détection d'intrus (12 items), tâche de suppression de syllabe (6 items).

Lecture de mots

Dans une épreuve de connaissance du lexique orthographique, l'enfant doit entourer le mot cible énoncé parmi 5 items tests. Les items, outre le mot cible, sont composés de 4 intrus dont un distracteur phonologique (10 items).

Lecture de pseudo- mots

Dans une épreuve d'identification visuelle de mots inconsistants, l'enfant doit entourer le mot cible dicté parmi 5 items tests. Les items, outre le mot cible, sont composés de 4 intrus (10 items).

Mathématiques

Descriptif des exercices

Les exercices numériques, issus des épreuves DEPP Panel CP 1997⁵, incluaient 9 items ainsi conçus :

a) Un exercice de complèment de la chaîne écrite : les élèves devaient ajouter dans des cases prévues à cet effet : 3, 17, 29 (3 items).

⁵ *Les compétences des élèves à l'entrée du cours préparatoire, études réalisées à partir du panel d'écoliers recrutés en 1997, dossier 132, août 2002, MEN*

- b) Un exercice de comparaisons de quantités, dans lesquelles deux alignements de points étaient présentés. Ces deux alignements comportaient soit le même nombre soit des nombres différents de jetons. Par ailleurs la longueur des alignements était soit congruente avec la numérosité soit non congruente avec celle-ci (3 items).
- c) Des problèmes arithmétiques simples (à une opération) dont les enfants devaient marquer le résultat en entourant pour chacun un nombre parmi 6 présentés (3 items).
- d) Un exercice de dénombrement dans laquelle les enfants devaient dénombrer 17 objets et sélectionner le nombre obtenu parmi 6 présentés à la fois sous forme de dominos et en écriture de chiffres arabes (1 item).

Justification des épreuves numériques

Les traitements numériques nécessitent que les dénominations des nombres soient connues ainsi que leur transcription en chiffres arabes. La batterie élaborée comprend ainsi une première série d'épreuves qui vise à évaluer la maîtrise de cette connaissance par les enfants. Elle s'attache plus particulièrement à la capacité de compléter la suite des nombres par des items occupant des positions particulières : 3 est particulièrement facile ; 17 marque le passage de l'utilisation d'un mot unique (quinze, seize) pour désigner une quantité à une combinaison de mots (dix sept) faisant intervenir une dizaine (dix) et une unité (sept) dans un ordre précis. Le nombre 29 évalue la même capacité (vingt et neuf) à un niveau plus avancé de la chaîne, juste avant le passage à la dizaine supérieure (30) Ces épreuves permettent de déterminer si les enfants connaissent les chiffres arabes et disposent de connaissances quant à leur fonctionnement. Il est clair que ces épreuves ne permettent en rien de savoir si les enfants sont en mesure de se représenter les quantités correspondantes.

Les traitements numériques nécessitent ensuite une certaine connaissance de la cardinalité. Notamment celle-ci doit être indépendante des caractéristiques perceptives des collections, longueur et densité entre autres. La deuxième série d'épreuves cherche précisément à évaluer les poids respectifs de la cardinalité et de la perception dans des exercices de comparaison numérique (est-ce qu'il y a pareil de points en dessus et en dessous de la ligne ?). Ces épreuves sont directement inspirées des travaux ayant trait à la conservation du nombre. Elles en constituent une adaptation destinée à une passation collective. Dans certaines épreuves, la comparaison peut réussir en s'appuyant sur l'évaluation des longueurs des rangées ou de la densité des points ou enfin sur la détermination de la quantité numérique.

Les traitements numériques nécessitent encore que les enfants soient capables d'imaginer les situations évoquées par de brefs problèmes arithmétiques et d'effectuer sur les données numériques des transformations simples, additions ou soustractions. Afin d'évaluer ces capacités, deux problèmes arithmétiques requérant chacun une seule opération ont été proposés. Le premier correspond à une addition ($7 + 2$), l'autre renvoie à une "soustraction" dont il n'est pas sûr qu'elle soit en général plus difficile que l'addition sur le plan conceptuel alors qu'elle le devient très vite pour ce qui concerne la résolution des opérations.

Les traitements numériques font enfin appel au dénombrement. Celui-ci fait l'objet d'un item) dans lequel les enfants doivent dénombrer des objets et choisir parmi plusieurs nombres (indiqués par des dominos et des chiffres arabes) lequel correspond au cardinal de la collection.

Instruments construits pour l'évaluation du projet Lecture en novembre 2012 en classe de Cours Préparatoire (T3 CP).

Procédure

Les épreuves administrées collectivement (groupe classe) nécessitent une passation qui respecte strictement les consignes. Les épreuves étaient administrées en trois séquences de 30 minutes afin d'éviter les effets de fatigue.

Descriptif des épreuves

Afin d'assurer un ancrage, certaines épreuves constituent des reprises du temps de mesure précédent.

Pour ce troisième moment de mesure, six domaines sont évalués, la connaissance des lettres, les habiletés phonologiques, la compréhension orale de récit et le vocabulaire, la lecture de mots et de pseudo-mots et des éléments de mathématiques.

Vocabulaire

Nouvelle épreuve construite (12 items)

Chaque item est constitué de deux tâches :

- 1- retrouver l'image d'un mot dicté parmi 4 (reprise phase 1&2)
- 2- question fermée de vocabulaire sur le mot cible (profondeur du vocabulaire)

Lettres

Nouvelle épreuve : connaissances des lettres. Il s'agit pour l'élève d'écrire des lettres dictées par l'enseignant (15 items).

Compréhension de récit

Reprise de l'épreuve T1 (12 items).

Habiletés phonologiques

Une nouvelle tâche est proposée, il s'agit de segmentation phonémique. L'élève doit dessiner autant de ronds que de sons d'un pseudo-mot dicté par l'enseignant (10 items).

Lecture de mots

Reprise de l'épreuve T2 (10 items).

Lecture de pseudo- mots

Reprise de l'épreuve T2 (10 items).

Mathématiques

Reprise de l'épreuve T2 (9 items).

Instruments utilisés pour l'évaluation du projet Lecture en juin 2013 en classe de cours préparatoire (T4 CP).

Procédure

Les épreuves administrées collectivement (groupe classe) nécessitent une passation qui respecte strictement les consignes. Les épreuves étaient administrées en quatre séquences de 30 minutes afin d'éviter les effets de fatigue.

Descriptif des épreuves

Afin d'assurer un ancrage, certaines épreuves constituent des reprises du temps de mesure précédent.

La batterie de test a été complétée par d'autres épreuves permettant d'analyser plus finement les compétences prédictives de la réussite de l'apprentissage de la lecture et de l'écriture. Pour ce quatrième moment de mesure, huit domaines sont évalués, la connaissance des lettres, les habiletés phonologiques, la compréhension orale et écrite de récit, le vocabulaire, la lecture de mots et de pseudo-mots, l'orthographe lexicale, la fluence et des éléments de mathématiques.

Vocabulaire

Reprise partielle de l'épreuve T3 (4 items).

Nouvelle épreuve (30 items) testant les connaissances lexicales des élèves.

Compréhension orale de récit

Reprise de l'épreuve T1 (12 items).

Compréhension écrite de récit

Nouvelle épreuve. Il s'agit pour l'élève de lire seul un texte silencieusement (texte de 10 phrases, 110 mots) puis de répondre à des questions à choix forcé. Quatre types de questions sont proposées, littérales, inférentielles (Cain & Oakhill, 1999). Pour les inférences, on distingue les inférences de cohésion locale, de cohésion globale nécessaires et non nécessaires (12 items).

Habiletés phonologiques

Reprise de l'épreuve T3 (10 items).

Production orthographique

Nouvelle épreuve. L'élève doit écrire sous la dictée des mots (12 items).

Lecture de mots

Reprise de l'épreuve T2 (10 items).

Lecture de pseudo- mots

Reprise de l'épreuve T2 (10 items) + (6 items nouveaux)

Mathématiques

Reprise de l'épreuve T2 complétée de 10 items évaluant la connaissance des nombres, les additions posées et la résolution de problèmes soit au total 19 items.

Fluence

La tâche consiste à dire si une paire d'items constituée d'un mot et d'un pseudo-mot se prononcent de la même façon (pain/paim; classe/clave). Le temps imparti est de 1 min.



ANNEXE 2 : SYNTHÈSE D'ENQUÊTE

Tableau synthétique T1T2 : GS (2011-2012)

Sélection des deux groupes	Mode de sélection	Les écoles de l'échantillon témoin T1T2 sont tirées selon un plan de sondage stratifié, de façon à présenter des caractéristiques proches de celles du groupe expérimental. Dans les écoles tirées, tous les élèves de grande section de maternelle sont évalués.
	Unité de sélection	Écoles
	Critères de sélection	La stratification est faite selon 3 dimensions : <ul style="list-style-type: none"> - éducation prioritaire ou non ; - département géographique ; - taille de l'école en termes d'effectif de GS. L'échantillon témoin est tiré de manière à ce que la répartition des écoles dans les strates soit la même que pour le groupe expérimental.
Comparabilité des groupes	Nombre d'enquêtes réalisées	2 moments de mesure : <ul style="list-style-type: none"> - 14 au 25 novembre 2011 - 4 juin au 15 juin 2012
	Taux de réponse	Il ressort que sur les 5 900 élèves attendus, 4 468 élèves de GS ont participé au temps 1 (soit 76 %) et 3 885 élèves au temps 2 (66 %). Au final, 3 569 élèves ont participé aux deux sessions d'évaluation (60 %).
	Taux d'attrition	Parmi les 4 468 élèves ayant participé au 1er temps de mesure, 899 (soit 20 %) n'ont pas participé au 2ème temps de mesure (19 % pour le groupe expérimental, et 21 % pour le groupe témoin) Cette attrition est due pour plus de la moitié à une non réponse au niveau école : 6 écoles du groupe expérimental et 6 écoles du groupe témoin n'ont pas du tout participé au 2ème temps de mesure, ce qui représente 499 élèves qui avaient participé au 1er temps de mesure (247 élèves du groupe expérimental et 252 élèves du groupe témoin). On observe également une non-réponse au niveau classe : pour 3 écoles du groupe expérimental et 3 écoles du groupe témoin la différence d'effectif entre les deux temps de mesure laisse penser qu'une ou plusieurs classes entières n'ont pas participé au 2ème temps (ce qui représente 195 élèves, dont 126 du groupe expérimental et 69 du groupe témoin). Le reste de l'attrition se situe au niveau élève.

Recueil des données	Fonction du passateur	Enseignant des classes sélectionnées
	Mode et conditions de passation	C'est l'enseignant de la classe qui a assuré la passation des épreuves. Aucune passation ne devait être confiée à une ou plusieurs personnes extérieures à la classe. Les épreuves ont été administrées semi-collectivement (groupes de 4 à 8 enfants).

Tableau synthétique T3T4 : CP (2012-2013)

Sélection des deux groupes	Mode de sélection	Les écoles de l'échantillon témoin T3T4 sont tirées selon un plan de sondage stratifié, de façon à présenter des caractéristiques proches de celles du groupe expérimental. Dans les écoles tirées, tous les élèves de CP sont évalués.
	Unité de sélection	Écoles
	Critères de sélection	La stratification est faite selon 3 dimensions : <ul style="list-style-type: none"> - éducation prioritaire ou non ; - département géographique (4 départements) ; - score total moyen aux évaluations CE1 2011 (4 groupes constitués selon les quartiles observés dans la base obtenue). L'échantillon témoin est tiré de manière à ce que la répartition des écoles dans les strates soit la même que pour le groupe expérimental.
Comparabilité des groupes	Nombre d'enquêtes réalisées	2 moments de mesure : <ul style="list-style-type: none"> - 13 au 24 novembre 2012 - 3 juin au 14 juin 2013
	Taux de réponse	Au sein des 28 écoles concernées par l'expérimentation, 1 113 élèves ont participé aux évaluations des temps 3 (début de CP) et 4 (fin de CP). Parmi eux, 728 élèves ont suivi le protocole expérimental en GS et en CP, et 369 élèves l'ont suivi en CP mais pas en GS. Ont également été évalués 16 élèves qui auraient du suivre le protocole expérimental en CP mais dont la classe n'a finalement pas participé, dont 3 élèves qui l'avaient suivi en GS et 13 élèves qui ne l'ont finalement jamais suivi. Parmi les 91 écoles tirées pour l'échantillon témoin, 85 écoles avaient participé au T3, et 78 écoles ont finalement participé aux T3 et T4, soit 3 534 élèves.
	Taux d'attrition	Au final, 4 647 élèves ont participé aux deux sessions d'évaluation (73,2 %), l'attrition totale est de 26,8 %. Sur les 1 821 élèves prévus concernés par l'expérimentation au T3, 1 113 ont finalement participé aux deux temps de mesure (attrition 39 %) ; concernant l'échantillon témoin sur les 4 531 élèves attendus au T3, 3 534 élèves ont participé aux deux temps de mesure (attrition 22 %)

Recueil des données	Fonction passateur	du	Enseignant des classes sélectionnées
	Mode conditions passation	et de	C'est l'enseignant de la classe qui a assuré la passation des épreuves. Aucune passation ne devait être confiée à une ou plusieurs personnes extérieures à la classe. Les épreuves ont été administrées collectivement (groupes classe).



ANNEXE 3 : ÉCHANTILLONNAGE - STRATIFICATION

Tableau 1 : Quartiles du score total moyen aux évaluations CE1 2011

1er quartile	53.61
médiane	60.48
3e quartile	67.43

Tableau 2 : Répartition des écoles de la base de sondage selon les strates

strate	ep2011	dep	quartile	Nb écoles	%
1	0	059	1	65	4.2 %
2	0	059	2	83	5.4 %
3	0	059	3	121	7.8 %
4	0	059	4	154	9.9 %
5	0	062	1	29	1.9 %
6	0	062	2	74	4.8 %
7	0	062	3	76	4.9 %
8	0	062	4	76	4.9 %
9	0	092	1	12	0.8 %
10	0	092	2	29	1.9 %
11	0	092	3	67	4.3 %
12	0	092	4	106	6.8 %
13	0	093	1	63	4.1 %
14	0	093	2	56	3.6 %
15	0	093	3	49	3.2 %
16	0	093	4	20	1.3 %
17	1	059	1	72	4.6 %
18	1	059	2	63	4.1 %
19	1	059	3	31	2.0 %
20	1	059	4	13	0.8 %
21	1	062	1	45	2.9 %
22	1	062	2	21	1.4 %
23	1	062	3	18	1.2 %
24	1	062	4	12	0.8 %
25	1	092	1	18	1.2 %
26	1	092	2	23	1.5 %
27	1	092	3	14	0.9 %
28	1	092	4	3	0.2 %
29	1	093	1	83	5.4 %
30	1	093	2	39	2.5 %
31	1	093	3	12	0.8 %
32	1	093	4	4	0.3 %

Tableau 3 : Répartition des écoles du groupe expérimental selon les strates

	ep2011	dep	quartile	strate	Nb écoles	%
1	0	062	1	5	1	3.3 %
2	0	062	3	7	1	3.3 %
3	0	092	3	11	6	20.0 %
4	0	092	4	12	2	6.7 %
5	0	093	1	13	1	3.3 %
6	0	093	2	14	2	6.7 %
7	0	093	3	15	2	6.7 %
8	1	059	1	17	2	6.7 %
9	1	059	4	20	1	3.3 %
10	1	062	1	21	2	6.7 %
11	1	062	2	22	1	3.3 %
12	1	092	1	25	2	6.7 %
13	1	092	3	27	2	6.7 %
14	1	093	1	29	2	6.7 %
15	1	093	2	30	3	10.0 %

Tableau 4 : Répartition des écoles de l'échantillon témoin selon les strates

	strate	ep2011	dep	quartile	Nb écoles	%	Nb élèves 2011
1	5	0	062	1	3	3.3 %	94
2	7	0	062	3	3	3.3 %	88
3	11	0	092	3	18	20.0 %	1127
4	12	0	092	4	6	6.7 %	282
5	13	0	093	1	3	3.3 %	195
6	14	0	093	2	6	6.7 %	355
7	15	0	093	3	6	6.7 %	334
8	17	1	059	1	6	6.7 %	173
9	20	1	059	4	3	3.3 %	83
10	21	1	062	1	6	6.7 %	184
11	22	1	062	2	3	3.3 %	104
12	25	1	092	1	6	6.7 %	325
13	27	1	092	3	6	6.7 %	457
14	29	1	093	1	6	6.7 %	302
15	30	1	093	2	9	10.0 %	428



ANNEXE 4 : ANALYSE DES ITEMS T1T2

Analyse des items T1T2

Temps 1

Séquence1 : vocabulaire (4 499 élèves)

23 items. Alpha = 0.82

Aucun Rbis inférieur à 0.2

Séquence2 : reconnaissance des lettres (4 483 élèves)

26 items. Alpha = 0.93

Aucun Rbis inférieur à 0.2

Séquence3 : compréhension orale (4 426 élèves)

12 items. Alpha = 0.72

Aucun Rbis inférieur à 0.2

Séquence4 : habiletés phonologiques - intrus (4 390 élèves)

12 items. Alpha = 0.61

3 items avec un Rbis inférieur à 0.2

pour info	item	Rbis
Phf bande	intrus4	0.19
Phf robe	intrus6	0.09
Phi cœur	intrus9	0.18

Séquence5 : habiletés phonologiques - suppression syllabes (4 347 élèves)

6 items. Alpha = 0.78

Aucun Rbis inférieur à 0.2

Séquences 4 et 5 : habiletés phonologiques (3 854 élèves)

18 items. Alpha = 0.76

2 items avec un Rbis inférieur à 0.2

pour info	item	Rbis
Phf robe	intrus6	0.09
Phi cœur	intrus9	0.18

Ensemble sauf séquence2 (reconnaissance des lettres) (4 578 élèves)

53 items. Alpha = 0.88

3 items avec un Rbis inférieur à 0.2

séquence	pour info	item	Rbis
3 compréhension orale	9.3 InfGn3	Comp9	0.13
4 intrus	Phf robe	intrus6	0.08

4 intrus	Phi cœur	intrus9	0.16
----------	----------	----------------	-------------

Temps 2

Séquence1 : vocabulaire (3 885 élèves)

13 items. Alpha = 0.72

Aucun Rbis inférieur à 0.2

Séquence2 : connaissance des lettres (3 848 élèves)

15 items. Alpha = 0.86

Aucun Rbis inférieur à 0.2

Séquence3 : compréhension orale (3 797 élèves)

12 items. Alpha = 0.73

Aucun Rbis inférieur à 0.2

Séquence4 : habiletés phonologiques - intrus (3 806 élèves)

12 items. Alpha = 0.71

Un item avec un Rbis inférieur à 0.2

pour info	item	Rbis
Phf robe	intrus6	0.15

Séquence5 : habiletés phonologiques - suppression syllabe (3 793 élèves)

6 items. Alpha = 0.77

Aucun Rbis inférieur à 0.2

Séquences 4 et 5 : habiletés phonologiques

18 items. Alpha = 0.79

1 item avec Rbis inférieur à 0.2

pour info	item	Rbis
Phf robe	intrus6	0.16

Séquence6 : Lecture voie lexicale (3 713 élèves)

10 items. Alpha = 0.58

2 items avec un Rbis inférieur à 0.2

pour info	item	Rbis
Rc4F-3l nid	LVLRC4	0.13
Rc6F-5l table	LVLRC6	0.19

Séquence7 : lecture voie non lexicale (3 740 élèves)

10 items. Alpha = 0.77

Aucun Rbis inférieur à 0.2

Séquence8 : maths (3 807 élèves)

10 items. Alpha = 0.72

Aucun Rbis inférieur à 0.2

Conclusion analyse des items T1T2

Séquence 4 habiletés phonologiques – intrus

L'item *intrus6* présente un mauvais Rbis aux deux temps de mesure → **on le supprime aux 2 temps**

pour info	item	Rbis temps1	Rbis temps2
Phf robe	intrus6	0.16	0.15

Les items *intrus4* et *intrus9* ont un mauvais Rbis au temps1, mais pas au temps 2 → **on les conserve aux 2 temps**

pour info	item	Rbis temps1	Rbis Temps2
Phf bande	intrus4	0.19	0.32
Phi cœur	intrus9	0.18	0.43

Séquence6 lecture voie lexicale (temps2 uniquement)

Les items *LVLRC4* et *LVLRC6* ont un mauvais Rbis → **on les supprime**

pour info	item	Rbis
Rc4F-3I nid	LVLRC4	0.13
Rc6F-5I table	LVLRC6	0.19

On supprime les items *intrus6* de la séquence 4 et *LVLRC4* et *LVLRC6* de la séquence 6.

Temps 1 après suppression des mauvais items

Séquence4 : habiletés phonologiques - intrus (4 390 élèves)

11 items. Alpha = 0.63

2 items avec un Rbis inférieur à 0.2

pour info	item	Rbis
Phf bande	intrus4	0.19
Phi cœur	Intrus9	0.18

Séquences 4 et 5 : habiletés phonologiques (3 854 élèves)

16 items. Alpha = 0.77

1 item avec un Rbis inférieur à 0.2

pour info	item	Rbis
Phi cœur	Intrus9	0.18

Synthèse

	séquence	nb d'items au départ	nb d'items conservés	alpha	Facteur d'allongement ¹	Alpha à nb d'items équivalent ²
1	vocabulaire	23	23	0.82	1.13	0.84
2	connaissance des lettres	26	26	0.93	1.00	0.93
3	compréhension orale	12	12	0.72	2.17	0.85
4	intrus	12	11	0.63	2.36	0.80
5	suppression de syllabes	6	6	0.78	4.33	0.94
4 et 5	phonologie	18	17	0.77	1.53	0.84

¹ On reporte toutes les séquences à 26 items. Facteur d'allongement $n = 26/\text{nb d'items de la séquence}$

² Alpha à nb d'items équivalent = $\frac{n\alpha}{1+(n-1)\alpha}$

Temps 2 après suppression des mauvais items

Séquence4 : habiletés phonologiques - intrus

11 items. Alpha = 0.71

Aucun Rbis inférieur à 0.2.

Séquences 4 et 5 : habiletés phonologiques

17 items. Alpha = 0.80

Aucun Rbis inférieur à 0.2.

Séquence6 lecture voie lexicale

8 items. Alpha = 0.58

Aucun Rbis inférieur à 0.2.

	séquence	nb d'items au départ	nb d'items conservés	alpha	Facteur d'allongement ¹	Alpha à nb d'items équivalent ²
1	vocabulaire	13	13	0.72	1.31	0.77
2	connaissance des lettres	15	15	0.86	1.13	0.87
3	compréhension orale	12	12	0.73	1.42	0.79
4	intrus	12	11	0.71	1.55	0.79
5	suppression de syllabes	6	6	0.77	2.83	0.90
4 et 5	phonologie	18	17	0.80	1.00	0.80
6	Lecture voie lexicale	10	8	0.58	2.13	0.75
7	lecture voie non lexicale	10	10	0.77	1.70	0.85
8	mathématiques	10	10	0.72	1.70	0.81

¹ On reporte toutes les séquences à 17 items. Facteur d'allongement $n = 17/\text{nb d'items de la séquence}$

² Alpha à nb d'items équivalent = $\frac{n\alpha}{1+(n-1)\alpha}$

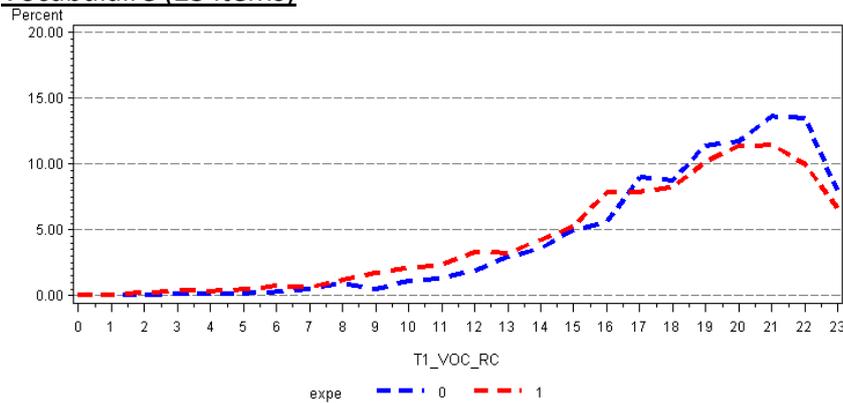


ANNEXE 5 : DISTRIBUTION DES SCORES T1T2

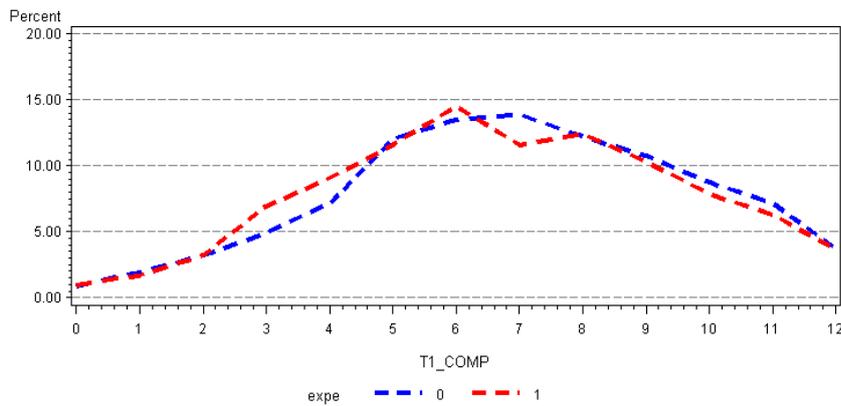
Distribution des scores bruts T1T2

Scores bruts de l'ensemble des élèves ayant participé au temps1 selon le groupe expé / témoin

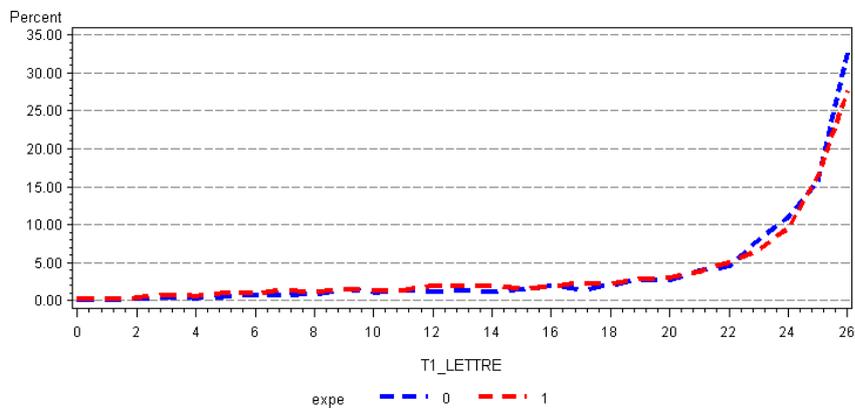
Vocabulaire (23 items)



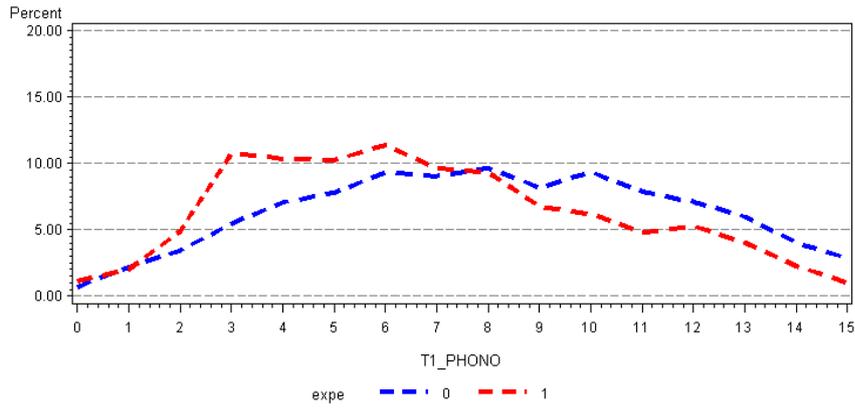
Compréhension orale (12 items)



Connaissance des lettres (26 items)

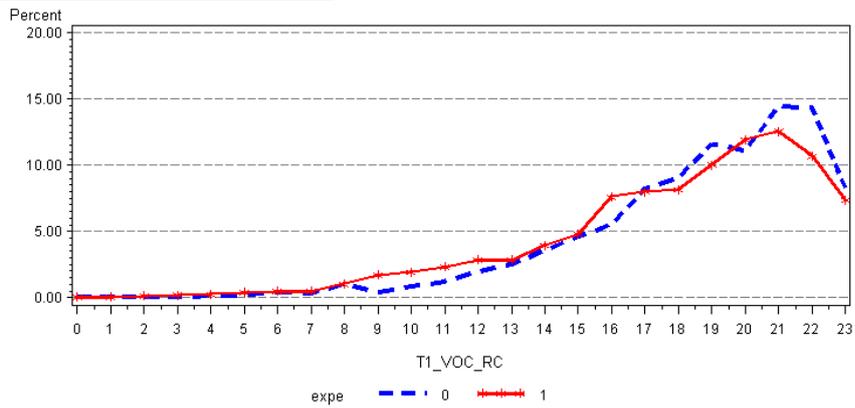


Phonologie (17 items)

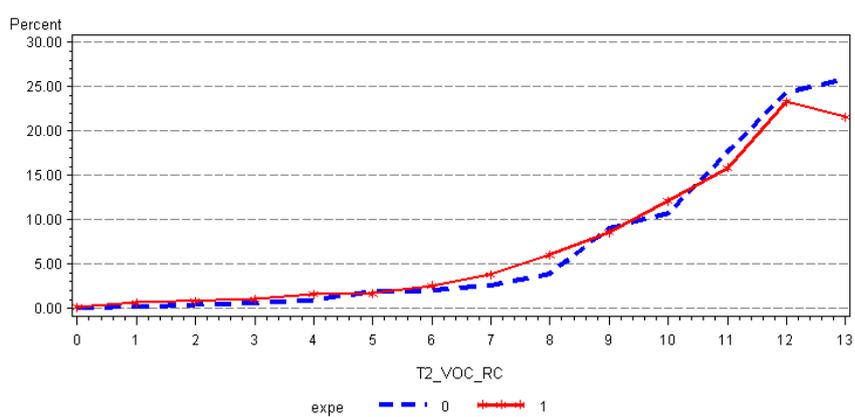


Scores bruts des élèves ayant participé aux deux temps selon le groupe expé / témoin

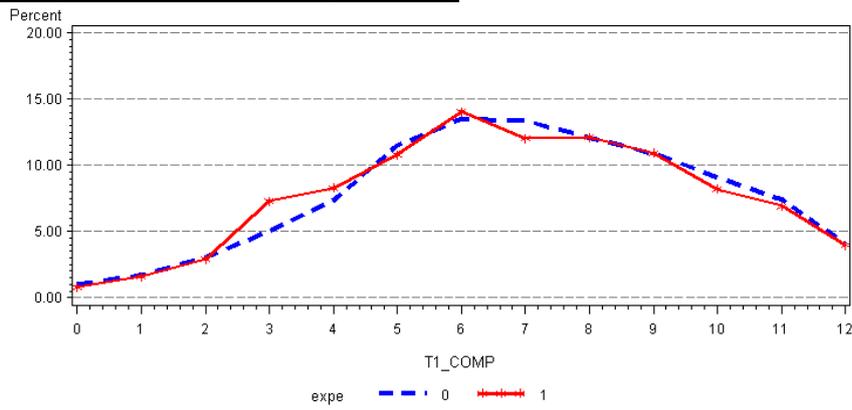
Vocabulaire T1 (23 items)



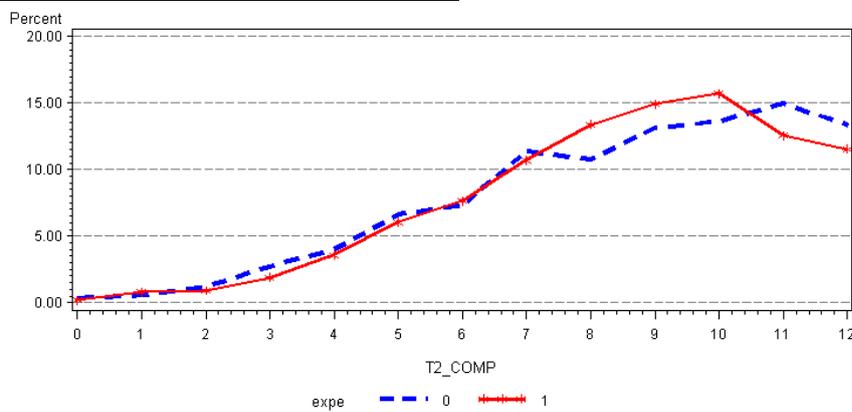
Vocabulaire T2 (13 items)



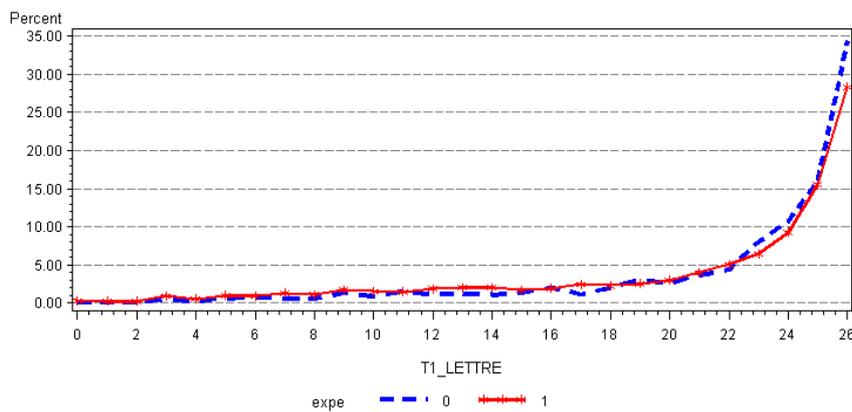
Compréhension orale T1 (12 items)



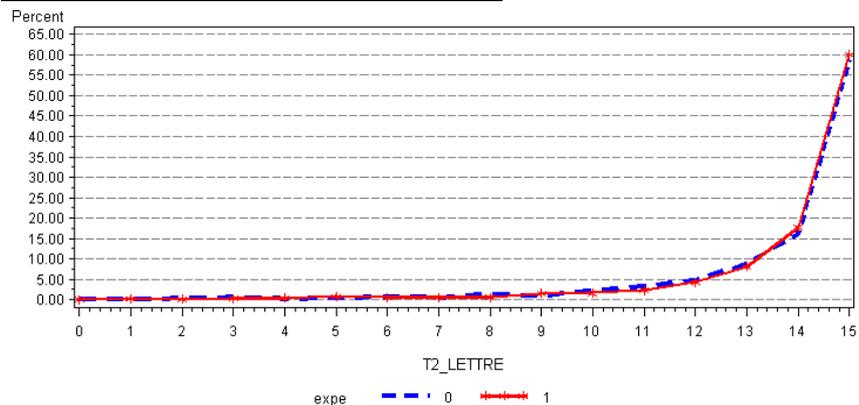
Compréhension orale T2 (12 items)



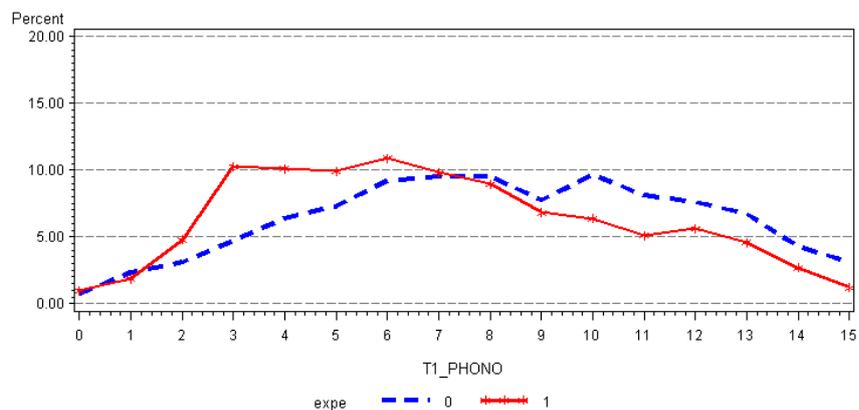
Connaissance des lettres T1 (26 items)



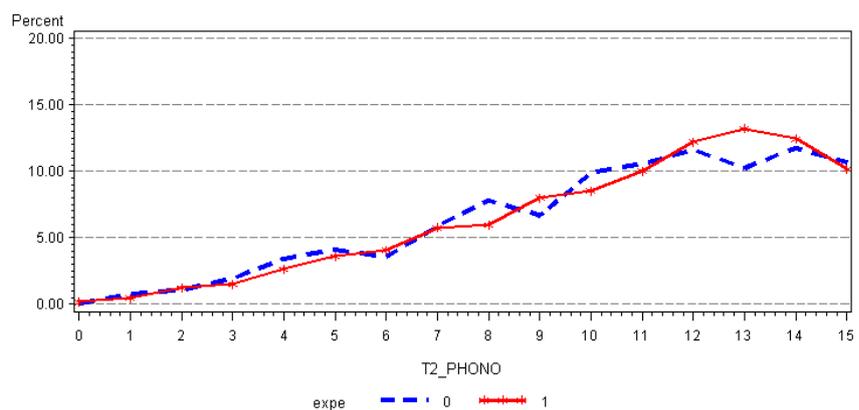
Connaissance des lettres T2 (15 items)



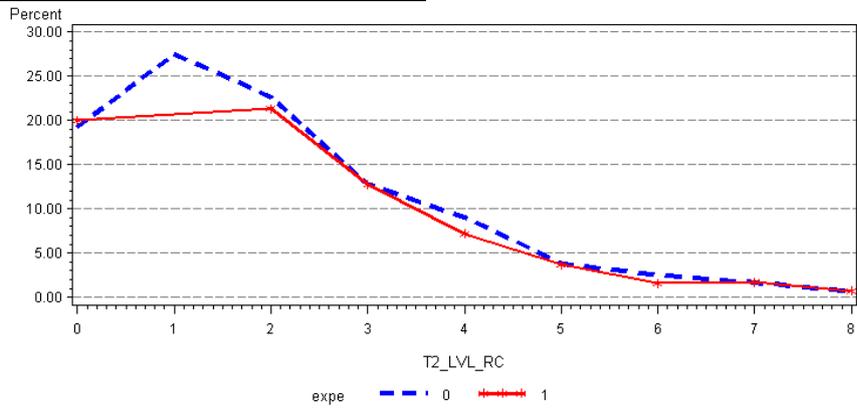
Phonologie T1 (17 items)



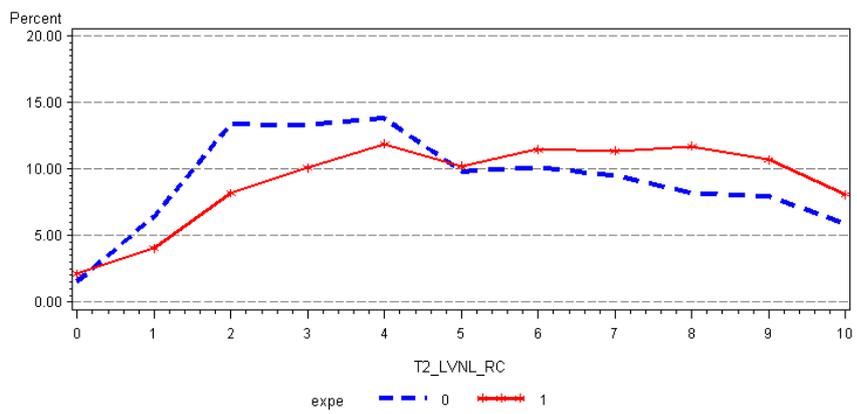
Phonologie T2 (17 items)



Lecture voie lexicale T2 (8 items)



Lecture voie non lexicale T2 (10 items)





ANNEXE 6 : ANALYSE DES ITEMS T3T4

Temps 3

Écriture de lettres

15 items, alpha = 0.84

Compréhension orale de récit

12 items, alpha = 0.71

Vocabulaire

L'épreuve de vocabulaire se compose de 13 items qui sont une reprise des 13 items de vocabulaire du temps 2, mais avec une mise en forme différente (renommés V2) auxquels s'ajoute à chaque fois un item consistant en une question portant sur le sens du mot considéré (items nommés V3).

Si on considère les 26 items pris ensemble l'indice alpha vaut 0.76 et 3 items ont un Rbis inférieur à 0.2

NAME	N	MEAN	Rbis
xV3_6	5 068	0.55	0.12
xV3_11	5 068	0.31	0.03
xV3_14	5 068	0.22	-0.02

Est-ce qu'on peut voir une cascade dans la plaine ?

Est-ce que plonger ça veut dire sauter ?

Est-ce qu'un gland est un fruit ?

Si on considère uniquement les items V2 (13 items), l'indice alpha vaut 0.75.

Si on considère uniquement les items V3 (13 items), l'indice alpha vaut 0.55 et 3 items ont un Rbis inférieur à 0.2.

NAME	N	MEAN	Rbis
xV3_6	5 068	0.55	0.13
xV3_11	5 068	0.31	0.08
xV3_14	5 068	0.22	-0.01

Lecture par voie lexicale

10 items, alpha = 0.63

1 item a un Rbis inférieur à 0.2.

NAME	N	MEAN	Rbis
xLVL4	5 061	0.20	0.06

nid

Lecture par voie non lexicale

10 items, alpha = 0.72

Habiletés phonologiques

12 items, alpha = 0.83

1 item a un Rbis inférieur à 0.2

NAME	N	MEAN	Rbis
xNBSONS1	5 045	0.85	0.15

 at**Connaissance du son des lettres**

15 items, alpha = 0.86

Mathématiques

10 items, alpha = 0.68

Synthèse.

Séquence	Nb d'items de départ	Nb d'items Conservés	Alpha	Facteur d'allongement ¹	Alpha à nb d'items équivalent ²
Écriture	15	15	0.84	1.53	0.89
Compréhension	12	12	0.71	1.92	0.82
Vocabulaire	26	23	0.79	1.00	0.79
Items V2	13	13	0.75	1.77	0.84
Items V3	13	10	0.61	2.30	0.78
Lecture voie lexicale	10	9	0.65	2.56	0.83
Lecture voie non lexicale	10	10	0.72	2.30	0.86
Phonologie	12	11	0.83	2.09	0.91
Sons des lettres	15	15	0.86	1.53	0.91
Mathématiques	10	10	0.68	2.30	0.83

¹ On reporte toutes les séquences à 23 items. Facteur d'allongement $n = 23/\text{nb d'items de la séquence}$

² Alpha à nb d'items équivalent = $\frac{n\alpha}{1+(n-1)\alpha}$

Temps 4

Compréhension orale

9 items, alpha = 0.67

Vocabulaire 1

La séquence vocabulaire 1 se compose de 4 items qui sont repris parmi les 13 items de vocabulaire du temps 2, mais avec une mise en forme différente comme au temps 3 (renommés V2) auxquels s'ajoute à chaque fois, comme au temps 3, un item consistant en une question portant sur le sens du mot considéré (items nommés V3)

8 items, alpha = 0.45

4 items ont un Rbis < 0.2, dont un < 0.15.

NAME	N	MEAN	Rbis
xV2_20	4 631	0.87	0.18
xV3_15	4 631	0.73	0.18
xV3_16	4 631	0.71	0.15
xV3_20	4 631	0.91	0.18

On supprime l'item xV3_16 dont le Rbis est inférieur à 0.15.

NAME	N	MEAN	Rbis
xV2_20	4 631	0.87	0.19
xV3_15	4 631	0.73	0.17
xV3_20	4 631	0.91	0.18

Vocabulaire 2

La séquence vocabulaire 2 est une nouvelle séquence.

30 items, alpha = 0.46

Seulement 5 items ont un Rbis >= 0.2

Production orthographique

12 items, alpha = 0.84

Lecture voie lexicale

16 items, alpha = 0.75

Un item a un Rbis inférieur à 0.2.

NAME	N	MEAN	Rbis
xLVL11	4 622	0.54	0.20

Lecture voie non lexicale

10 items, alpha = 0.69

Habilités phonologiques

12 items, alpha = 0.85

Un item a un Rbis inférieur à 0.2.

NAME	N	MEAN	Rbis
xNBSONS1	4 612	0.86	0.18

Compréhension en lecture

12 items, alpha = 0.73

Mathématiques

20 items, alpha = 0.83

Synthèse

séquence	Nb items origine	Nb items conservés	Alpha	Facteur d'allongement ¹	Alpha à nb d'items équivalent ²
Compréhension orale	9	9	0.67	3.33	0.87
Vocabulaire 1	8	7	0.45	4.29	0.78
Vocabulaire 2	30	30	0.46	1.00	0.46
Orthographe	12	12	0.84	2.50	0.93
Lecture voie lexicale	16	16	0.75	1.88	0.85
Lecture voie non lexicale	10	10	0.69	3.00	0.87
Habilités phonologiques	12	12	0.85	2.50	0.93
Compréhension en lecture	12	12	0.73	2.50	0.87
Mathématiques	20	20	0.83	1.50	0.88

¹ On rapporte toutes les séquences à 30 items. Facteur d'allongement $n = 30/\text{nb d'items de la séquence}$

² Alpha à nb d'items équivalent = $\frac{n\alpha}{1+(n-1)\alpha}$

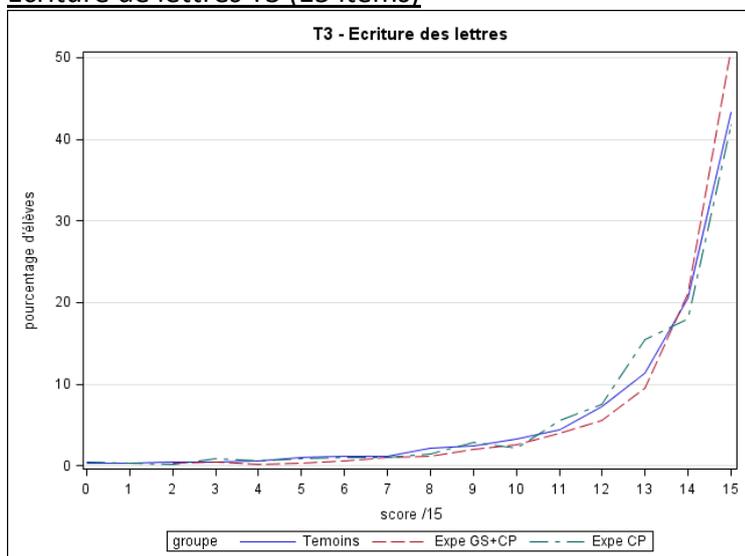
Pour la séquence de vocabulaire 2, étant donné le trop grand nombre d'items ayant un mauvais indice de discrimination, les items sont exclus de l'analyse, il n'y a pas de score calculé.



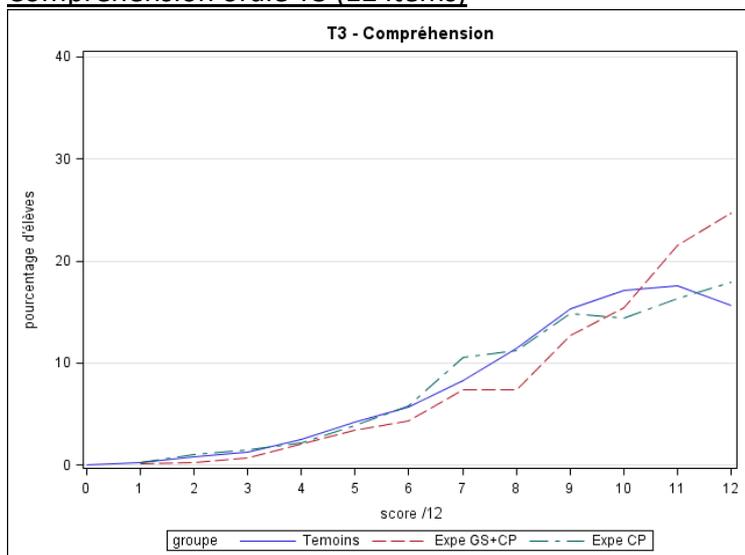
ANNEXE 7 : DISTRIBUTION DES SCORES T3T4

Scores bruts de l'ensemble des élèves ayant participé au temps 3 selon le groupe expé GS/CP/ expé CP / témoin

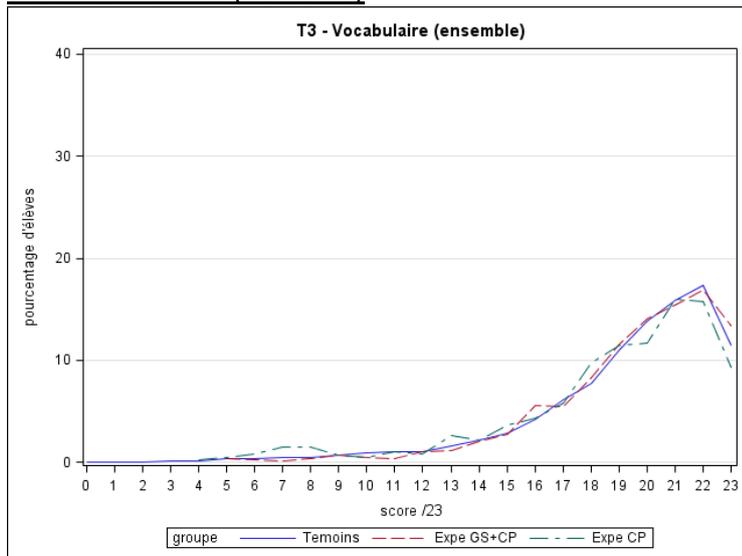
Écriture de lettres T3 (15 items)



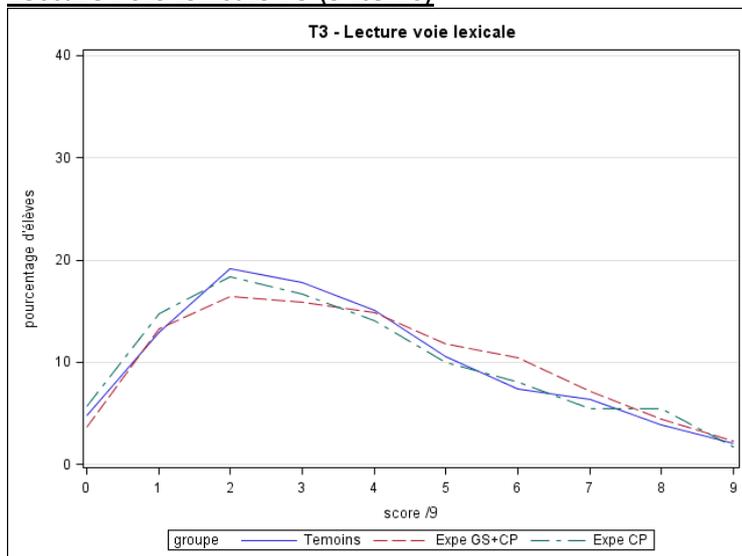
Compréhension orale T3 (12 items)



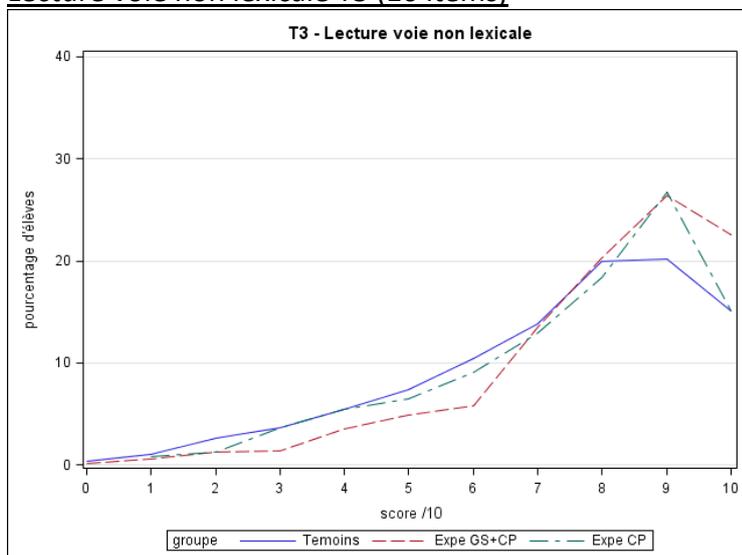
Vocabulaire 1 T3 (23 items)



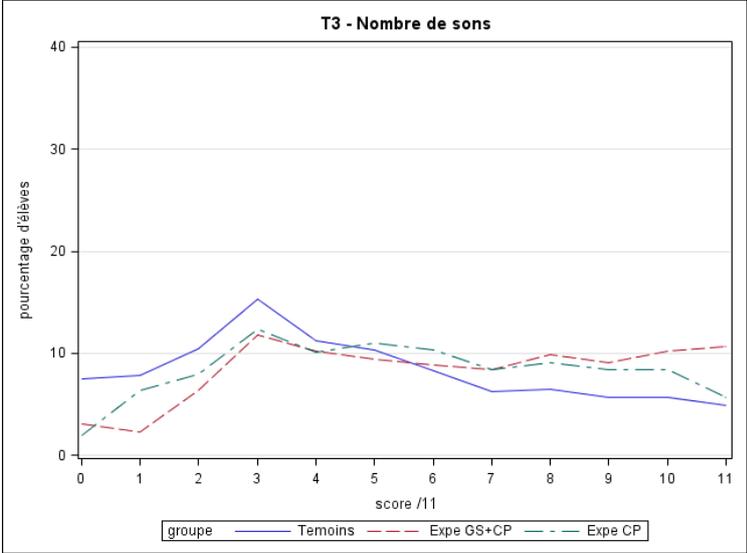
Lecture voie lexicale T3 (9 items)



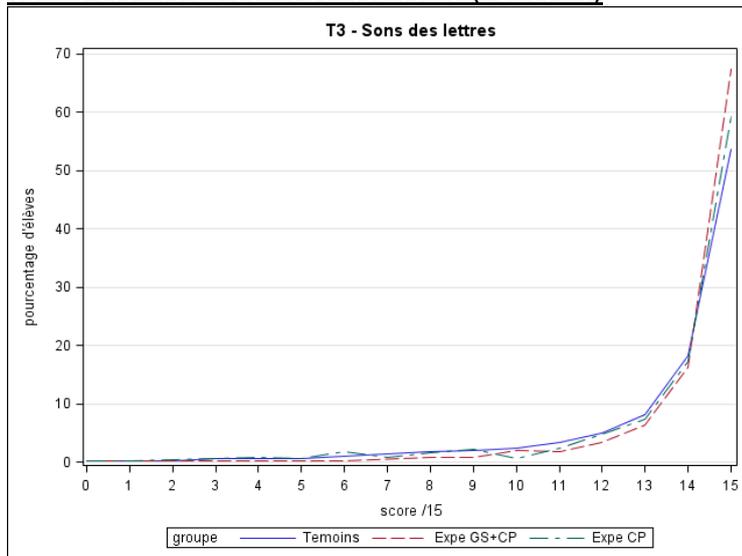
Lecture voie non lexicale T3 (10 items)



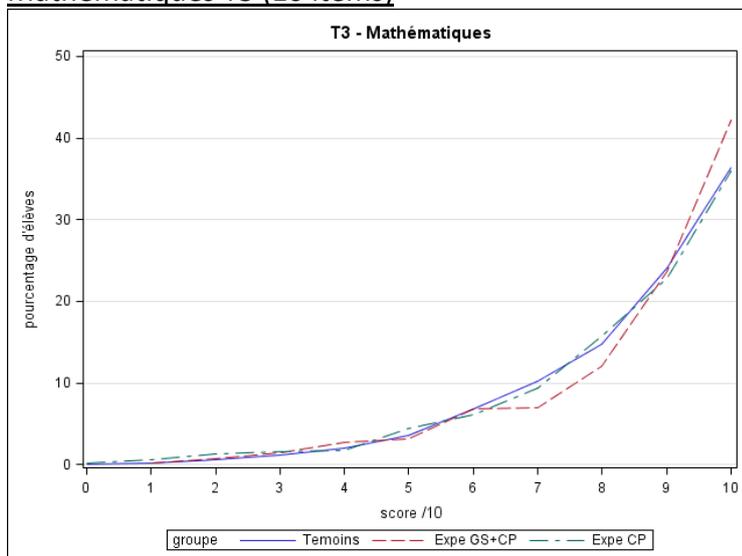
Habilités phonologiques T3 (11 items)



Connaissance du son des lettres T3 (15 items)

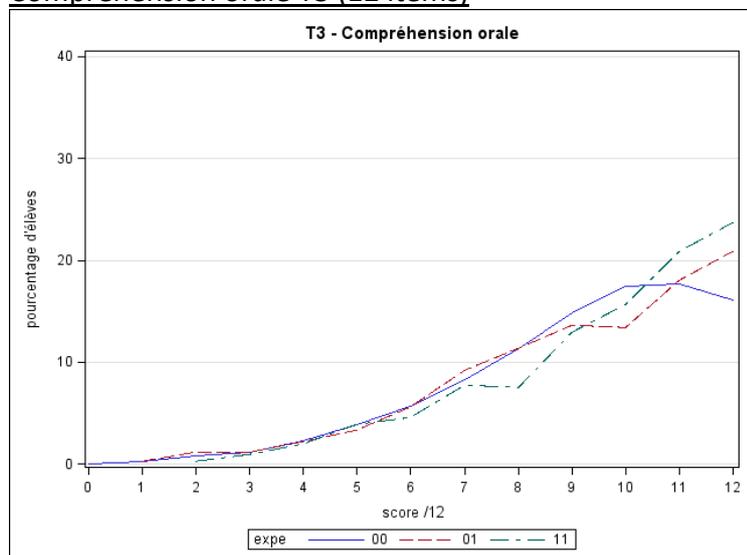


Mathématiques T3 (10 items)



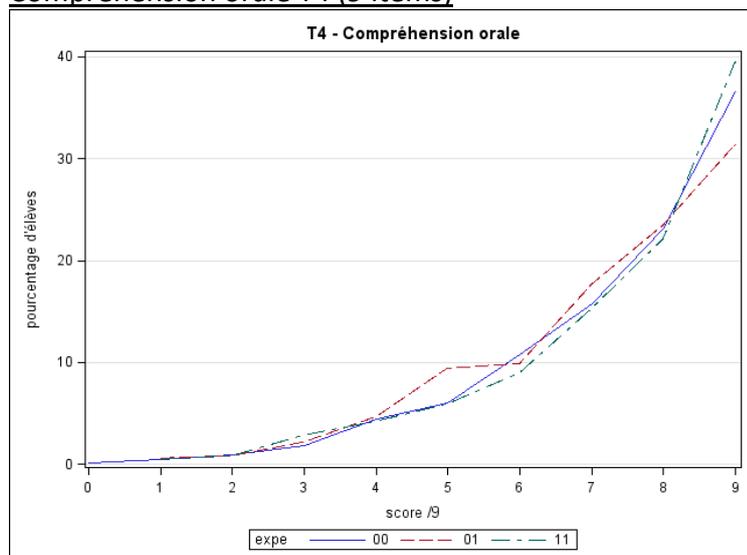
Scores bruts des élèves ayant participé aux temps 3 et 4 selon le groupe expé GS/CP/ expé CP / témoin

Compréhension orale T3 (12 items)

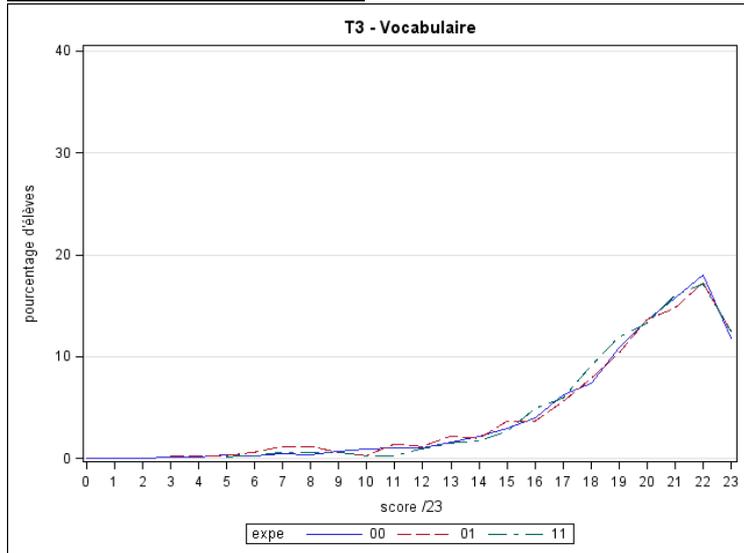


Expé GS/CP = 11 Expé CP= 01 Témoin = 00

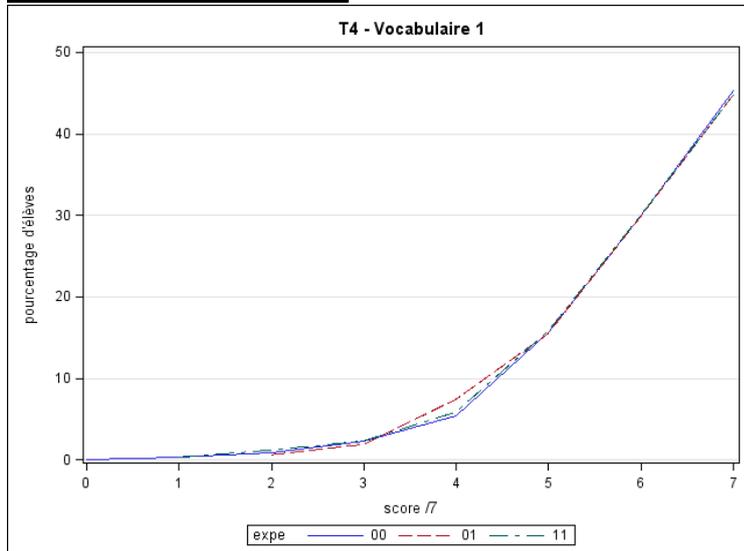
Compréhension orale T4 (9 items)



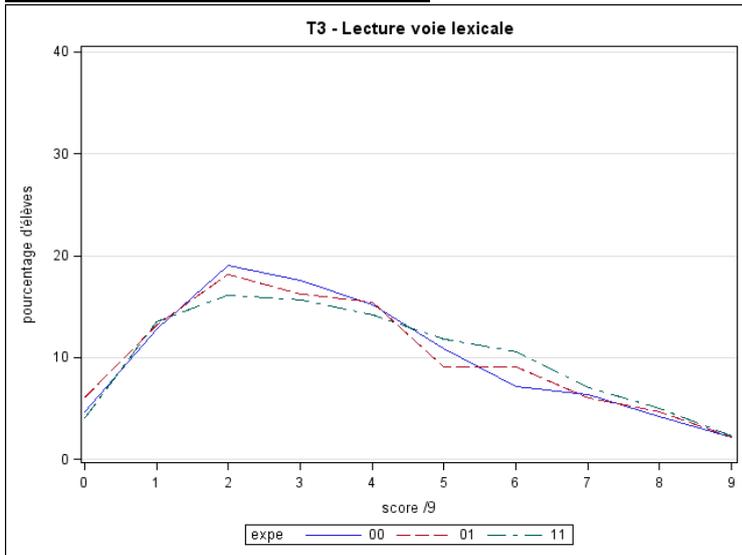
Vocabulaire 1 T3 (23 items)



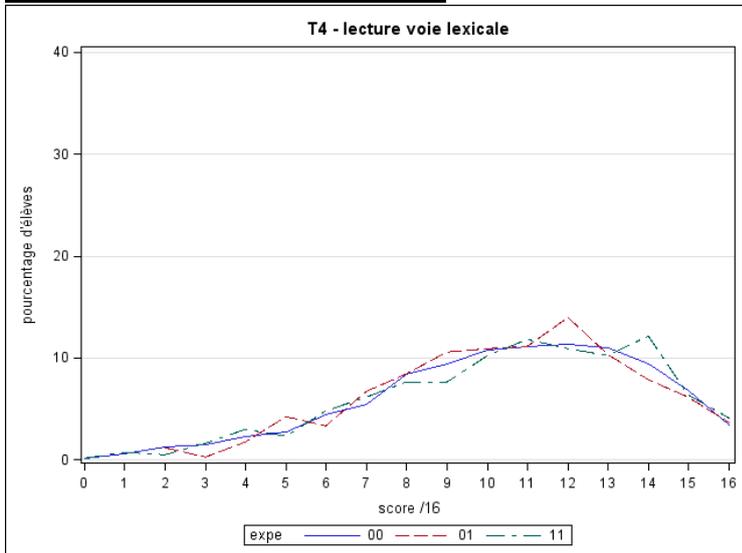
Vocabulaire 1 T4 (7 items)



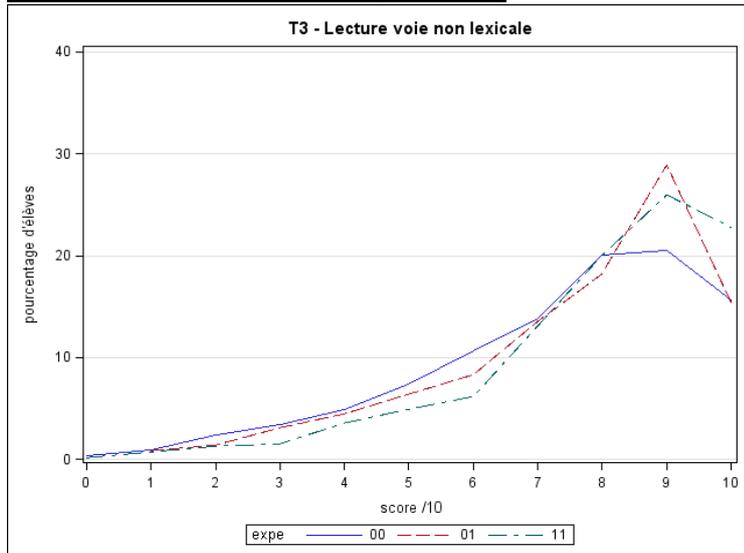
Lecture voix lexicale T3 (9 items)



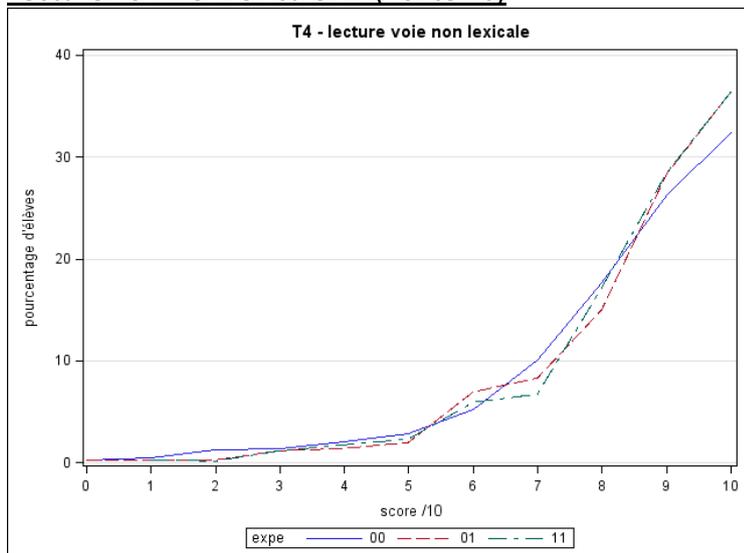
Lecture voix lexicale T4 (16 items)



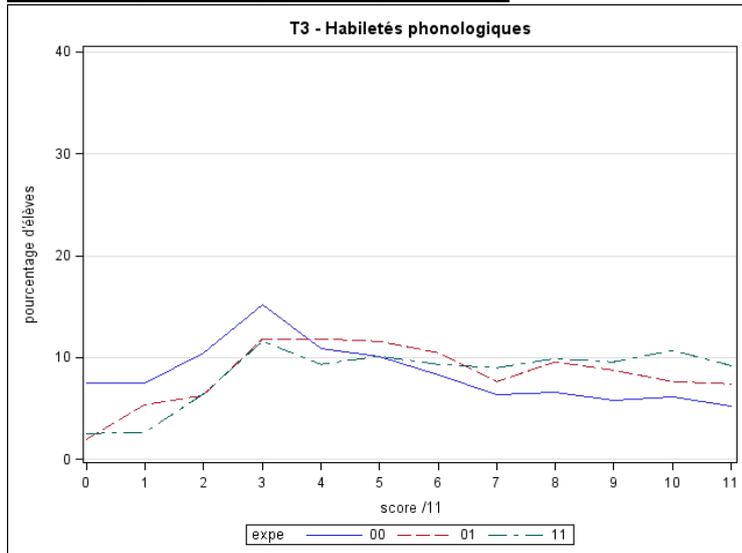
Lecture voix non lexicale T3 (10 items)



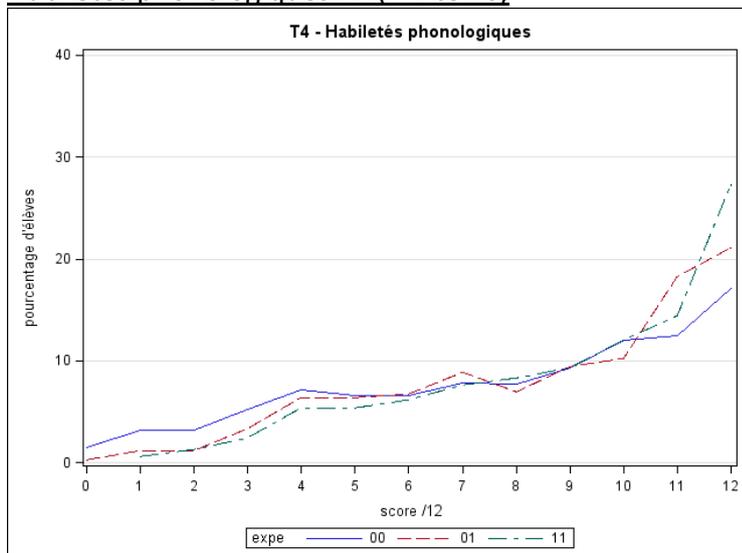
Lecture voix non lexicale T4 (10 items)



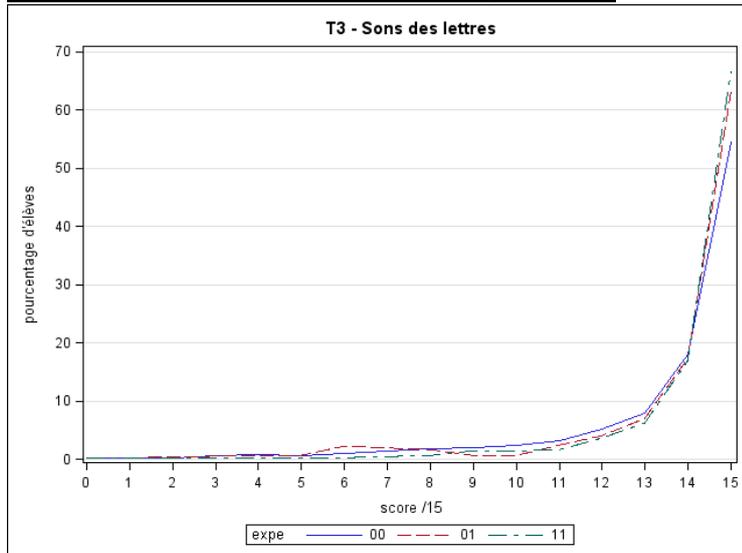
Habilités phonologiques T3 (11 items)



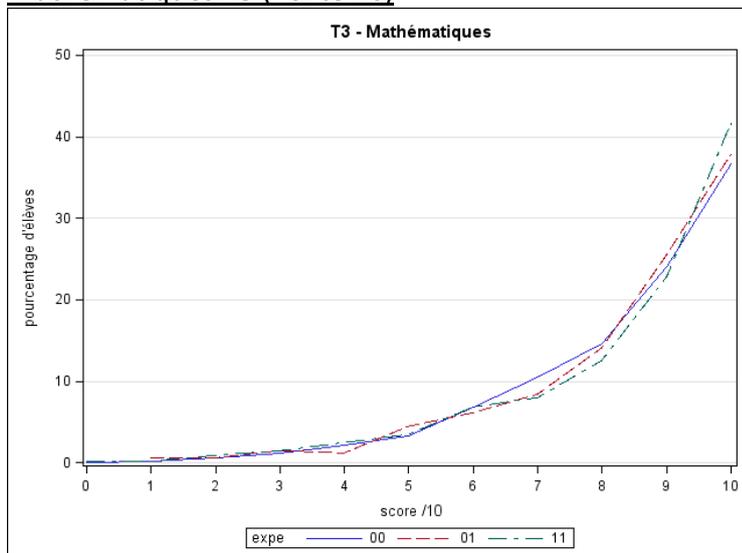
Habilités phonologiques T4 (12 items)



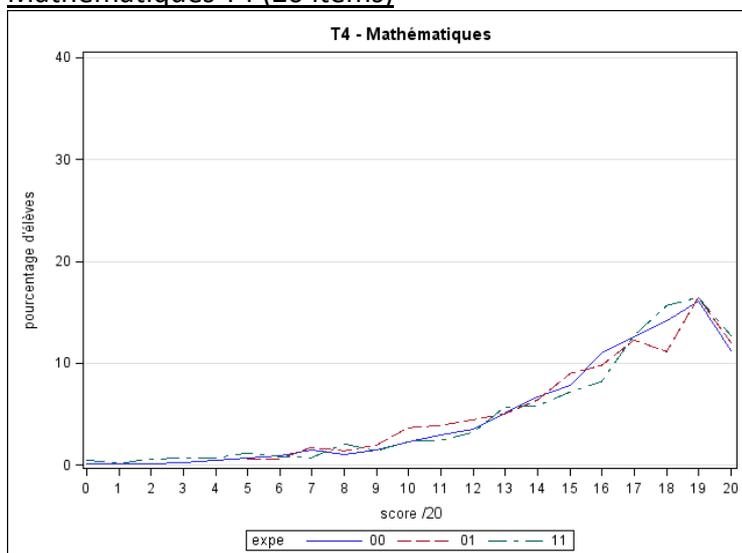
Connaissances du son des lettres T3 (15 items)



Mathématiques T3 (10 items)



Mathématiques T4 (20 items)



Fonds d'Expérimentation pour la Jeunesse

Ministère des Sports, de la Jeunesse, de l'Éducation Populaire et de la Vie Associative
Direction de la Jeunesse, de l'Éducation Populaire et de la Vie Associative

Mission d'animation du Fonds d'Expérimentation pour la Jeunesse

95, avenue de France 75 650 Paris Cedex 13

Téléphone : 01 40 45 93 22

www.experimentation.jeunes.gouv.fr

