



PROGRAMME D'ACCOMPAGNEMENT 1001MOTS

ÉVALUATION D'IMPACT

Eval-Lab
www.eval-lab.com

Rendre l'excellence scientifique accessible aux acteurs du changement social

Table des matières

Table des matières

Synthèse	4
Un mot sur Eval-Lab	6
1. Description de l'intervention	7
a. Modèle théorique.....	7
b. Contenu du programme 1001mots	8
c. Inscription des familles	10
2. État des lieux : portrait des familles et des enfants bénéficiaires	10
a. Profil socio-démographique des familles bénéficiaires	11
b. Croyances et pratiques parentales	13
i) Mesurer les croyances et pratiques parentales	13
ii) Portrait des croyances et des pratiques parentales	14
c. Développement de l'enfant.....	15
i) Mesurer le niveau de développement des enfants	15
ii) Portrait du niveau de développement des enfants	16
d. Déterminants du développement de l'enfant	18
i) Filles et garçons : des pratiques déclarées similaires mais un niveau de développement différent	19
ii) Ce qui prédit le niveau de développement des enfants	19
3. Comment évaluer l'impact du programme 1001mots ?	21
a. Impact et évaluation d'impact.....	21
i) Qu'est-ce que l'impact d'un programme ?	21
ii) Observer une évolution ne suffit pas à démontrer un impact	22
iii) Un groupe de comparaison pour approcher ce qui se serait passé sans 1001mots	22
b. Un groupe de comparaison composé des futurs bénéficiaires	24
d. Les participants à l'étude	27
e. Comparabilité des groupes	27
f. Méthodes statistiques pour estimer l'impact du programme	29
i) <i>Méthode par ajustement statistique (régression multivariée)</i>	29
ii) <i>Méthode par appariement statistique (matching)</i>	30
iii) <i>Méthode par repondération statistique (entropy balancing)</i>	31
iv) <i>Robustesse : sensibilité aux facteurs inobservables</i>	32
4. L'impact du programme 1001mots	33
a. Impact sur les croyances et pratiques parentales	33

b. Impact sur le développement des enfants	35
c. Hétérogénéité des effets	39
5. Ce que disent les parents bénéficiaires : analyse des retours qualitatifs	41
6. Limites de l'évaluation	43
a. On ne compare pas des familles tirées au sort	43
b. Les informations sont déclarées par les parents	44
c. Un effet de cohorte lié à la temporalité du recrutement.....	44
Conclusion.....	45
Annexe A – Tableaux d'hétérogénéité des effets	47
Annexe B - Impact du programme sur chaque item de l'IDE.....	52

Synthèse

1001mots est une association qui lutte contre les inégalités en proposant un an d'accompagnement parental à distance aux familles ayant un petit budget, afin de favoriser l'éveil langagier de leur jeune enfant. Le programme repose sur trois piliers : l'envoi de trois SMS par semaine, l'envoi de six livres et quatre appels d'accompagnement parental. Les premières évaluations randomisées du programme avaient montré un effet positif sur les pratiques parentales, mais pas sur le développement des enfants. La présente évaluation mesure l'impact de la nouvelle version du programme sur les pratiques parentales et le développement des enfants.

Un public fragile, aux besoins importants

Parmi les 802 familles cibles interrogées avant leur accompagnement, 83 % vivent sous le seuil de pauvreté. Leurs enfants, âgés de 24 mois en moyenne, ont un âge de développement langagier inférieur de trois à quatre mois à leur âge réel : 20 % présentent un très haut risque de retard de langage en expression et 25 % en compréhension. Seules 30 % des familles lisent quotidiennement avec leur enfant, et 39 % des enfants regardent la télévision tous les jours.

Une méthodologie rigoureuse pour isoler l'effet propre du programme

L'évaluation compare deux groupes : 224 familles engagées dans le programme, interrogées à la fin de leur accompagnement, et 340 familles interrogées avant de commencer le programme. Ces dernières ont ensuite bénéficié de l'accompagnement et ont été retenues pour la comparaison parce qu'elles ont atteint un niveau d'engagement minimal, similaire à celui des familles du premier groupe.

Le développement des enfants a été mesuré à l'aide de l'Inventaire du Développement de l'Enfant (IDE), un outil standardisé et validé. Pour tenir compte des différences de profil entre les deux groupes, quatre méthodes statistiques ont été utilisées, en prenant en compte les caractéristiques socio-démographiques mesurées. Ces méthodes aboutissent à des résultats convergents. Un test de robustesse renforce également la confiance dans les résultats en examinant leur sensibilité à d'éventuelles différences non mesurées entre les groupes.

Des pratiques parentales nettement modifiées

- **Plus de lecture** : les parents bénéficiaires lisent davantage avec leur enfant (4,6 fois par semaine contre 3,4). La part de ceux qui lisent tous les jours passe de 26 % à 47 % et celle des parents qui ne lisent jamais de 12 % à 5 %. L'âge auquel les parents jugent utile de commencer à regarder des livres avec un enfant recule de 9,4 à 6,1 mois.
- **Moins d'écrans** : le temps d'exposition aux écrans est divisé par 2 (de 4h30 à 2h17 par semaine) chez les familles bénéficiaires. La part de familles dont la télévision reste souvent allumée en fond passe de 41 % à 25 %.

- **Plus de conversations** : la part de parents qui décrivent tout ce qu'ils font avec l'enfant au quotidien passe de 39 % à 55 % entre les deux groupes.

Un effet marqué sur le développement des enfants

- **Un meilleur niveau de langage** : l'âge de développement langagier des enfants bénéficiaires est supérieur de 2,3 mois en expression et de 1,8 mois en compréhension. Le vocabulaire et la syntaxe s'améliorent avec des effets de +0,44 et +0,50 écart-type respectivement.
- **Moins d'enfants à risque de retard de langage** : la part d'enfants à très haut risque de retard de langage est divisée par deux (de 21 % à 9 % en expression, de 26 % à 13 % en compréhension) chez les enfants bénéficiaires.
- **De meilleures compétences socio-émotionnelles** : les compétences socio-émotionnelles des enfants bénéficiaires augmentent de +0,42 écart-type. Cet effet est notamment porté par une amélioration de la concentration et de l'expression des émotions.

Des effets homogènes entre familles, plus forts chez les enfants plus âgés

Les analyses par sous-groupes montrent que les différentes caractéristiques socio-économiques ne modifient quasiment pas l'ampleur des effets : le programme est aussi efficace auprès des familles les plus défavorisées qu'auprès des autres. Les effets sont en revanche plus marqués pour les enfants plus âgés (plus de 24 mois) et l'adoption de la lecture partagée est surtout portée par les parents nés à l'étranger.

En conclusion

La nouvelle version du programme 1001mots a un fort impact à la fois sur les pratiques des parents accompagnés, mais aussi sur le développement langagier et socio-émotionnel de leurs enfants. La convergence des quatre méthodes statistiques et du test de robustesse atteste de la solidité de ces résultats. L'accompagnement 1001mots permet effectivement aux enfants accompagnés d'entrer à l'école maternelle avec de bien meilleures chances de réussite.

Un mot sur Eval-Lab

Ce rapport a été réalisé par [Eval-Lab](#), structure spécialisée dans l'évaluation d'impact social, la recherche et le conseil, et qui se distingue en France par la rigueur de ses méthodes et son éthique. Notre mission est de rendre l'excellence scientifique accessible aux acteurs du changement social.

La rédaction de ce rapport a été réalisée par Quentin Daviot, fondateur et directeur général d'Eval-Lab. Titulaire d'un doctorat en économie, Quentin est spécialiste de l'évaluation d'impact des programmes éducatifs. Il a travaillé près de dix ans au laboratoire J-PAL, créé par la lauréate du prix Nobel Esther Duflo, et huit ans à la Banque Mondiale, où il s'est consacré à l'évaluation de programmes éducatifs. Quentin a participé à de nombreux projets en éducation, en France et à l'international. Ses travaux de recherche portent sur la petite enfance, les apprentissages, la formation des enseignants, le développement des compétences socio-émotionnelles, ainsi que sur les trajectoires scolaires et professionnelles. Quentin enseigne l'évaluation d'impact à Sciences Po Paris depuis 2021.

1. Description de l'intervention

1001mots est une association qui a pour mission de lutter contre les inégalités dès le plus jeune âge, en proposant un accompagnement parental à distance aux familles ayant un petit budget. La littérature scientifique a mis en évidence qu'en fonction du milieu social d'origine, les enfants ne présentent pas le même niveau de développement du langage au moment de l'entrée à l'école, ce qui crée un écart dès le début de leur scolarité, qui va croissant avec le temps. L'objectif de 1001mots est d'accompagner les parents dans l'éveil langagier de leur enfant afin de faire grandir son langage et de lui donner toutes ses chances de réussite avant d'entrer à l'école maternelle.

a. Modèle théorique

Le programme 1001mots repose sur un modèle théorique issu de la synthèse de différents modèles utilisés en psychologie dans le champ du changement de comportement ou "*Behaviour Change*". Ces théories sont largement utilisées pour répondre à des problématiques de santé publique telles que le fait de trop fumer ou de ne pas manger sainement. Ce modèle a été adapté à la problématique cœur de l'action de 1001mots : amener les parents à modifier leurs pratiques parentales afin qu'elles deviennent plus stimulantes pour le développement de l'enfant, et particulièrement de son langage. Pour amener le parent à changer son comportement, il est nécessaire qu'il soit suffisamment motivé à le faire. Pour cela, le programme travaille sur trois différents types de croyances que le parent peut avoir vis-à-vis de ce comportement et qui influencent directement son intention de le changer.

- L'**attitude envers le comportement** regroupe toutes les croyances sur les bénéfices à plus ou moins long terme du comportement : plus le parent aura de connaissances sur les bénéfices du comportement, plus son intention de l'adopter sera élevée.
- La **norme sociale** regroupe toutes les croyances sur ce que l'entourage du parent pense du comportement. Plus les personnes dont l'opinion compte beaucoup pour le parent (famille, amis, professionnels, autres parents...) sont favorables au comportement en question, plus l'intention de le mettre en place sera élevée.
- Le **sentiment de capacité personnelle** regroupe toutes les croyances concernant la capacité du parent à mettre en place facilement le comportement ciblé. Plus le parent sait comment mettre concrètement en place le comportement et se sent capable de le faire malgré les difficultés potentielles, plus il aura l'intention de le mettre en pratique.

Concrètement, nous avons sélectionné plusieurs techniques comportementales issues de la pratique en psychologie pour modifier ces différents types de croyances

afin d'accompagner les parents vers l'adoption des comportements ciblés. Ces techniques sont utilisées dans tous les contenus développés dans notre accompagnement. Chaque élément présent pour chaque thématique a une motivation théorique.

b. Contenu du programme 1001mots

Le programme 1001mots est un programme d'accompagnement parental à distance qui dure un an. Il est composé d'une succession de sept modules thématiques durant chacun cinq semaines, qui sont séparés par deux semaines de pause. Un module comprend une séquence de SMS et un livre sur un thème défini, pour une tranche d'âge précise (tranche de six mois) et pour un niveau de pratiques parentales donné (débutant, intermédiaire, avancé). Le thème des deux premiers modules est imposé par 1001mots (parler avec mon bébé et intéresser mon enfant aux livres). Le thème des cinq modules suivants est choisi par la famille ou, en l'absence de choix, attribué automatiquement. Parmi les thématiques proposées figurent notamment : chanter avec mon bébé, des idées pour jouer avec mon bébé, découvrir le monde avec mon enfant pendant les sorties, parler plusieurs langues à la maison, occuper mon enfant sans les écrans et comprendre ou encore gérer sa colère. L'ensemble des messages est strictement adapté à l'âge de l'enfant. Ainsi, les parents d'un enfant de six mois reçoivent des contenus différents de ceux d'un enfant de 13 mois.

Le programme repose sur trois composantes principales :

L'envoi de contenu personnalisé par SMS : le parent reçoit trois SMS par semaine sur le thème du module en cours. Ces messages comprennent des conseils pratiques, des témoignages de parents, des idées d'activités ou encore des informations un peu plus théoriques. Ces contenus peuvent prendre différents formats : texte, image ou vidéo (via un lien intégré dans le message renvoyant vers une vidéo hébergée sur une [application web](#)). Selon les préférences des familles, un ou deux parents peuvent recevoir les messages, ce choix leur appartient.

L'envoi de six livres par voie postale (à raison d'un livre tous les deux mois). À partir du module sur le thème de la lecture (deuxième module reçu), les parents reçoivent un livre à domicile pour chaque nouveau module. Chaque livre est sélectionné en fonction de l'âge de l'enfant et de la thématique du module choisie par la famille.

Quatre appels téléphoniques d'accompagnement parental sont prévus au cours du programme. Ils sont réalisés par des spécialistes de la petite enfance (orthophoniste, psychologue ou éducatrice de jeunes enfants). Chaque famille est suivie par une seule personne de référence au cours de son accompagnement, appelée "accompagnante". Ces appels visent à accompagner les parents dans l'adoption de pratiques ayant un fort impact sur le développement langagier, tel que le

langage adressé à l'enfant, la lecture interactive, ou encore le fait de parler des actions que l'enfant est en train de faire et de suivre son intérêt, tout en adoptant une posture bienveillante et la plus horizontale possible.

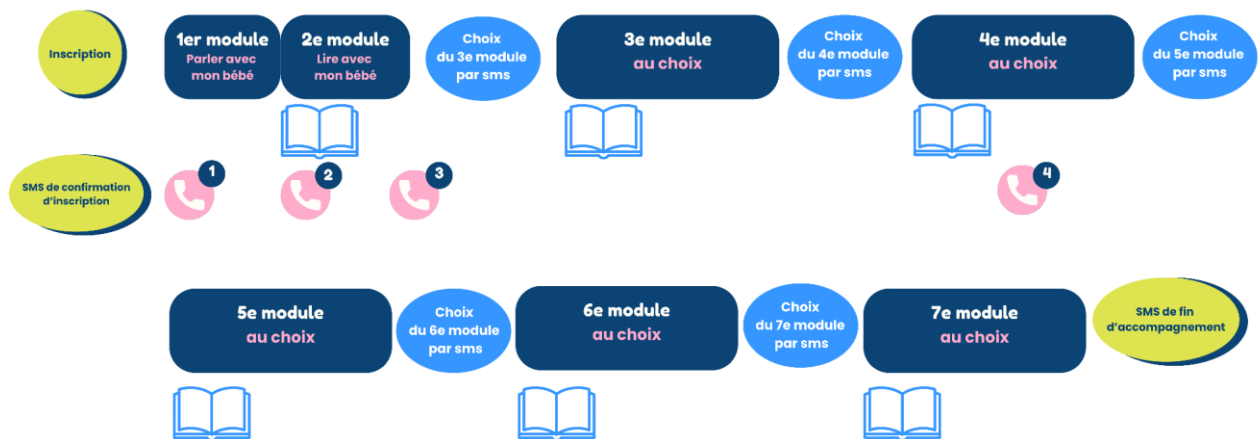
Le premier appel a lieu en tout début d'accompagnement, au début du module sur les conversations du quotidien, le second appel a lieu environ un mois plus tard, au début du module sur la lecture et de la réception du premier livre. Le troisième appel a lieu à la fin de ce module lecture et permet notamment d'expliquer au parent qu'il pourra choisir les thèmes suivants. Enfin le quatrième et dernier appel a lieu plus tard dans l'accompagnement, pendant le quatrième module, avant que le parent poursuive l'accompagnement plus en autonomie.

Pour chaque appel, 1001mots a développé un script afin de guider les accompagnantes dans le déroulé de l'appel ainsi que de s'assurer d'une homogénéité dans la qualité des appels pour toutes les familles. Tous les scripts d'appel sont fondés sur la technique de l'entretien motivationnel, souvent utilisé dans le champ de la psychologie. L'objectif de cette technique est d'accompagner le parent vers un changement de pratique, lorsqu'il s'y sent prêt. Les accompagnantes sont formées à repérer dans son discours les signes d'une volonté de changement ou, au contraire, les réticences à changer. Elles peuvent ainsi adapter leurs échanges au niveau de motivation de chaque famille. Dans la plupart des appels, l'accompagnante propose au parent de visionner une courte vidéo 1001mots en lien avec le thème abordé. Cette vidéo sert de point de départ à l'échange : le parent peut réagir à ce qu'il a découvert et parler de ce qu'il fait déjà avec son enfant. À partir de cet échange, l'accompagnante l'aide, s'il s'y sent prêt, à identifier une nouvelle chose simple à mettre en place au quotidien avec son enfant : la « petite mission ». L'objectif est que cette petite mission vienne autant que possible du parent et soit co-construite avec l'accompagnante, plutôt que de lui être proposée toute faite. Elle correspond ainsi davantage à ses envies et à sa situation, ce qui favorise sa mise en pratique.

Les parents accompagnés ont aussi la possibilité de rejoindre le [compte Instagram de 1001mots](#) animé par deux orthophonistes et proposant du contenu complémentaire à l'accompagnement plusieurs fois par semaine. On retrouve notamment des conseils et astuces pratiques, des idées d'activité, des comptines, des témoignages de parents ou encore des bons plans.

Enfin, sur l'[application web](#) qui héberge les vidéos envoyées par SMS, le parent peut avoir accès à un réservoir d'activités sur de nombreuses thématiques, et adaptées à différentes tranches d'âge, que les parents peuvent consulter à tout moment.

Figure 1 : déroulement de l'accompagnement 1001mots



c. Inscription des familles

La plupart des familles s'inscrivent au programme par l'intermédiaire des partenaires de 1001mots : les CAF et les PMI. Les modalités diffèrent pour chacun d'entre eux.

- Les CAF envoient un mail ou un SMS aux familles ayant un enfant dans la bonne tranche d'âge et dont les ressources sont inférieures à un seuil défini. Les familles font ensuite elles-mêmes la démarche de s'inscrire en remplissant un formulaire en ligne joint dans le message reçu.
- Les PMI proposent l'accompagnement aux familles qu'elles repèrent comme étant dans la cible (faibles ressources, faibles pratiques parentales, etc.) lors d'un rendez-vous de suivi. Si elles acceptent, ce sont les professionnels de PMI qui inscrivent la famille à l'aide d'un formulaire qui leur est réservé.

2. État des lieux : portrait des familles et des enfants bénéficiaires

Pour caractériser au mieux le profil des familles et des enfants bénéficiaires, nous avons développé avec 1001mots un questionnaire qui permet de mesurer trois catégories de données :

- les caractéristiques socio-démographiques des familles ;
- les croyances et les pratiques parentales en lien avec le développement de l'enfant et ;

- le développement langagier (en termes d'expression et de compréhension) et le développement socio-émotionnel des enfants.

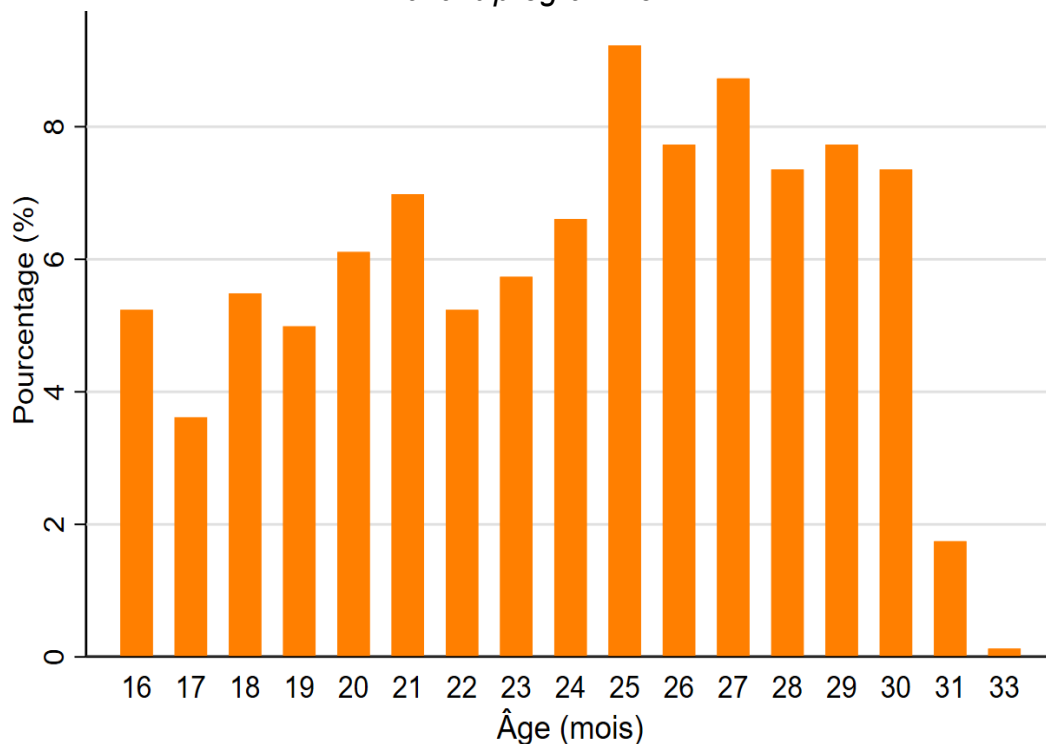
Ces données ont été collectées auprès de 1 026 répondants : 802 y ont répondu avant de bénéficier du programme et 224 y ont répondu après leur participation au programme.

a. Profil socio-démographique des familles bénéficiaires

Cette section dresse le portrait des 802 familles qui ont répondu à notre questionnaire avant de démarrer leur accompagnement, ce qui donne une image fidèle des ménages que le programme 1001mots cherche à accompagner. L'ensemble des caractéristiques sont détaillées dans le tableau 1.

Les enfants. Au moment de la passation du questionnaire, les enfants ont en moyenne 23,9 mois, dans une fourchette allant de 16 à 33 mois (voir figure 2). La répartition entre garçons (52,7 %) et filles (47,3 %) est équilibrée.

Figure 2 : distribution des âges des enfants en mois au moment du questionnaire avant programme

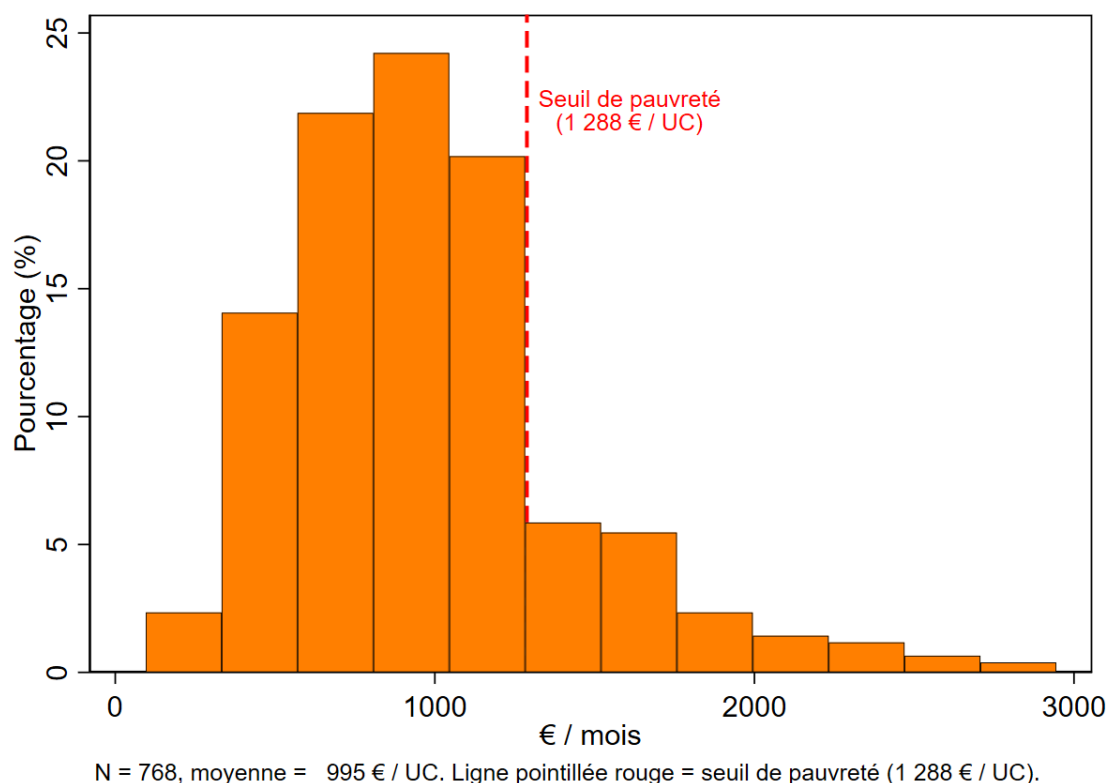


N = 802, moyenne = 23.9 mois

Un contexte familial et économique marqué par la précarité. Les ménages comptent en moyenne 4,1 personnes et 2,4 enfants ; 63 % sont biparentaux. Surtout,

ces familles vivent dans une grande précarité économique : le revenu médian par unité de consommation est d'environ 917 € par mois, et 83 % des familles se situent sous le seuil de pauvreté (fixé à 1 288 € par unité de consommation en 2023, soit 60% du niveau de vie médian¹) (voir figure 3). L'accès à l'emploi est limité : 53 % des pères et 28 % des mères occupent un emploi. La garde de l'enfant est assurée majoritairement par les parents eux-mêmes (70 %), le recours à la crèche restant minoritaire (26 %).

Figure 3 : distribution du revenu par unité de consommation et seuil de pauvreté des parents avant programme



Un niveau d'éducation contrasté. 38 % des parents interrogés ont un diplôme supérieur au baccalauréat, 33 % un niveau baccalauréat, et 29 % un niveau inférieur.

Une population largement multilingue et souvent d'origine étrangère. Moins de la moitié des répondants (44 %) sont nés en France, et les foyers parlent environ 2 langues en moyenne. Le français est la langue la plus parlée au domicile pour 73 % des familles, et 83 % des répondants déclarent le maîtriser couramment.

¹ INSEE, 2025

Tableau 1 : caractéristiques socio-démographiques des familles répondantes au questionnaire avant programme

	Moyenne / %	Écart-type	Min	Max	N
<i>Panel A : Enfant</i>					
Âge de l'enfant à la passation (mois)	23,9	4,3	16	33	802
<i>Panel B : Ménage</i>					
Taille du ménage (nombre de personnes)	4,1	1,5	2	10	796
Foyer biparental (%)	62,6	48,4	0	1	802
Répondant né en France (%)	43,6	49,6	0	1	802
<i>Panel C : Revenu</i>					
Revenu par unité de consommation (/mois)	995	450	95	2944	768
Sous le seuil de pauvreté (%)	82,7	37,9	0	1	768
<i>Panel D : Mode de garde</i>					
Parent(s) (%)	69,8	45,9	0	1	802
Crèche (%)	26,1	43,9	0	1	802
<i>Panel E : Emploi</i>					
Mère en emploi (%)	28,3	45,1	0	1	802
Père en emploi (%)	53,4	49,9	0	1	802
<i>Panel F : Langues parlées à la maison</i>					
Nombre de langues	1,9	0,7	1	4	802
Français parlé le plus souvent (%)	72,8	44,5	0	1	802
Niveau de français : courant (%)	82,8	37,8	0	1	715
Niveau de français : intermédiaire (%)	15,0	35,7	0	1	715
Niveau de français : débutant (%)	2,2	14,8	0	1	715
<i>Panel G : Niveau d'éducation du répondant</i>					
Inférieur au baccalauréat (%)	28,8	45,3	0	1	800
Baccalauréat (%)	33,2	47,1	0	1	800
Supérieur au baccalauréat (%)	38,0	48,6	0	1	800

Notes : 802 familles ayant répondu au questionnaire avant le démarrage de leur accompagnement. Le revenu par unité de consommation (UC) suit l'échelle d'équivalence INSEE modifiée (1 UC pour le répondant, 0,5 pour un conjoint, 0,3 par enfant de moins de 14 ans, 0,5 au-delà). Seuil de pauvreté fixé à 1 288 € par UC et par mois (60 % du revenu médian français, 2023).

En résumé, 1001mots s'adresse à une population jeune, multilingue et économiquement très fragile : plus de huit familles sur dix vivent sous le seuil de pauvreté. C'est dans ce contexte, où l'éveil langagier de l'enfant peut être fragilisé par la situation socio-économique des familles, que s'inscrit l'accompagnement 1001mots.

b. Croyances et pratiques parentales

i) Mesurer les croyances et pratiques parentales

L'objectif du programme 1001mots est d'accompagner progressivement les parents bénéficiaires dans le changement de leurs pratiques parentales, afin qu'ils adoptent des comportements favorables au développement de l'enfant.

Il est donc indispensable d'évaluer dans quelle mesure l'intervention conduit effectivement à un changement des pratiques parentales, cette première étape étant une condition nécessaire à l'observation d'un impact sur le développement des

enfants. En effet, l'intervention de 1001mots vise un impact sur l'enfant, mais elle passe entièrement par le parent qui en est le bénéficiaire direct. Pour évaluer ce premier niveau d'impact, nous avons mesuré :

- la perception du parent sur l'âge auquel il est jugé utile de commencer à parler et lire à son enfant ;
- la fréquence à laquelle les parents jouent, chantent, regardent un livre et décrivent ce qu'ils font avec leur enfant ;
- l'exposition aux écrans des enfants (télévision, tablette et/ou téléphone portable) ;
- la satisfaction des parents par rapport à l'intervention.

ii) Portrait des croyances et des pratiques parentales

Avant d'évaluer l'impact du programme, il est utile de décrire les croyances et les pratiques de l'ensemble des familles non encore accompagnées (n = 802). On constate chez ces familles un réel potentiel d'amélioration.

Croyances parentales. Les parents sont déjà très majoritairement convaincus de l'importance de parler tôt à leur enfant : 78 % estiment qu'il faut commencer dès la naissance. En revanche, ils situent bien plus tardivement l'intérêt de regarder des livres avec l'enfant : en moyenne à 10 mois. Enfin, interrogés sur le rôle des écrans pour l'acquisition du langage et la préparation à l'entrée à l'école, les parents sont partagés : 39 % considèrent que les écrans(applications, dessins animés sur la télévision ou le téléphone) « n'aident pas » l'enfant à apprendre à parler, 46 % qu'ils « aident un peu » et 13 % qu'ils « aident beaucoup », signe que les croyances sur le rôle des écrans sont hétérogènes.

Pratiques parentales. La lecture partagée entre le parent et l'enfant, l'un des leviers les plus efficaces pour l'éveil langagier, reste peu fréquente : les parents lisent en moyenne 3,6 fois par semaine avec leur enfant, mais cette moyenne masque de fortes disparités : seuls 30 % lisent quotidiennement, tandis que 13 % ne lisent jamais. Les autres types de pratiques stimulantes sont plus ancrées : les parents jouent avec leur enfant en moyenne 5,2 fois par semaine et chantent 4,4 fois par semaine. En revanche, le fait de décrire à voix haute ce que l'on est en train de faire avec l'enfant, une pratique simple et bénéfique pour le développement langagier, n'est adopté quotidiennement que par 42,8 % des parents.

Exposition aux écrans. L'exposition des enfants à la télévision est élevée : près de 4 heures par semaine en moyenne (environ 219 minutes), et 39 % des enfants y sont exposés au quotidien. L'usage du téléphone et de la tablette est en revanche beaucoup plus limité et plus rare à cet âge : plus d'une famille sur deux (56 %) n'y a jamais recours.

En résumé, ces données mettent en évidence trois axes de progression que le programme 1001mots adresse directement :

- un potentiel important sur la lecture, pratiquée quotidiennement par moins d'un tiers des familles et jamais pratiquée par plus d'une sur dix ;
- des croyances à faire évoluer sur l'âge auquel introduire les premiers livres, jugé trop tardif (10 mois en moyenne) ;
- une exposition à la télévision à réduire, avec près de quatre enfants sur dix exposés quotidiennement.

c. Développement de l'enfant

i) Mesurer le niveau de développement des enfants

Le développement de l'enfant est mesuré par l'Inventaire du Développement de l'Enfant (IDE), qui constitue l'adaptation française du *Child Development Inventory* (CDI). L'IDE est un questionnaire hétéro-rapporté par le parent qui permet de mesurer différentes dimensions du développement du jeune enfant et de l'enfant, de 15 mois à 6,5 ans. Ce questionnaire consiste à répondre par oui ou par non à un ensemble de questions sur ce que l'enfant est capable ou non de faire dans chacune des dimensions. Nous avons utilisé la version IDE-Langage afin d'avoir une mesure détaillée des compétences langagières des enfants. Cela nous a permis de mesurer les éléments suivants :

- Le niveau de langage en expression
- Le niveau de langage en compréhension
- Le développement socio-émotionnel, à l'aide de quelques items sélectionnés dans le questionnaire.

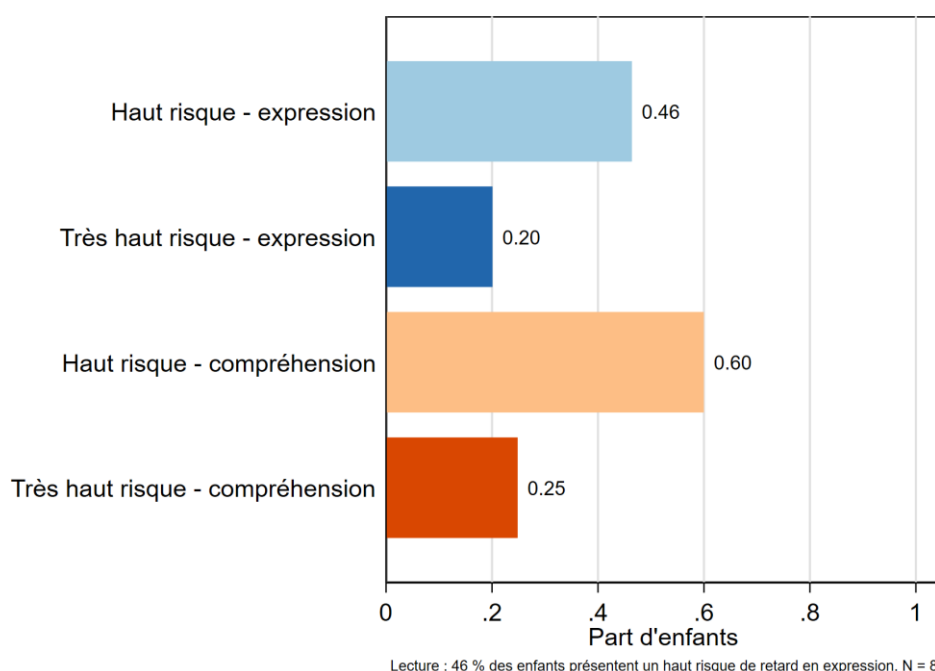
Le système de cotation de cet outil permet de calculer l'âge de développement langagier de l'enfant en expression ainsi qu'en compréhension, à partir des résultats obtenus. Grâce à un calcul prévu dans le manuel, il est ensuite possible de déterminer si l'enfant est à haut risque ou à très haut risque de retard de langage, en expression et en compréhension.

ii) Portrait du niveau de développement des enfants

De la même manière que pour les pratiques parentales, cette section décrit le niveau de développement des 802 enfants des familles non encore accompagnées. Les données révèlent une situation préoccupante : **une part importante des enfants présentent des risques élevés, voire très élevés, d'avoir un retard de langage, aussi bien en expression qu'en compréhension.**

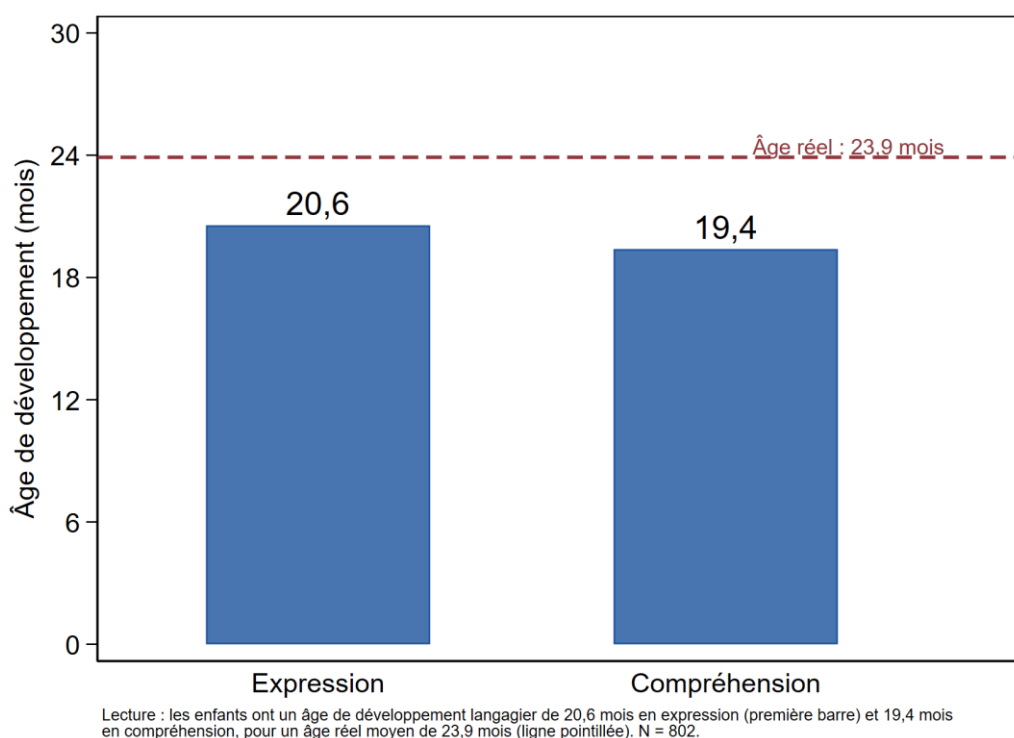
Sur le versant expressif du langage, 46 % des enfants présentent un haut risque de retard de langage, et 20 % un très haut risque. La compréhension est encore plus touchée : 60 % des enfants présentent un haut risque de retard de langage et 25 % un très haut risque (voir figure 4).

Figure 4 : part d'enfants à risque et très haut risque de retard de langage



Ces difficultés se retrouvent dans l'écart entre l'âge de développement langagier et l'âge réel des enfants : alors qu'ils ont en moyenne 24 mois, leur âge de développement langagier n'est en moyenne que de 21 mois en expression et de 19 mois en compréhension, soit un retard moyen de l'ordre de trois à cinq mois (voir figure 5).

Figure 5 : comparaison entre l'âge réel et l'âge de développement langagier des enfants



Ce constat atteste de l'ampleur des besoins auxquels le programme a pour objectif de répondre, dans une population où l'éveil langagier est fragilisé dès les premiers mois.

Le tableau 2 présente plusieurs items de compétences pour lesquels la proportion d'enfants ayant acquis la compétence, selon les déclarations de leurs parents, varie fortement.

Tableau 2 : exemple d'items issus de l'IDE et taux de réussite correspondant à ces items chez les enfants non encore accompagnés

Item de développement langagier	% maîtrise	N
Répond à son nom	99 %	800
Dit « maman »/« papa » (ou équivalent)	98 %	802
Comprend des demandes simples	95 %	801
Dit ≥ 2 autres mots (hors maman/papa)	94 %	794
Comprend « mettre » vs « enlever »	90 %	799
Comprend « ouvre/ferme » et fait le geste	82 %	795
Demande boire/manger avec mots ou sons	79 %	801
Dit ≥ 10 mots	72 %	802
Répond correctement par oui/non	60 %	799
Dit ≥ 5 noms d'objets usuels	54 %	801
Fait des phrases de 2-3 mots ou plus	47 %	801
Dit au moins son prénom	36 %	798
Parole claire (compréhensible la plupart du temps)	33 %	663
Utilise des phrases négatives	28 %	802
Identifie correctement au moins une couleur	27 %	662
Dit correctement son âge	25 %	383
Fait des phrases avec ≥ 4 mots	19 %	801
Récite des comptines (ex. 1,2,3, ...)	17 %	386
Fait des phrases de ≥ 6 mots	14 %	386
Montre droite/gauche correctement	9 %	202
Fait des phrases longues/complexes (≥ 10 mots)	7 %	203
Parle du futur (au-delà de la journée)	5 %	203

Lecture : le premier item (« Répond à son nom ») est maîtrisé par 99 % des 800 enfants dont le parent a répondu à cet item. *Note* : part d'enfants maîtrisant chaque item (réponse « oui »). Ensemble des 802 familles dont les données ont été collectées avant le démarrage du programme. Le questionnaire IDE est adaptatif et de difficulté croissante : un enfant qui n'atteint pas un item le laisse **manquant** - jamais compté comme un échec. La part est donc calculée parmi les seuls **répondants** (colonne N), dont le nombre diminue à mesure que les items se complexifient ; pour les items les plus avancés, elle s'interprète *conditionnellement* au fait d'avoir atteint ce niveau.

d. Déterminants du développement de l'enfant

Avant d'estimer l'impact du programme, les données collectées permettent d'éclairer une question plus large : qu'est-ce qui, indépendamment de 1001mots, prédit le niveau de développement d'un enfant ? Pour y répondre, nous étudions l'ensemble des 802 familles qui ont répondu à notre questionnaire avant de commencer leur accompagnement. Ces analyses sont corrélationnelles : elles décrivent des associations, et non des effets causaux, mais elles sont utiles pour comprendre le contexte dans lequel le programme intervient.

i) Filles et garçons : des pratiques déclarées similaires mais un niveau de développement différent

Le premier résultat est frappant (voir tableau 3). **Les parents ne font aucune distinction entre filles et garçons, ni dans leurs croyances, ni dans leurs pratiques** (en tout cas, dans celles qui ont été mesurées). Ils jugent utile de parler et de lire à leur enfant au même âge, portent le même regard sur le rôle des écrans dans l'acquisition du langage, et lisent, jouent, chantent, décrivent leurs actions et exposent leur enfant aux écrans dans des proportions statistiquement similaires, que l'enfant soit une fille ou un garçon. Cette absence de différence n'est d'ailleurs pas un artefact des moyennes : la distribution complète des pratiques est elle aussi identique : la part de parents qui lisent tous les jours, comme celle qui ne lisent jamais, ne varie pas selon le sexe de l'enfant.

Et pourtant, les filles sont significativement plus avancées sur l'ensemble des dimensions du développement. Leur âge de développement langagier est supérieur d'environ deux mois en expression et d'un mois et demi en compréhension ; leur risque de retard de langage est nettement plus faible (par exemple 37 % de haut risque en expression, contre 54 % pour les garçons, et deux fois moins de très haut risque) ; et elles devancent également les garçons sur le développement socio-émotionnel. Toutefois, on peut noter que cette différence entre filles et garçons dans le domaine du langage est régulièrement décrite dans la littérature scientifique. Il ne s'agit pas d'une particularité inhabituelle spécifique à notre échantillon.

ii) Ce qui prédit le niveau de développement des enfants

Afin de déterminer quels éléments pourraient prédire le niveau de développement des enfants, nous avons réalisé des régressions de ce niveau de développement sur les principales caractéristiques du ménage, puis nous avons ajouté les pratiques parentales (voir tableau 4). Trois enseignements se dégagent :

- **Le capital éducatif prime sur le revenu.** Le diplôme du parent interrogé est un prédicteur fort du développement (langagier comme socio-émotionnel), alors que le revenu, une fois le diplôme pris en compte, n'a plus d'association significative. C'est le niveau d'éducation, davantage que le pouvoir d'achat, qui influe sur le développement de l'enfant.
- **La lecture partagée est le levier le plus puissant.** Lire fréquemment avec son enfant est associé à un développement langagier nettement supérieur, et cette pratique absorbe une partie de l'effet du diplôme, signe que c'est en partie par la pratique de lecture que le capital éducatif se transmet.

- **Les écrans sont négativement associés au développement langagier de l'enfant**, après avoir contrôlé pour son environnement socio-économique. C'est-à-dire qu'une fois qu'on a tenu compte de l'environnement socio-économique, une exposition plus importante aux écrans est associée à un niveau de développement langagier plus limité.

Tableau 3 : différences de croyances et pratiques parentales et de niveau de développement entre filles et garçons

Dimension	Garçons (N=423)	Filles (N=379)	Diff. (F-G)	p-value
<i>Croyances parentales</i>				
Croit utile de parler dès la naissance (0/1)	0,79	0,77	-0,02	0,41
Âge jugé utile pour parler (mois)	1,7	2,1	+0,5	0,16
Âge jugé utile pour les livres (mois)	9,6	10,0	+0,4	0,48
Importance de lire aux tout-petits (≤ 6 mois)	0,44	0,48	+0,05	0,20
Croit que les écrans aident (0/1)	0,62	0,59	-0,02	0,53
<i>Pratiques parentales</i>				
Lecture (fois/sem.)	3,5	3,7	+0,3	0,14
dont : lit tous les jours	0,29	0,32	+0,03	0,32
dont : ne lit jamais	0,13	0,12	-0,02	0,52
Jeu (fois/sem.)	5,2	5,3	+0,0	0,89
Chant (fois/sem.)	4,2	4,5	+0,3	0,13
Décrit ses actions (0/1)	0,40	0,46	+0,05	0,13
Télévision (heures/sem.)	3,9	3,4	-0,5	0,17
Écrans, durée totale (min/sem.)	295	253	-42	0,14
TV laissée en fond (souvent/tjs)	0,39	0,40	+0,00	0,96
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	19,6	21,6	+2,0***	0,00
Âge langagier compréhension (mois)	18,7	20,2	+1,6***	0,00
Haut risque expression (0/1)	0,54	0,37	-0,17***	0,00
Très haut risque expression (0/1)	0,26	0,13	-0,14***	0,00
Haut risque compréhension (0/1)	0,68	0,51	-0,18***	0,00
Très haut risque compréhension (0/1)	0,30	0,19	-0,12***	0,00
Vocabulaire (é.-t.)	-0,17	0,19	+0,36***	0,00
Syntaxe (é.-t.)	-0,15	0,17	+0,31***	0,00
Comptines (é.-t.)	-0,14	0,16	+0,30***	0,00
Indice socio-émotionnel (é.-t.)	-0,07	0,08	+0,15**	0,04

Lecture : par exemple, 62 % des parents de garçons et 59 % des parents de filles pensent que les écrans aident à parler (différence non significative, $p = 0,53$) ; à l'inverse, l'âge de développement langagier en expression est de 19,6 mois chez les garçons contre 21,6 mois chez les filles (différence significative, $p < 0,01$). *Note* : moyennes par sexe, ensemble des 802 familles ayant répondu au questionnaire avant le début de leur accompagnement. Diff. = filles moins garçons ; p-value d'un test t. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Ces derniers points sont au cœur de l'intervention de 1001mots. En effet, parmi les déterminants du développement de l'enfant, certains sont des caractéristiques données, sur lesquelles le programme 1001mots ne peut pas agir, comme le sexe de l'enfant à la naissance ou le niveau de diplôme de ses parents. La fréquence de lecture, en revanche, est directement modifiable avec un accompagnement adapté, et c'est précisément un des principaux leviers que le programme cherche à actionner.

Tableau 4 : analyse des déterminants du développement langagier et socio-émotionnel des enfants

	(1) Langage	(2) Langage	(3) Socio-émo.	(4) Socio-émo.
Âge de l'enfant (mois)	0,111*** (0,007)	0,113*** (0,007)	0,057*** (0,008)	0,058*** (0,008)
Fille (0/1)	0,297*** (0,061)	0,284*** (0,061)	0,149** (0,071)	0,134* (0,071)
Revenu par UC (milliers €)	0,120 (0,075)	0,052 (0,078)	0,121 (0,086)	0,070 (0,088)
Diplôme $\geq$ bac (0/1)	0,315*** (0,064)	0,270*** (0,062)	0,192** (0,075)	0,169** (0,076)
Foyer biparental (0/1)	-0,016 (0,064)	0,020 (0,064)	-0,094 (0,072)	-0,076 (0,073)
Nombre de langues	-0,003 (0,050)	-0,002 (0,049)	0,042 (0,055)	0,038 (0,054)
Lecture tous les jours		0,354*** (0,071)		0,311*** (0,077)
Écrans élevés ($\geq$ 7h/sem.)		-0,152** (0,074)		-0,102 (0,086)
Constant	-3,121*** (0,207)	-3,142*** (0,205)	-1,725*** (0,254)	-1,754*** (0,253)
N	767	754	767	754
R ²	0,287	0,319	0,080	0,106

Lecture : chaque coefficient indique l'association entre une caractéristique et le niveau de développement (exprimé en écart-type de l'indice), les autres facteurs étant tenus constants. Par exemple, à caractéristiques égales, être une fille est associé à un développement langagier supérieur de 0,297 écart-type (colonne 1, significatif à 1 %). *Note* : L'indice de développement langagier est un indice standardisé qui combine les deux âges de développement mesurés par l'IDE (expression et compréhension). Les coefficients se lisent donc en écart-type. L'indice socio-émotionnel est construit de la même façon. « Lecture tous les jours » = lecture partagée 7 fois/semaine (30 % des familles). « Écrans élevés » = exposition $\geq$ 7 h/semaine ($\approx$ 1 h/jour), soit le 79^e centile de la distribution (21 % des familles au-dessus). Erreurs-types robustes entre parenthèses. * p<0,10 ; ** p<0,05 ; *** p<0,01.

3. Comment évaluer l'impact du programme 1001mots ?

a. Impact et évaluation d'impact

i) Qu'est-ce que l'impact d'un programme ?

L'impact d'un programme peut se résumer à une question très simple : **que sont devenus les bénéficiaires grâce au programme, par rapport à ce qu'ils seraient devenus sans lui ?**

Dans le cas du programme 1001mots, mesurer l'impact revient par exemple à comparer le développement d'un enfant ayant bénéficié de l'accompagnement à celui qu'aurait connu ce même enfant, au même moment, s'il n'en avait pas bénéficié.

La difficulté est évidente : ces deux situations ne peuvent jamais être observées en même temps. Nous pouvons observer le développement d'un enfant qui a bénéficié du programme 1001mots, mais nous ne saurons jamais directement ce qu'il serait devenu sans le programme. Cette situation non observée est appelée le **contrefactuel**.

Tout l'enjeu de l'évaluation d'impact consiste donc à reconstruire ce contrefactuel de la manière la plus crédible possible.

ii) Observer une évolution ne suffit pas à démontrer un impact

Une première approche pourrait consister à mesurer les enfants avant leur participation à 1001mots, puis quelques mois plus tard. Si leurs compétences se sont améliorées entre les deux mesures, pourrait-on en conclure que le programme a eu un impact ?

Non. Et cette distinction est particulièrement importante dans le champ de la petite enfance.

Entre deux mesures, un jeune enfant grandit, apprend de nouveaux mots, développe ses capacités cognitives et motrices et interagit avec son environnement. Même sans aucune intervention, son développement évolue rapidement.

Une comparaison « avant/après » permet donc de mesurer une évolution, mais pas d'identifier ce qui, dans cette évolution, est réellement attribuable à 1001mots. Une progression observée après le programme est associée à la participation au programme, mais cela ne signifie pas nécessairement qu'elle a été causée par celui-ci.

C'est toute la différence entre corrélation et causalité : observer deux phénomènes qui vont ensemble ne suffit pas à démontrer que l'un est à l'origine de l'autre.

iii) Un groupe de comparaison pour approcher ce qui se serait passé sans 1001mots

Puisqu'il est impossible d'observer directement ce que seraient devenus les mêmes enfants sans 1001mots, l'évaluation d'impact s'appuie généralement sur un groupe de comparaison.

L'idée est de trouver un ensemble d'enfants et de familles permettant d'approcher au mieux ce qui serait arrivé aux bénéficiaires en l'absence du programme. L'évolution de ce groupe sert alors de référence pour distinguer l'effet de 1001mots des évolutions qui auraient eu lieu de toute façon.

Mais disposer d'un groupe de comparaison ne suffit pas : encore faut-il qu'il constitue une comparaison crédible.

Par exemple, comparer les enfants accompagnés par 1001mots à des enfants issus de familles très différentes pourrait conduire à attribuer au programme des écarts qui s'expliquent en réalité par le niveau socio-économique des familles, l'environnement éducatif des enfants ou d'autres caractéristiques préexistantes.

C'est là que réside une grande partie de la difficulté, et de l'intérêt, de l'évaluation d'impact : construire un groupe de comparaison qui permette réellement d'isoler l'effet causal du programme. Toutes les évaluations ne fournissent donc pas le même niveau de preuve. La méthode utilisée pour construire ce contrefactuel détermine en grande partie la crédibilité des résultats obtenus.

Demander aux parents s'ils perçoivent des progrès permet-il de mesurer l'impact ?

Il peut être tentant de demander directement aux parents : « Trouvez-vous que votre enfant a progressé depuis qu'il participe au programme 1001mots ? » Une telle question peut être intéressante pour comprendre l'expérience et la perception des familles, **mais elle ne permet pas de mesurer l'impact du programme.**

La raison est la même que pour une comparaison avant/après. Un jeune enfant se développe naturellement au fil du temps. Un parent peut donc constater, à juste titre, que son enfant parle davantage ou maîtrise de nouvelles compétences, sans pouvoir déterminer quelle part de cette progression est due à 1001mots et quelle part aurait eu lieu de toute façon.

À cela s'ajoutent d'autres difficultés. Les réponses peuvent notamment être influencées par le souhait de donner une réponse positive à une organisation qui accompagne la famille (biais de désirabilité sociale), par la manière dont chacun interprète la notion de « progrès », ou encore par des biais de mémoire.

Les questions de perception ne sont donc pas inutiles : elles peuvent apporter des informations précieuses sur la satisfaction des familles, leur expérience du programme ou les changements qu'elles pensent observer. Mais une perception positive ne constitue pas une preuve d'impact.

C'est pourquoi l'évaluation présentée ici ne repose pas sur des questions ad hoc demandant aux parents d'estimer l'impact de 1001mots. Elle cherche à mesurer le développement des enfants à partir d'outils standardisés et scientifiquement validés et, surtout, à comparer leur situation à un contrefactuel construit de manière aussi crédible que possible.

b. Un groupe de comparaison composé des futurs bénéficiaires

Il existe plusieurs méthodes d'évaluation permettant d'estimer l'effet causal de l'intervention de 1001mots sur les pratiques parentales et le développement des enfants. La méthode de référence est l'évaluation randomisée contrôlée (ERC), qui consiste à composer, par tirage au sort, un groupe bénéficiaire et un groupe de comparaison statistiquement comparables en moyenne, sous réserve d'une taille d'échantillon suffisante. Cette approche avait déjà été utilisée lors des deux précédentes évaluations du programme 1001mots, la première à sa création en 2019 (de Chaisemartin et al., 2021), et la seconde en 2021. Les résultats avaient mis en évidence une amélioration significative des pratiques parentales, sans toutefois montrer d'effet statistiquement significatif sur le développement langagier des enfants.

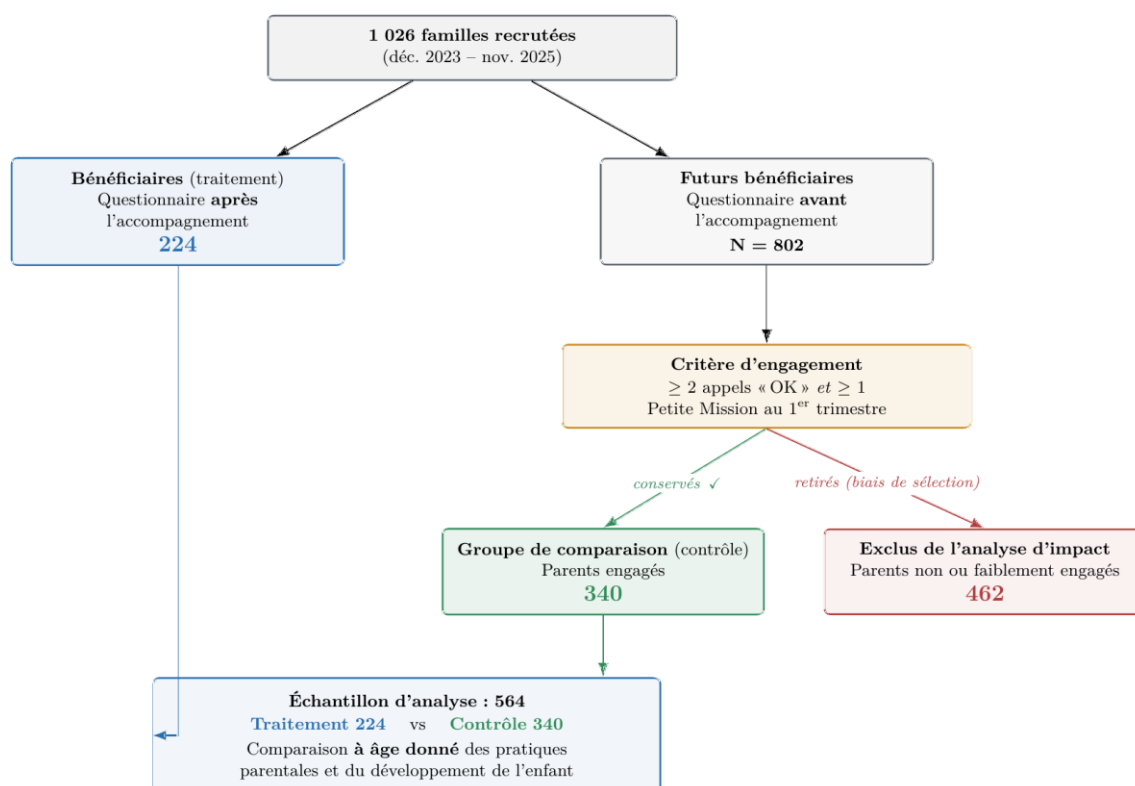
Six ans après sa création, le programme 1001mots a connu des évolutions importantes. La présente évaluation vise à démontrer l'efficacité de cette nouvelle version du programme et à consolider les preuves disponibles avant d'envisager une nouvelle évaluation randomisée contrôlée à grande échelle. Avant d'investir dans un tel dispositif, il est en effet essentiel de vérifier que le programme est correctement mis en œuvre et qu'il présente un potentiel d'impact suffisant.

Dans cette perspective, nous avons conçu, en étroite collaboration avec 1001mots, une évaluation quasi-expérimentale permettant d'estimer l'impact du programme de la manière la plus rigoureuse possible, tout en tenant compte des contraintes budgétaires et opérationnelles de l'association. À l'issue de nombreux échanges méthodologiques, nous avons élaboré un protocole conciliant faisabilité et robustesse scientifique, en cherchant à limiter autant que possible les biais de sélection.

Concrètement, l'évaluation repose sur une comparaison entre deux groupes : d'une part, des familles ayant bénéficié de l'accompagnement 1001mots, dont les données sont recueillies à l'issue du programme, et d'autre part, des familles qui intégreront prochainement le dispositif, mais dont les données sont collectées avant le début de leur accompagnement. Cette stratégie permet de comparer les pratiques parentales et le développement des enfants entre des bénéficiaires et des « futurs bénéficiaires ». Ces derniers constituent un groupe de comparaison particulièrement pertinent,

puisqu'ils présentent des caractéristiques proches des familles déjà accompagnées tout en n'ayant pas encore été exposés à l'intervention.

Figure 6 : schéma traduisant la constitution des deux groupes, bénéficiaire et de comparaison



Groupe de comparaison et niveau d'engagement. Un enjeu méthodologique important concerne le niveau d'engagement des familles dans le programme. Parmi les parents qui souhaitent initialement bénéficier de l'accompagnement 1001mots, tous ne participent pas avec la même intensité. Certains ne répondent à aucun appel d'accompagnement, ou à un seul. D'autres participent aux appels mais ne souhaitent pas se fixer de « petite mission » (PM), c'est-à-dire une nouvelle pratique simple à essayer au quotidien avec leur enfant. Enfin, certains parents se fixent une petite mission mais ne la mettent finalement pas en pratique.

Les raisons de ce moindre engagement peuvent être diverses : un besoin qui évolue ou disparaît, un manque de disponibilité, une perte d'intérêt au fil du temps ou encore un intérêt initial principalement lié à la possibilité de recevoir des livres gratuitement, sans souhait particulier de bénéficier de l'ensemble de l'accompagnement proposé.

Cette différence de niveau d'engagement entre les familles peut créer un biais de sélection. En effet, les familles qui répondent à plusieurs appels et qui essaient ou réussissent au moins une petite mission peuvent être différentes de celles qui

participent peu au programme. Elles peuvent, par exemple, être plus motivées, plus disponibles ou avoir une organisation familiale différente. Or, ces caractéristiques peuvent elles-mêmes influencer les pratiques parentales ou le développement de l'enfant, indépendamment de l'accompagnement 1001mots.

Il ne serait donc pas pertinent de comparer directement les familles les plus engagées dans le programme 1001mots à l'ensemble des futurs bénéficiaires. Le premier groupe serait composé uniquement de familles qui se sont effectivement investies dans l'accompagnement, tandis que le second comprendrait aussi bien des familles qui s'investiront fortement que des familles qui participeront peu ou pas au programme lorsqu'il leur sera proposé.

Si des différences étaient ensuite observées entre ces deux groupes, il serait alors difficile de savoir si elles sont dues à l'effet de l'accompagnement 1001mots ou simplement au fait que les familles les plus engagées étaient différentes dès le départ.

Afin de limiter ce risque, nous appliquons au groupe de comparaison les mêmes critères d'engagement que ceux retenus pour le groupe de bénéficiaires : avoir répondu à au moins deux appels et avoir essayé ou réussi au moins une petite mission. Ainsi, parmi les 802 familles interrogées avant leur entrée dans le dispositif, seules celles qui remplissent ensuite ces critères au cours de leur accompagnement sont conservées dans le groupe de comparaison. Cette sélection est réalisée a posteriori, une fois le premier trimestre d'accompagnement achevé. Elle permet de comparer des familles présentant des niveaux d'engagement similaires et ainsi de limiter le biais de sélection décrit précédemment.

Enfin, compte tenu des ressources disponibles pour l'évaluation, il n'a pas été possible d'interroger chaque famille à deux reprises, avant puis après son accompagnement. Le protocole repose donc sur la comparaison de deux groupes distincts : d'un côté, des bénéficiaires interrogés à l'issue de leur accompagnement ; de l'autre, de futurs bénéficiaires interrogés avant leur entrée dans le programme, mais dont on vérifie par la suite qu'ils présentent un niveau d'engagement comparable.

L'objectif est ainsi de comparer des familles qui se ressemblent en termes d'engagement dans le programme 1001mots, mais qui se distinguent au moment de l'enquête par un élément essentiel : **les unes ont déjà bénéficié de l'accompagnement, tandis que les autres ne l'ont pas encore commencé.**

Comparabilité de l'âge des enfants. L'entrée dans le programme 1001mots n'est pas conditionnée à un âge précis de l'enfant. Ainsi, à un même âge, certains enfants arrivent au terme de l'accompagnement tandis que d'autres commencent tout juste à en bénéficier. Il est donc possible de comparer des enfants de même âge, mais qui se trouvent à des étapes différentes de leur parcours dans le programme.

d. Les participants à l'étude

Les données ont été recueillies sur un échantillon de 1 026 participants, répartis en trois catégories selon l'état de leur accompagnement. Les participants ont été recrutés dans l'étude entre mars 2025 et mai 2026, elles se sont toutes inscrites au programme suite à un mail de la CAF de Seine-Saint-Denis. Nous avons choisi d'évaluer uniquement des parents bien engagés dans le programme, afin de mesurer l'impact du programme 1001mots quand il est pleinement utilisé par les participants. Pour s'assurer du niveau effectif d'engagement des parents dans le programme, l'équipe 1001mots s'est basée sur sa réponse aux appels d'accompagnement et le fait d'essayer ou non les petites missions co-construites lors de ces appels, comme décrit dans la section précédente. L'inclusion ou l'exclusion des familles dans l'expérimentation repose ainsi sur un critère d'engagement prédéfini, une famille est incluse dans l'étude lorsqu'elle satisfait à la condition suivante : avoir bénéficié d'au moins deux appels d'accompagnement complets et avoir essayé ou réussi au moins une petite mission au cours du premier trimestre de l'accompagnement.

En fonction du moment de passation du questionnaire et du niveau d'engagement des parents, les 1 026 répondants sont répartis dans les catégories suivantes :

- **Individus bénéficiaires** (traitement) : 224 participants ayant répondu au questionnaire après leur accompagnement et répondant au critère d'engagement.
- **Individus de comparaison** (contrôle) : 340 participants ayant répondu au questionnaire avant le démarrage de leur accompagnement, puis répondant au critère d'engagement lors de leur accompagnement, proposé à la suite de leur participation.
- **Individus exclus de l'étude** : 462 participants ayant répondu au questionnaire avant l'accompagnement, mais n'ayant pas satisfait le critère d'engagement lors de leur accompagnement, proposé à la suite de leur participation. Ils correspondent aux observations perdues du groupe de comparaison (i.e., les parents "non-engagés" supprimés pour limiter le biais de sélection).

L'échantillon mobilisé pour l'évaluation comprend donc 564 participants (224 individus bénéficiaires et 340 individus de comparaison).

e. Comparabilité des groupes

Le tableau 5 ci-dessous compare les caractéristiques des familles bénéficiaires et de celles du groupe de comparaison. Il met en évidence plusieurs différences socio-démographiques statistiquement significatives entre les deux groupes.

Premièrement, les familles bénéficiaires apparaissent en moyenne plus favorisées sur le plan socio-économique : elles sont plus souvent biparentales, disposent d'un revenu par unité de consommation supérieur de 167 €, les parents sont plus souvent en emploi et légèrement plus diplômés, les familles ont un peu moins d'enfants et recourent davantage à la crèche comme mode de garde principal. Si ces différences sont statistiquement significatives, leur ampleur reste modérée pour certaines caractéristiques. Deuxièmement, les enfants du groupe bénéficiaire sont plus souvent des filles.

Ces différences sont importantes à prendre en compte car plusieurs de ces caractéristiques sont susceptibles d'être liées au développement de l'enfant. Les analyses présentées en section 2.d le confirment : dans nos données, certaines des caractéristiques qui distinguent les deux groupes sont également associées aux indicateurs de développement que nous cherchons à mesurer. Une comparaison directe des résultats des deux groupes risquerait donc d'attribuer à 1001 mots des différences qui pourraient, au moins en partie, s'expliquer par leurs caractéristiques initiales.

Ce constat souligne l'importance de la stratégie d'ajustement mise en œuvre dans l'analyse. Le questionnaire a été conçu pour recueillir un ensemble riche de caractéristiques des enfants et de leur environnement familial, notamment celles que la littérature identifie comme liées au développement de l'enfant. Les méthodes présentées ci-après permettent de tenir compte de ces différences observées afin de rendre les deux groupes aussi comparables que possible sur ces dimensions.

Cette approche ne permet toutefois pas d'exclure entièrement l'existence de différences non observées entre les deux groupes. Elle permet en revanche de corriger les principaux déséquilibres que nous pouvons observer et mesurer. Le fait que la restriction aux familles engagées ne suffise pas, à elle seule, à rendre les deux groupes comparables souligne ainsi l'importance de combiner cette restriction avec un ajustement statistique rigoureux.

Tableau 5 : tableau de comparaison entre le groupe de comparaison (contrôle) et le groupe de bénéficiaires (traitement)

Variable	Contrôle (N = 340)	Traitement (N = 224)	Différence	p
Âge de l'enfant (mois)	23,4	23,4	+ 0,0	0,999
Fille	0,43	0,51	+0,08*	0,064
Foyer biparental	0,62	0,86	+0,24***	0,000
Revenu par UC (€)	1007	1175	+ 167***	0,000
Diplôme ≥ bac	0,72	0,80	+0,08**	0,023
Père en emploi	0,55	0,73	+0,19***	0,000
Mère en emploi	0,28	0,39	+0,11***	0,007
Garde en crèche	0,26	0,41	+0,14***	0,000
Nombre de langues	1,92	2,02	+0,10*	0,065
Né en France	0,45	0,40	-0,05	0,289
Nombre d'enfants	2,57	2,37	-0,19*	0,099

Lecture : par exemple, 86 % des familles bénéficiaires sont biparentales contre 62 % du groupe de comparaison (écart significatif à 1 %). *Notes* : 564 familles : 340 familles de comparaison, engagées et ayant montré un niveau d'engagement suffisant, et 224 familles bénéficiaires ayant terminé leur accompagnement. Différence = Traitement – Contrôle. *p* correspond à la p-value d'un test t de la différence. Seuils de significativité : * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

f. Méthodes statistiques pour estimer l'impact du programme

Comme les deux groupes présentent des différences significatives dans leurs caractéristiques initiales, une comparaison directe ne permettrait pas d'isoler l'impact du programme : les écarts observés pourraient refléter à la fois l'effet de 1001mots et les différences qui existaient déjà entre les deux groupes. Ainsi, pour améliorer cette comparaison et limiter ce biais de sélection, nous utilisons trois modélisations statistiques :

i) Méthode par ajustement statistique (régression multivariée)

La régression multivariée est une méthode statistique qui permet de comparer deux groupes en tenant compte simultanément de leurs différences de départ. Elle répond à la question : "A profil identique, quel est l'effet du programme ?"

Concrètement, pour chaque famille, on dispose d'informations sur sa situation socio-économique (type de foyer, niveau de revenus, emploi des parents, etc.). La

régression utilise ces informations pour corriger statistiquement les différences entre groupes, et isoler la part de l'effet qui est vraiment attribuable au programme.

Dans cette analyse, nous ne limitons pas les variables de correction à celles qui diffèrent significativement entre les groupes. Nous contrôlons au contraire l'ensemble des caractéristiques pré-traitement susceptibles d'influencer le développement de l'enfant, qu'elles soient ou non significativement différentes entre bénéficiaires et contrôles : l'âge et le sexe de l'enfant, la structure familiale, le revenu par unité de consommation, le niveau d'éducation du répondant, l'emploi du père et de la mère, le mode de garde, le nombre d'enfants du foyer, le nombre de langues parlées, le niveau de français et le fait d'être né en France ou à l'étranger.

Ce choix est délibéré : sélectionner les variables de contrôle sur la seule base d'un test de différence serait fragile, car la significativité d'un tel test dépend avant tout de la taille de l'échantillon, et non de l'ampleur réelle du facteur : une variable peut différer substantiellement sans être « significative », et inversement. Nous retenons donc toutes les variables indépendantes du programme mesurées dans notre questionnaire, ce qui rend l'ajustement d'autant plus fiable.

Le résultat est un effet estimé du programme qui est « nettoyé » de ces différences de départ. La limite naturelle de cette méthode est qu'elle corrige les différences que l'on peut mesurer, mais ne peut pas corriger des différences que l'on n'a pas mesurées. Nous répondons à cette limite de deux façons : d'une part en recoupant ce résultat avec d'autres méthodes d'ajustement, d'autre part au moyen d'un test de robustesse spécifiquement dédié à cette question, les bornes d'Oster, détaillées ci-dessous.

ii) Méthode par appariement statistique (matching)

Le *matching*, ou appariement, consiste à composer des paires de familles très similaires, une du groupe traitement et une du groupe contrôle, et à ne comparer que ces paires. Les familles du groupe contrôle qui n'ont pas de "jumeau" comparable dans le groupe traitement sont écartées de l'analyse.

On obtient ainsi deux groupes qui se ressemblent beaucoup sur toutes les caractéristiques mesurées, et la comparaison entre eux est plus directe.

Pour cette analyse, nous retenons deux modèles d'appariement :

- **Coarsened Exact Matching (CEM)** : Le CEM regroupe les familles en "cases" selon leur profil et s'assure que chaque case contient à la fois des familles bénéficiaires et des familles de comparaison. Dans cette analyse, 153 strates ont été constituées, retenant 182 familles bénéficiaires et 218 familles de comparaison, soit 400 familles sur les 564 de l'échantillon total. Cette méthode

est particulièrement intéressante car elle traite naturellement les variables continues et les variables binaires, et qu'elle fonctionne bien avec des échantillons de taille modeste.

- **Propensity Score Matching (PSM)** : Le PSM est une autre méthode d'appariement. Plutôt que de grouper les familles case par case comme le CEM, le PSM calcule pour chaque famille une probabilité d'avoir été dans le groupe traitement, ce qu'on appelle le "score de propension", puis compare des familles du groupe de comparaison ayant une probabilité similaire. Plutôt que de chercher une seule famille "jumelle" pour chaque famille traitée, ce qui gaspillerait beaucoup d'observations, la variante utilisée ici (dite "à noyau") utilise toutes les familles de comparaison en leur donnant un poids proportionnel à leur proximité : une famille de comparaison très similaire compte beaucoup, une famille moins similaire compte moins. En pratique, cette méthode retient 204 familles traitées et 325 contrôles, soit presque la totalité de l'échantillon (529 familles sur 564 familles éligibles, c'est-à-dire celles où nous avons des informations pour les différentes variables de contrôle). Ce taux de rétention élevé indique que les deux groupes ont beaucoup de caractéristiques en commun et que le biais de sélection n'est pas extrême.

iii) Méthode par repondération statistique (entropy balancing)

La repondération part d'une idée différente de l'appariement : plutôt que d'écarter des familles trop éloignées ou de chercher des « jumeaux », on conserve toutes les familles de comparaison et on ajuste leur poids dans le calcul, c'est-à-dire l'importance qu'on leur accorde.

L'intuition est la suivante : on souhaite que le groupe de comparaison ressemble, en moyenne, exactement au groupe bénéficiaire. Pour y parvenir, on accorde un peu plus de poids aux familles de comparaison dont le profil ressemble à celui des bénéficiaires, et un peu moins à celles qui s'en éloignent. C'est le même principe que le redressement d'un sondage, où l'on pondère les répondants pour que l'échantillon reflète fidèlement la population.

L'*entropy balancing* effectue cette opération de façon particulièrement rigoureuse : il calcule les poids de sorte qu'une fois pondéré, le groupe de comparaison présente exactement la même moyenne que le groupe bénéficiaire sur chacune des caractéristiques mesurées, simultanément : même revenu moyen, même part de filles, même niveau d'éducation, même mode de garde, etc. Les déséquilibres constatés dans le tableau 2 disparaissent alors par construction : les deux groupes sont rendus identiques sur l'ensemble des variables observées, sans qu'il ne reste rien à corriger.

Cette approche présente deux atouts. D'une part, elle ne perd aucune famille : contrairement à l'appariement qui écarte les familles sans correspondant, l'ensemble

des observations est conservé. D'autre part, elle vise directement l'équilibre recherché, sans passer par l'estimation préalable d'un modèle. Par ailleurs, elle choisit les poids les plus proches possibles de poids égaux, ce qui évite de faire reposer le résultat sur une poignée de familles au poids démesuré.

Comme pour les deux méthodes statistiques précédentes, cette méthode tient compte de l'ensemble des caractéristiques mesurées, dont l'âge et le sexe de l'enfant.

iv) Robustesse : sensibilité aux facteurs inobservables

Les quatre méthodes précédentes corrigent les différences observées entre les familles bénéficiaires et le groupe de comparaison. La crédibilité de l'analyse repose cependant sur une hypothèse plus exigeante : qu'il ne subsiste pas de facteur inobservé - une caractéristique que nous n'aurions pas mesurée - qui expliquerait à la fois l'engagement des familles et le développement de l'enfant, et donc une partie des écarts attribués au programme.

Pour éprouver cette hypothèse, nous mobilisons le test des bornes d'Oster². L'idée est la suivante : on observe de combien l'effet estimé se déplace lorsqu'on ajoute les facteurs que l'on sait mesurer (âge, sexe, revenu, diplôme, emploi, langue...), puis on extrapole ce mouvement au cas où l'on observerait aussi les facteurs cachés. On en déduit la « force » que devrait avoir un confondeur inobservé pour annuler complètement l'effet du programme.

Le résultat est rassurant. Sur le développement de l'enfant, un tel facteur caché devrait être 5 à 8 fois plus déterminant que l'ensemble de tout ce que nous avons déjà mesuré réuni : âge, sexe, revenu, diplôme, emploi et langue pris ensemble. Autrement dit, il faudrait qu'un élément totalement passé sous les radars pèse bien plus lourd que tous ces déterminants à la fois, un scénario assez peu probable. Les effets estimés sur le développement de l'enfant apparaissent donc robustes à un éventuel biais de sélection sur l'inobservable (Tableau 6).

² Oster, E. (2019). Unobservable Selection and Coefficient Stability. *Journal of Business & Economic Statistics*, 37(2), 187-204.

Tableau 6 : bornes d'Oster sur les dimensions de développement de l'enfant

Dimension	Effet	δ^* (recommandé)	δ^* (conservateur)
Âge langagier — expression	+2,28	6,1	1,6
Âge langagier — compréhension	+1,80	7,5	1,6
Vocabulaire (é.-t.)	+0,44	6,1	1,3
Syntaxe (é.-t.)	+0,50	7,7	1,8
Comptines (é.-t.)	+0,62	5,6	0,8
Socio-émotionnel (é.-t.)	+0,42	5,1	0,6

Lecture (première ligne) : le programme augmente l'âge langagier en expression de +2,28 mois (colonne « Effet »). Pour qu'un biais de variables omises annule entièrement cet effet, il faudrait que la sélection sur les caractéristiques *non observées* soit 6,1 fois plus forte que la sélection sur les caractéristiques observées sous l'hypothèse recommandée, ou 1,6 fois plus forte sous l'hypothèse conservatrice — dans les deux cas très au-delà de 1, ce qui signale un effet robuste.

Note : δ^* = force de la sélection sur les caractéristiques inobservées (relative à celle sur les observées) qui suffirait à ramener l'effet à zéro ; l'effet est jugé robuste lorsque $\delta^* > 1$ (Oster, 2019). Contrôles = spécification principale (12 covariables). Borne « recommandée » : $R_{\max} = 1,3 \times R^2$; borne « conservatrice » : $R_{\max} = 1$. Les bornes calculées sur les résultats parentaux ne sont pas reportées : leur R^2 est faible (0,02–0,14), ce qui rend δ^* peu informatif.

4. L'impact du programme 1001mots

a. Impact sur les croyances et pratiques parentales

La tableau 7 ci-dessous présente les mesures de l'impact du programme 1001mots sur les croyances et les pratiques parentales. La première colonne estime l'impact du programme avec un modèle OLS de régression multivariée, la seconde colonne avec un modèle matching de type Coarsened Exact Matching (CEM), la troisième colonne avec un modèle matching de type Propensity Score Matching (PSM) kernel et la quatrième colonne présente les résultats du modèle par *entropy balancing* (EB). La convergence des quatre méthodes renforce la confiance dans les résultats.

En résumé, la tableau 7 indique que :

- **Les parents bénéficiaires améliorent leurs croyances, en particulier dans le domaine de la lecture.** L'âge auquel ils jugent utile de commencer à regarder des livres avec les bébés est réduit significativement : -3,3 mois, soit une réduction de 35 % par rapport au groupe de comparaison (de 9,4 à 6,1 mois). Autrement dit, les parents bénéficiaires sont nettement plus nombreux à juger utile de lire à un tout-petit (dès 6 mois ou avant) : 67 % contre 48 % dans le groupe de comparaison, soit un écart de +19 points (significatif à 1 %). Le programme ne réduit pas seulement l'âge moyen jugé utile pour la lecture : il fait basculer une part importante des familles vers l'idée qu'on peut commencer

la lecture dès les tout premiers mois de l'enfant. Les parents bénéficiaires sont également un peu moins nombreux à penser que les écrans aident l'enfant à apprendre à parler : cette réponse recule d'environ la moitié (–52 %, de 12 % à 6 % entre les deux groupes).

« On lit beaucoup mais je le faisais pas aussi tôt, et j'ai découvert qu'on pouvait le faire même quand ils comprennent pas. Au fil des mois, on voit que c'est encaissé au fond d'elle, et maintenant elle sort tout ! »

— enfant de 30 mois

- **Les parents bénéficiaires adoptent de meilleures pratiques de lecture et déclarent lire davantage avec leurs enfants** : la fréquence de lecture augmente d'environ une unité (+37 %), passant de 3,4 à 4,6 fois par semaine. Surtout, cet effet moyen comprend un rééquilibrage aux deux extrémités de la distribution : la part de parents qui ne lisent jamais est divisée par plus de deux (–61 %, de 12 % à 5 %), tandis que la part de ceux qui lisent tous les jours augmente de plus de 80 % (de 26 % à 47 %). Le programme ne fait donc pas que déplacer légèrement la moyenne : il fait réellement basculer une part importante des familles vers une lecture quotidienne. Les autres pratiques stimulantes progressent également : la description de ce que l'on fait au quotidien avec l'enfant (+41 %, de 39 % à 55 %) et, dans une moindre mesure, le chant (+15 %).

« Je me suis mise à montrer des livres grâce à 1001 mots, je ne le faisais pas avant. J'ai compris que c'était important que mon enfant ait des livres bien adaptés à son âge pour que ça puisse l'intéresser. Avant, je n'en avais pas de bien adaptés. C'est un grand bienfait tout ce que mon enfant a pu découvrir grâce aux livres 1001 mots. Grâce aux livres, elle a pu découvrir plein de nouvelles choses qu'elle n'avait jamais vu avec nous. »

— enfant de 29 mois

- **Les parents bénéficiaires réduisent significativement l'exposition aux écrans de leurs enfants**. La durée totale d'écran diminue de près de deux heures par semaine, soit une baisse de 41 % (spécification OLS), et jusqu'à environ 50 % selon les méthodes d'appariement (d'environ 270 minutes à 140–160 minutes). De plus, la part d'enfants fortement exposés (au moins 7 heures par semaine) recule de 45 % (de 23 % à 13 %). Cette baisse concerne essentiellement la télévision, à la fois en durée (–41 %, d'environ 225 à 130 minutes par semaine) et en fréquence de visionnage (–23 %, de 3,8 à 2,9 fois par semaine). Cette baisse ne concerne pas seulement le temps d'écran « actif

» : elle touche aussi les écrans passifs. La part de familles dont la télévision reste souvent ou tout le temps allumée en fond, même sans être activement regardée, est plus faible chez les bénéficiaires : 25 % contre 41 % dans le groupe de comparaison (-16 points), un effet significatif à 1 %. En revanche, aucune différence significative n'apparaît concernant l'exposition aux téléphones ou tablettes, cette pratique étant déjà très faible initialement (environ 45 minutes par semaine dans le groupe de comparaison).

« Avant ça m'intéressait pas les livres, et maintenant j'en achète pas mal ! [...] Elle regarde même plus les écrans ma grande grâce à vous. »

— enfant de 22 mois

b. Impact sur le développement des enfants

La tableau 8 ci-dessous présente les mesures d'impact du programme 1001mots sur le développement des enfants. En résumé, nous observons une amélioration significative et importante du niveau de développement des enfants :

- **Que ce soit en expression ou en compréhension, l'âge de développement langagier des enfants bénéficiaires est significativement plus élevé** : +2,3 mois en expression (de 19,8 à 22,1 mois, soit +12 %) et +1,8 mois en compréhension (de 18,8 à 20,6 mois, soit +10 %). L'écart entre l'âge de développement langagier et l'âge réel (23,4 mois en moyenne) se réduit donc nettement : en expression, il passe de 3,6 à 1,3 mois (près des deux tiers en moins) et en compréhension, il passe de 4,6 à 2,8 mois.
- **Nous observons une diminution marquée de la part d'enfants à haut ou très haut risque de retard dans le groupe bénéficiaire**. En expression, le haut risque de retard de langage recule de 47 % à 28 % (-40 %) et le très haut risque de 21 % à 9 % (-56 %). En compréhension, le haut risque passe de 62 % à 43 % (-30 %) et le très haut risque de 26 % à 13 % (-50 %). Tous ces effets sont significatifs à 1 %

« Avant, je ne disais pas forcément ce que je faisais pendant la douche. Quand j'ai eu Capucine au téléphone, elle me disait que je pouvais lui nommer tout ce que je faisais. C'est à partir de là que j'ai commencé à voir qu'il connaissait les parties du corps. [...] Capucine, elle m'avait aussi dit de citer tout ce qu'on voit (ca c'est un arbre, c'est vert, c'est le bus, le toboggan, le chien...), plein de choses qu'il peut voir. Il commence à les citer. »

— enfant de 30 mois

Tableau 7 : impact du programme 1001mots sur les croyances et les pratiques parentales

Résultat parental	Moy. comparaison	OLS	CEM	PSM	EB
<i>Croyances</i>					
Écrans aident à parler (0/1)	0,12	-0,06** (0,03) [519]	-0,04 (0,03) [373]	-0,06** (0,03) [519]	-0,07** (0,03) [519]
Âge utile pour parler (mois)	1,43	-0,35 (0,29) [517]	-0,07 (0,34) [369]	-0,33 (0,31) [517]	-0,32 (0,31) [517]
Âge utile pour les livres (mois)	9,41	-3,34*** (0,57) [511]	-2,76*** (0,72) [364]	-3,33*** (0,60) [511]	-3,30*** (0,62) [511]
Importance de lire aux tout-petits (≤ 6 mois)	0,48	0,19*** (0,05) [511]	0,17*** (0,06) [364]	0,19*** (0,05) [511]	0,18*** (0,05) [511]
<i>Pratiques de stimulation</i>					
Lecture (fois/sem.)	3,38	1,24*** (0,23) [528]	1,12*** (0,28) [377]	1,29*** (0,24) [528]	1,28*** (0,24) [528]
dont : ne lit jamais (0/1)	0,12	-0,07*** (0,02) [528]	-0,05* (0,02) [377]	-0,07*** (0,02) [528]	-0,07*** (0,03) [528]
dont : lit tous les jours (0/1)	0,26	0,21*** (0,04) [528]	0,17*** (0,06) [377]	0,22*** (0,05) [528]	0,21*** (0,05) [528]
Jeu (fois/sem.)	5,09	-0,16 (0,21) [527]	-0,15 (0,26) [376]	-0,29 (0,22) [527]	-0,20 (0,23) [527]
Chant (fois/sem.)	4,48	0,66*** (0,24) [527]	0,35 (0,28) [377]	0,65*** (0,25) [527]	0,60** (0,25) [527]
Décrit ses actions (0/1)	0,39	0,16*** (0,05) [526]	0,23*** (0,05) [376]	0,17*** (0,05) [526]	0,17*** (0,05) [526]
<i>Exposition aux écrans</i>					
Écrans : durée totale (min/sem.)	270,53	-111,27*** (33,64) [487]	-97,69** (40,02) [345]	-137,08*** (37,96) [487]	-136,83*** (38,04) [487]
Écrans : au moins 7h/semaine (0/1)	0,23	-0,10*** (0,04) [487]	-0,09* (0,05) [345]	-0,12*** (0,04) [487]	-0,12*** (0,04) [487]
TV : durée totale (min/sem.)	224,84	-93,07*** (27,84) [518]	-85,70*** (32,68) [370]	-112,05*** (31,07) [518]	-116,18*** (31,93) [518]
Téléphone : durée totale (min/sem.)	45,16	-16,46 (10,57) [496]	-13,05 (10,91) [351]	-21,23* (12,57) [496]	-16,92 (10,90) [496]
TV : fréquence (fois/sem.)	3,75	-0,87*** (0,27) [528]	-0,67** (0,31) [378]	-0,98*** (0,26) [528]	-1,00*** (0,27) [528]
TV laissée en fond souvent/tout le temps (0/1)	0,41	-0,16*** (0,04) [526]	-0,12** (0,06) [376]	-0,15*** (0,05) [526]	-0,14*** (0,04) [526]

Lecture : chaque ligne donne l'effet estimé du programme (écart entre bénéficiaires et groupe de comparaison). Par exemple, les parents bénéficiaires lisent 1,24 fois par semaine de plus que le groupe de comparaison (colonne OLS, significatif à 1%). *Notes* : la colonne *Moy. comparaison* donne la moyenne (non ajustée) de la variable dans le groupe de comparaison, servant de référence pour apprécier l'ampleur de l'effet estimé. Les colonnes suivantes rapportent l'effet estimé du traitement, avec l'erreur-type entre parenthèses et le nombre d'observations entre crochets. Les quatre colonnes correspondent à des méthodes d'estimation alternatives : OLS (*Ordinary Least Squares*), qui correspond à la régression multivariée contrôlant pour les caractéristiques socio-démographiques, CEM (*Coarsened Exact Matching*), PSM (*Propensity Score Matching*) et EB (*Entropy Balancing*). Seuils de significativité : * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

- En analysant les sous-scores du questionnaire, **nous observons une amélioration du niveau de vocabulaire** (+0,44 écart-type), **de syntaxe** (+0,50 écart-type) **et de maîtrise des comptines** (+0,62 écart-type) des enfants bénéficiaires : des effets substantiels, tous significatifs à 1 %.
- **Les enfants bénéficiaires présentent également une amélioration significative de leurs compétences socio-émotionnelles** (+0,42 écart-type, significatif à 1 %). Celle-ci est notamment portée par la capacité de concentration : la part d'enfants capables de rester concentrés au moins cinq minutes sur un livre ou une histoire progresse de 55 % à 75 % (+20 points). D'autres dimensions progressent aussi nettement, comme la participation aux conversations (+24 points) et le fait de pouvoir exprimer avec des mots ses émotions ou son mécontentement (+10 points chacun).

« Vous avez offert un livre génial qui montre les émotions des bébés. Il se met devant le livre, le lit, et j'avais filmé car il peut se concentrer longtemps dessus. »

— enfant de 19 mois

Les tableaux 7 et 8 présentent les effets estimés du programme selon quatre méthodes de mesure, ce qui permet d'en apprécier la robustesse mais rend la lecture de leur ampleur moins immédiate. Le tableau 9 propose une synthèse centrée sur l'intensité des effets. Pour chaque résultat, côté parent comme côté enfant, il met en regard la moyenne du groupe de comparaison et la valeur atteinte par le groupe bénéficiaire, puis exprime l'effet estimé (spécification OLS) en pourcentage de la moyenne du groupe de comparaison, ou directement en écart-type pour les indices standardisés. Cette présentation permet de comparer sur une même échelle relative des dimensions aussi diverses que les pratiques de lecture, l'exposition aux écrans ou le développement langagier de l'enfant, et de hiérarchiser les effets les plus marqués.

Tableau 8 : impact du programme 1001mots sur le développement langagier et socio-émotionnel des enfants

Résultat enfant	Moy. comparaison	OLS	CEM	PSM	EB
<i>Expression (LEX)</i>					
Âge langagier (mois)	19,78	2,28*** (0,43) [529]	2,57*** (0,51) [378]	2,39*** (0,42) [529]	2,40*** (0,43) [529]
Haut risque de retard	0,47	-0,19*** (0,04) [529]	-0,22*** (0,05) [378]	-0,20*** (0,05) [529]	-0,20*** (0,05) [529]
Très haut risque de retard	0,21	-0,12*** (0,03) [529]	-0,13*** (0,04) [378]	-0,13*** (0,03) [529]	-0,12*** (0,03) [529]
<i>Compréhension (LCO)</i>					
Âge langagier (mois)	18,85	1,80*** (0,37) [529]	1,99*** (0,45) [378]	1,87*** (0,36) [529]	1,96*** (0,37) [529]
Haut risque de retard	0,62	-0,19*** (0,04) [529]	-0,18*** (0,05) [378]	-0,19*** (0,05) [529]	-0,21*** (0,05) [529]
Très haut risque de retard	0,26	-0,13*** (0,03) [529]	-0,14*** (0,04) [378]	-0,13*** (0,03) [529]	-0,14*** (0,03) [529]
<i>Sous-scores (écart-type)</i>					
Vocabulaire	0,00	0,44*** (0,08) [529]	0,51*** (0,09) [378]	0,45*** (0,08) [529]	0,46*** (0,08) [529]
Syntaxe	0,00	0,50*** (0,09) [529]	0,56*** (0,11) [378]	0,54*** (0,09) [529]	0,54*** (0,09) [529]
Comptines	0,00	0,62*** (0,12) [529]	0,71*** (0,13) [378]	0,64*** (0,12) [529]	0,66*** (0,12) [529]
<i>Socio-émotionnel (écart-type)</i>					
Score global	0,00	0,42*** (0,09) [529]	0,41*** (0,10) [378]	0,43*** (0,09) [529]	0,45*** (0,09) [529]

Lecture : par exemple, l'âge de développement langagier en expression des enfants bénéficiaires est supérieur de 2,28 mois à celui du groupe de comparaison (colonne OLS, significatif à 1 %). *Notes* : la colonne *Moy. comparaison* donne la moyenne (non ajustée) de l'indicateur dans le groupe de comparaison ; les âges langagiers sont en mois et les risques de retard en proportion. Les sous-scores et l'indice socio-émotionnel sont standardisés sur l'écart-type du groupe de comparaison (moyenne 0 et écart-type 1 par construction), de sorte que le coefficient se lit directement en écart-type. Les colonnes suivantes rapportent l'effet estimé du traitement, avec l'erreur-type entre parenthèses et le nombre d'observations entre crochets. Les quatre colonnes correspondent à des méthodes d'estimation alternatives : OLS (*Ordinary Least Squares*), qui correspond à la régression multivariée contrôlant pour les caractéristiques socio-démographiques, CEM (*Coarsened Exact Matching*), PSM (*Propensity Score Matching*) et EB (*Entropy Balancing*). Seuils de significativité : * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau 9 : synthèse des tailles d'effets du programme 1001mots en comparaison avec les niveaux du groupe de comparaison

Résultat	Groupe de comparaison	Groupe traité	Intensité de l'effet	p
<i>Croyances et pratiques parentales</i>				
Écrans aident à parler (0/1)	12 %	6 %	−52 %	0,014
Âge utile pour les livres (mois)	9,4	6,1	−35 %	< 0,001
Importance de lire aux tout-petits (≤ 6 mois)	48 %	67 %	+39 %	< 0,001
Lecture (fois/sem.)	3,38	4,63	+37 %	< 0,001
dont : ne lit jamais (0/1)	12 %	5 %	−61 %	0,003
dont : lit tous les jours (0/1)	26 %	47 %	+81 %	< 0,001
Chant (fois/sem.)	4,48	5,14	+15 %	0,006
Décrit ses actions (0/1)	39 %	54 %	+41 %	< 0,001
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	271	159	−41 %	0,001
Écrans : au moins 7h/sem. (0/1)	23 %	13 %	−45 %	0,006
TV : durée (min/sem.)	225	132	−41 %	< 0,001
TV : fréquence (fois/sem.)	3,75	2,87	−23 %	0,001
TV laissée en fond (souvent/tjs)	41 %	25 %	−38 %	< 0,001
Téléphone : durée (min/sem.)	45	29	−36 %	0,120
<i>Développement langagier de l'enfant</i>				
Âge langagier – expression (mois)	19,8	22,1	+12 %	< 0,001
Haut risque – expression (0/1)	47 %	28 %	−40 %	< 0,001
Très haut risque – expression (0/1)	21 %	9 %	−56 %	< 0,001
Âge langagier – compréhension (mois)	18,8	20,6	+10 %	< 0,001
Haut risque – compréhension (0/1)	62 %	43 %	−30 %	< 0,001
Très haut risque – compréhension (0/1)	26 %	13 %	−50 %	< 0,001
<i>Qualité du langage et socio-émotionnel (écart-type)</i>				
Vocabulaire	0,00	0,44	+0,44 σ	< 0,001
Syntaxe	0,00	0,50	+0,50 σ	< 0,001
Comptines	0,00	0,62	+0,62 σ	< 0,001
Socio-émotionnel (score global)	0,00	0,42	+0,42 σ	< 0,001

Lecture : par exemple, la fréquence de lecture atteint 4,63 fois par semaine chez les bénéficiaires, contre 3,38 pour le groupe de comparaison, soit un effet relatif de +37 %. *Notes* : synthèse de l'intensité des effets du programme. Le *groupe traité* correspond à la moyenne du groupe de comparaison augmentée de l'effet estimé (spécification OLS avec contrôles). La colonne *Intensité de l'effet* exprime cet effet en pourcentage de la moyenne du groupe de comparaison, sauf pour les indices standardisés (marqués σ) exprimés directement en écart-type, leur moyenne de comparaison étant nulle par construction. Pour l'exposition totale aux écrans, la baisse atteint environ −51 % avec les méthodes d'appariement (CEM, PSM, EB). *p* : p-value du test de l'effet (OLS). Aucun effet significatif sur le temps passé sur les téléphones et tablettes.

c. Hétérogénéité des effets

Les effets présentés jusqu'ici sont des effets moyens. Il est utile de vérifier s'ils se concentrent uniquement sur certaines familles ou si, au contraire, ils bénéficient à tous de la même manière. Nous avons donc estimé, pour chaque résultat, l'effet du programme séparément dans deux sous-groupes, selon dix dimensions : l'âge de l'enfant, le revenu du foyer, le diplôme du répondant, le nombre de langues parlées, la structure familiale, la fratrie, le pays de naissance du parent, l'ancienneté d'arrivée en France (pour les familles immigrées), le niveau d'engagement des parents, et le sexe de l'enfant. Pour les variables continues (âge, revenu, engagement), la

séparation se fait à la médiane ; pour les autres variables, selon la catégorie. Chaque tableau (tableaux A1 à A10 dans l'annexe A) reporte l'effet dans le groupe A, l'effet dans le groupe B, et la p-value d'un test d'interaction indiquant si la différence entre les deux groupes est statistiquement significative.

Un effet remarquablement homogène selon le profil socio-économique des familles. C'est le principal enseignement : sur la très grande majorité des résultats, l'effet du programme est statistiquement similaire d'un sous-groupe à l'autre. Le revenu, le diplôme, le bilinguisme, la structure du foyer, la fratrie, l'ancienneté d'arrivée en France et le sexe de l'enfant ne modifient quasiment pas l'ampleur des effets : les tests d'interaction sont non significatifs presque partout. Concrètement, le programme améliore les croyances et les pratiques parentales, réduit l'exposition aux écrans et soutient le développement langagier et socio-émotionnel aussi bien dans les familles les plus défavorisées (faibles revenus, sans le bac, immigrées, non francophones) que dans les autres. C'est un résultat important, à la fois en termes d'équité et pour un futur passage à l'échelle : l'intervention ne « marche » pas seulement pour un public particulier.

La seule dimension nettement modératrice est l'âge de l'enfant (tableau A1). Pour presque tous les résultats (développement, lecture, écrans), l'effet est plus marqué pour les enfants les plus âgés (au-dessus de la médiane, soit environ 24 mois) que pour les plus jeunes. Par exemple, l'âge langagier en expression augmente de +1,0 mois chez les enfants plus jeunes, contre une augmentation de +3,7 mois chez les plus âgés. De même, l'effet mesuré sur le score de syntaxe passe de +0,2 à +0,8 écart-type et la réduction du temps d'écran, de -41 minutes (non significative) à -184 minutes par semaine entre ces deux sous-groupes. Il ne s'agit pas d'une inefficacité du programme chez les plus petits, les effets sont le plus souvent positifs et significatifs, mais bien d'une intensité croissante de ces effets avec l'âge.

Le pays de naissance du parent révèle un mécanisme intéressant sur la lecture (tableaux A8 et A9). Sur le développement de l'enfant, les effets sont homogènes selon que le parent soit né en France ou à l'étranger. En revanche, l'adoption de la lecture partagée est portée par les parents nés à l'étranger : la fréquence de lecture augmente de +1,7 fois par semaine chez eux (effet fortement significatif) contre un effet non significatif de +0,5 chez les parents nés en France. De même, l'absence de lecture recule presque exclusivement dans ce sous-groupe (les écarts d'interaction sont significatifs, $p < 0,01$). Autrement dit, ce sont surtout les familles immigrées qui font évoluer leurs pratiques de lecture grâce au programme, vraisemblablement parce qu'elles en étaient plus éloignées au départ. Parmi ces familles immigrées, en revanche, l'ancienneté d'arrivée en France ne joue pas : récemment arrivées ou installées de longue date, elles bénéficient du programme de façon comparable (aucune interaction significative, tableau A9).

Une gradation avec l'engagement des parents, à interpréter avec prudence (tableau 10). Les familles les plus engagées dans le programme (au-dessus de la médiane de l'indice d'engagement) tendent à présenter des effets plus forts sur certains résultats de développement (vocabulaire, compétences socio-émotionnelles) et sur la réduction de la non-lecture. Cette logique de « dose-réponse » est intuitive, mais elle demeure corrélationnelle : l'engagement n'est pas attribué aléatoirement, et les familles très engagées peuvent différer par ailleurs. Ce résultat est donc suggestif.

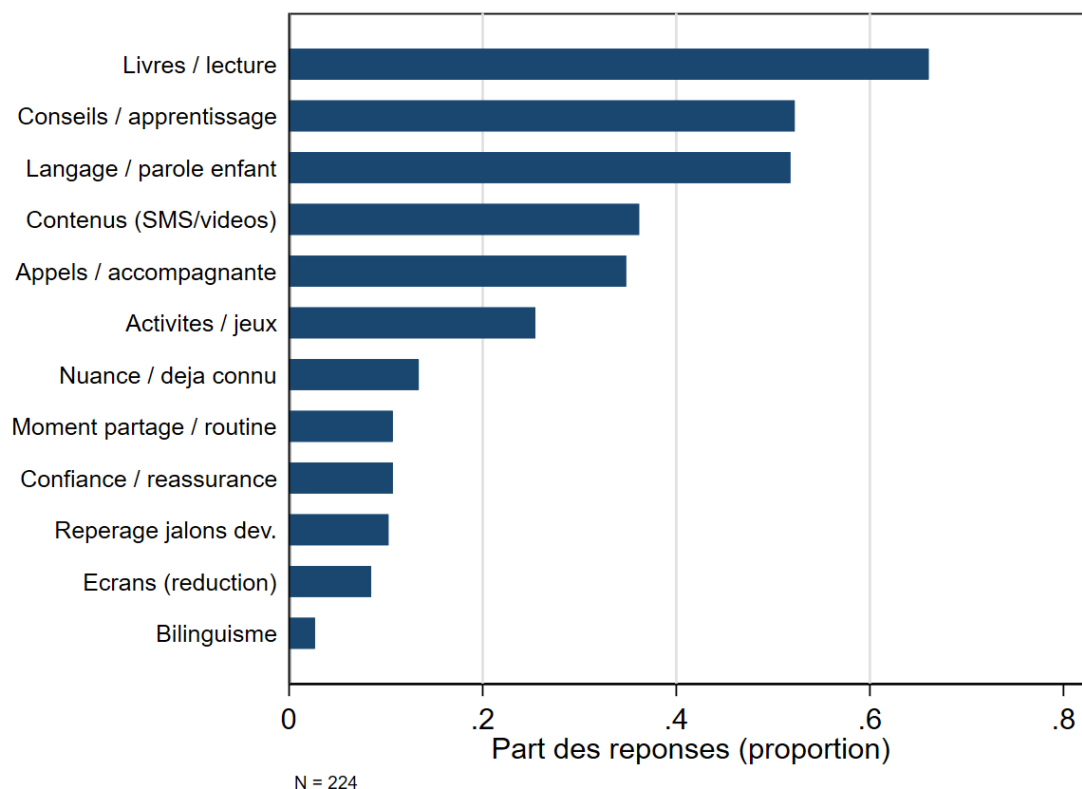
Quelques signaux isolés. Ponctuellement, certains effets diffèrent entre sous-groupes : sur la compréhension, l'effet apparaît un peu plus fort dans les foyers à un seul adulte, le programme semble alors davantage compensateur. On observe aussi que l'effet sur la syntaxe est plus net chez les parents diplômés du bac ou plus, ou que la baisse du temps d'écran est plus marquée chez les garçons. Ces différences ne portent chacune que sur un ou deux résultats et ne dessinent pas de tendance d'ensemble, elles doivent être considérées comme exploratoires.

Limites. Ces analyses de sous-groupes sont exploratoires. Elles sont moins puissantes statistiquement que l'analyse principale (les tests d'interaction exigent des échantillons plus grands), et le grand nombre de comparaisons réalisées (une vingtaine de résultats croisés avec dix dimensions) implique que quelques différences significatives soient estimées par le seul fait du hasard. Elles servent donc davantage à formuler des hypothèses qu'à établir des conclusions fermes. Le message robuste qui s'en dégage est double : un effet largement homogène entre les familles, et une intensité croissante de l'impact avec l'âge de l'enfant.

5. Ce que disent les parents bénéficiaires : analyse des retours qualitatifs

À la fin de l'accompagnement, les parents bénéficiaires ont répondu à une question ouverte sur ce que le programme a changé pour eux (224 réponses). Cette analyse est descriptive et exploratoire : elle porte seulement sur les bénéficiaires (ceux qui ont répondu au questionnaire à la fin de leur accompagnement 1001mots) et reflète des perceptions et non un impact causal. Elle éclaire néanmoins les mécanismes perçus par les familles, dont le récit est majoritairement très positif.

Figure 7 : thèmes abordés spontanément par les parents quand on leur demande leur avis sur le programme 1001mots



Le programme agirait d'abord par les livres (cités par 66 % des parents), les conseils et l'apprentissage du rôle parental (52 %) ainsi que l'attention au langage de l'enfant (52 %), relayés par les contenus (SMS/vidéos) et les appels d'accompagnement parental. Une dimension peu visible dans les chiffres mesurés précédemment ressort nettement : la confiance et la réassurance parentale. De façon cohérente, les parents d'un premier enfant et de tout-petits s'appuient davantage sur cette dimension relationnelle.

« Quand on est maman, on est un peu esseulée. [...] L'accompagnante m'a confirmé que ce que je faisais était bien. Je culpabilisais en voyant les parents sur les réseaux... elle m'a dit que lui parler et vivre avec lui, ça suffit pour qu'il se développe. »

— enfant de 18 mois

« Dès qu'il voit des livres, il va les chercher maintenant. Et je lui lis. [...] Il s'intéresse plus aux livres grâce à 1001mots. »

— enfant de 26 mois

« Ca nous a beaucoup aidé. Surtout avec les livres. Pour moi, c'était comme si j'avais des défis. Ca m'a appris à prendre du temps pour faire comprendre des choses à ma fille, se concentrer sur le langage, ça nous a rapprochées. »

— enfant de 28 mois

Les écrans : une baisse par substitution

Les écrans sont rarement cités spontanément (8,5 % des réponses), alors que les chiffres montrent une baisse marquée. Cela est lié au fait que le rappel spontané utilisé dans une question ouverte privilégie les ajouts de pratiques (« j'ai commencé à lire ») plutôt que les retraits, et la baisse d'exposition aux écrans apparaît surtout comme la conséquence de l'utilisation des livres et des activités qui prennent leur place :

« Elle regarde moins la TV et je lui donne plus souvent un livre. Quand elle était énervée, je la mettais devant la TV ; maintenant, je privilégie le langage et la communication. » — 16 mois

6. Limites de l'évaluation

Cette évaluation est robuste sur de nombreux plans, mais trois limites méritent d'être expliquées.

a. On ne compare pas des familles tirées au sort

Imaginons que les familles qui rejoignent le programme soient, dès le départ, plus attentives à l'éveil de leur enfant que les autres. Si leur enfant progresse davantage, comment être sûr que c'est grâce au programme, et non parce que ces familles étaient déjà différentes au départ ? C'est la difficulté centrale de toute évaluation où les participants ne sont pas tirés au sort. Nous y répondons de trois manières :

- **En comparant des familles cibles qui ont un niveau d'engagement similaire** : le groupe de comparaison n'est pas constitué de familles quelconques, mais de futurs bénéficiaires eux aussi engagés. On compare donc des familles qui se ressemblent sur ce qui compte le plus mais qu'on mesure mal : la motivation.
- **En tenant compte de tout ce qu'on peut mesurer** : revenu, diplôme, emploi, langue, mode de garde, âge et sexe de l'enfant... le tout avec quatre méthodes statistiques différentes. Le fait qu'elles donnent le même résultat est rassurant.

- **En mesurant à quel point un facteur caché pourrait tout remettre en cause** : un test dédié (les « bornes d'Oster »³) montre qu'il faudrait un facteur totalement oublié qui serait 5 à 8 fois plus déterminant que tout ce que nous avons déjà mesuré réuni pour annuler l'effet du programme sur le développement des enfants : un scénario très peu vraisemblable (voir l'annexe B).

Ces éléments réduisent fortement le risque de biais de sélection et améliore la crédibilité de la comparaison et donc de l'évaluation, sans pouvoir l'éliminer totalement : c'est la limite inhérente à ce type d'étude.

b. Les informations sont déclarées par les parents

Le langage des enfants a été mesuré à l'aide d'un questionnaire scientifiquement validé, l'Inventaire du Développement de l'Enfant (IDE) dans sa version Langage, conçu pour évaluer le développement langagier de façon standardisée et aussi objective que possible. Le recours à cet outil validé limite déjà nettement la part de subjectivité et constitue une garantie importante de la qualité de la mesure.

Il reste néanmoins que les pratiques parentales (lecture, écrans...) comme les items de l'IDE sont renseignés par les parents eux-mêmes, auprès de l'organisme qui délivre le programme. Une famille accompagnée peut donc, en toute bonne foi, avoir tendance à présenter les choses sous un jour un peu plus favorable, de sorte que l'impact estimé dans la section 4 puisse comprendre en partie ce biais conscient ou non. L'outil validé réduit ce risque, mais ne le supprime pas totalement.

c. Un effet de cohorte lié à la temporalité du recrutement

Une troisième limite tient à la temporalité du recrutement des deux groupes. Comme l'illustre la figure 7, les familles bénéficiaires (groupe traité) et les familles du groupe de comparaison ne sont pas recrutées aux mêmes périodes : le recrutement est largement séquentiel. Les familles traitées ont majoritairement rejoint le programme entre décembre 2023 et mars 2025, tandis que les familles de comparaison ont été recrutées entre avril 2025 et avril 2026, soit un décalage d'environ dix-sept mois entre les cohortes médianes des deux groupes.

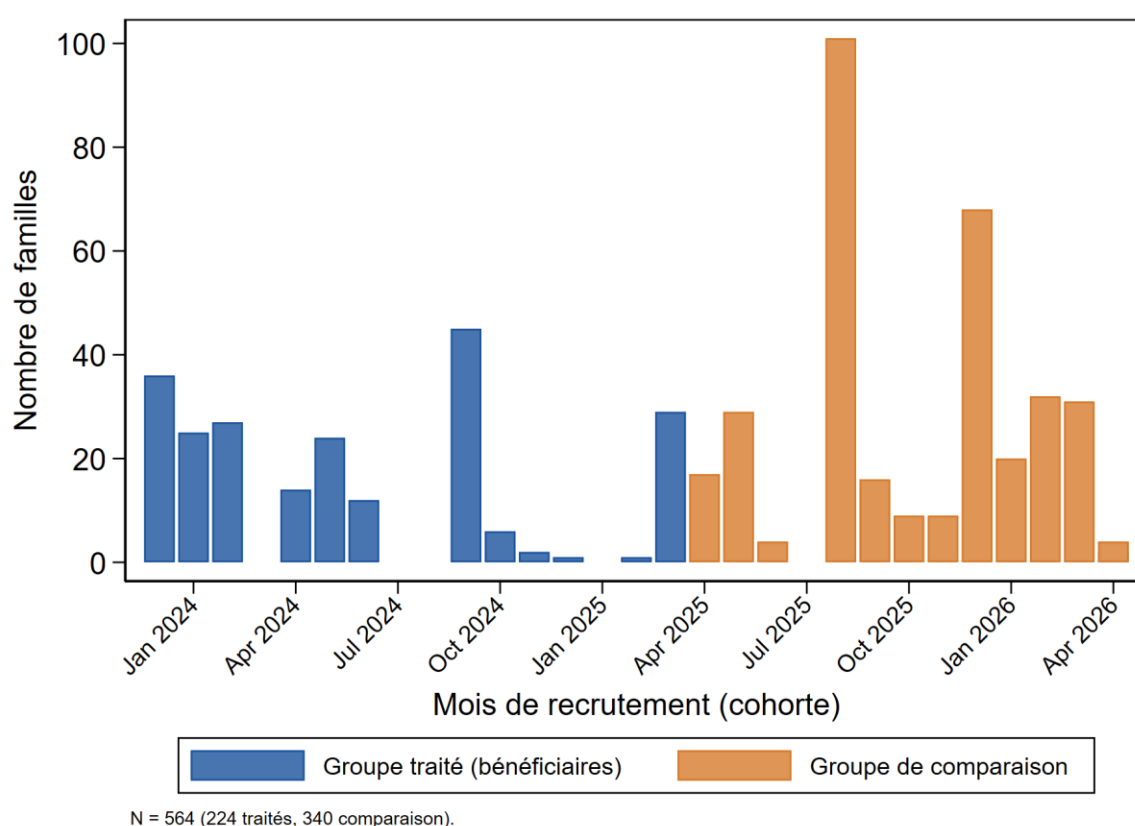
Ce décalage est en partie inhérent au design d'évaluation : le groupe traité est composé d'anciens bénéficiaires évalués après leur accompagnement, tandis que le groupe de comparaison est constitué de futures familles bénéficiaires, évaluées avant le leur. Il peut néanmoins induire un biais de cohorte si les familles recrutées à une période diffèrent de celles recrutées à une autre, que ce soit dans leurs caractéristiques (composition du public atteint) ou dans leur environnement

³ Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of business & economic statistics*, 37(2), 187-204.

(conjoncture, saisonnalité de la mesure du développement, évolutions du programme).

Deux éléments limitent toutefois la portée de cette problématique. D'une part, le processus de recrutement (canaux d'inscription, critères d'éligibilité, modalités d'inscription) est resté identique d'une année sur l'autre, ce qui réduit le risque d'un changement de composition entre les cohortes. D'autre part, l'analyse ne repose pas sur une comparaison à date calendaire identique mais sur une comparaison à âge de l'enfant donné, et l'appariement équilibre les caractéristiques observables des deux groupes (cf. tableau 5). Le risque résiduel se limite donc à d'éventuelles tendances temporelles portées par des facteurs inobservables corrélés au moment du recrutement, que la méthode ne peut entièrement exclure.

Figure 8 : temporalité de recrutement entre les cohortes bénéficiaires et de comparaison



Conclusion

Au terme de cette évaluation, les résultats dessinent un bilan cohérent et convergent. Le programme 1001mots agit d'abord sur les leviers qu'il vise : les parents bénéficiaires font évoluer leurs croyances et leurs pratiques - en particulier autour de la **lecture partagée** dès le plus jeune âge - et réduisent nettement l'**exposition de leurs enfants aux écrans**. Ces évolutions de l'environnement familial

s'accompagnent d'un meilleur **développement langagier et socio-émotionnel des enfants**.

Ces effets sont estimés de façon concordante par les quatre méthodes mobilisées (régression multivariée, appariement, repondération) et résistent à un test de sensibilité aux facteurs inobservables. Ils doivent se lire à la lumière des limites rappelées plus haut (absence de tirage au sort, mesures déclaratives, effet de cohorte) mais le faisceau d'indices concordants conforte l'idée que 1001mots produit des effets réels et positifs sur l'environnement langagier des jeunes enfants et sur leur développement.

Annexe A – Tableaux d'hétérogénéité des effets

Tableau A1 : hétérogénéité en fonction de l'âge de l'enfant

Dimension	Plus jeune (A)	Plus âgé (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,06	-0,06	252/267	0,955
Âge utile : parler (mois)	-0,43	-0,33	253/264	0,867
Âge utile : livres (mois)	-3,87***	-2,90***	249/262	0,414
Importance de lire aux tout-petits (≤ 6 mois)	0,23***	0,15**	249/262	0,409
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	0,82**	1,71***	257/271	0,052*
dont : ne lit jamais	-0,07*	-0,07**	257/271	0,976
dont : lit tous les jours	0,12*	0,32***	257/271	0,025**
Jeu (fois/sem.)	-0,35	0,02	257/270	0,393
Chant (fois/sem.)	0,48	0,78**	257/270	0,544
Décrit ses actions (0/1)	0,08	0,23***	257/269	0,103
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-41,04	-184,04***	234/253	0,035**
Écrans : au moins 7h/sem. (0/1)	-0,08	-0,12**	234/253	0,595
TV : durée (min/sem.)	-36,78	-153,15***	250/268	0,041**
Téléphone : durée (min/sem.)	-7,61	-23,53*	239/257	0,448
TV : fréquence (fois/sem.)	-0,71*	-1,00***	256/272	0,598
TV laissée en fond (souvent/tjs)	-0,18***	-0,14**	256/270	0,706
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	1,03**	3,65***	257/272	0,002***
Haut risque expression	-0,02	-0,34***	257/272	0,000***
Très haut risque expression	-0,07*	-0,15***	257/272	0,193
Âge langagier compréhension (mois)	1,13***	2,59***	257/272	0,045**
Haut risque compréhension	-0,12*	-0,25***	257/272	0,176
Très haut risque compréhension	-0,07*	-0,18***	257/272	0,095*
Vocabulaire (é.-t.)	0,28***	0,60***	257/272	0,057*
Syntaxe (é.-t.)	0,21**	0,81***	257/272	0,001***
Comptines (é.-t.)	0,20	1,00***	257/272	0,001***
Socio-émotionnel (é.-t.)	0,18*	0,64***	257/272	0,007***

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 0,82 dans « Plus jeune » (A) et 1,71 dans « Plus âgé » (B), différence significative ($p = 0,052$). *Note* : Séparation à la médiane de l'âge de l'enfant (24 mois) : (A) plus jeune (sous la médiane), (B) plus âgé (médiane ou au-dessus). Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interagi, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A2 : hétérogénéité en fonction du niveau d'engagement des familles dans le programme 1001mots

Dimension	Engagement bas (A)	Engagement élevé (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,05	-0,06**	226/293	0,801
Âge utile : parler (mois)	-0,40	-0,26	227/290	0,830
Âge utile : livres (mois)	-3,15***	-3,18***	229/282	0,982
Importance de lire aux tout-petits (≤ 6 mois)	0,21***	0,15**	229/282	0,504
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	0,93***	1,38***	233/295	0,349
dont : ne lit jamais	-0,01	-0,11***	233/295	0,040**
dont : lit tous les jours	0,17**	0,22***	233/295	0,559
Jeu (fois/sem.)	-0,08	-0,22	232/295	0,730
Chant (fois/sem.)	0,35	0,97***	232/295	0,203
Décrit ses actions (0/1)	0,09	0,20***	232/294	0,255
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-147,68***	-76,62	218/269	0,282
Écrans : au moins 7h/sem. (0/1)	-0,10*	-0,10**	218/269	0,997
TV : durée (min/sem.)	-104,83***	-80,74*	230/288	0,660
Téléphone : durée (min/sem.)	-37,80**	0,90	221/275	0,062*
TV : fréquence (fois/sem.)	-0,80**	-0,83**	234/294	0,954
TV laissée en fond (souvent/tjs)	-0,06	-0,23***	231/295	0,069*
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	1,65**	2,74***	234/295	0,210
Haut risque expression	-0,15**	-0,22***	234/295	0,452
Très haut risque expression	-0,10**	-0,13***	234/295	0,702
Âge langagier compréhension (mois)	1,25**	2,20***	234/295	0,206
Haut risque compréhension	-0,11	-0,25***	234/295	0,114
Très haut risque compréhension	-0,10**	-0,15***	234/295	0,437
Vocabulaire (é.-t.)	0,28**	0,57***	234/295	0,071*
Syntaxe (é.-t.)	0,52***	0,46***	234/295	0,778
Comptines (é.-t.)	0,55***	0,65***	234/295	0,671
Socio-émotionnel (é.-t.)	0,19	0,59***	234/295	0,021**

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 0,93 dans « Engagement bas » (A) et 1,38 dans « Engagement élevé » (B), différence non significative ($p = 0,349$). *Note* : Séparation à la médiane de l'indice d'engagement (indice standardisé mesuré à l'inscription) : (A) engagement faible, (B) engagement élevé. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interagi, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A3 : hétérogénéité en fonction du niveau de revenu par unité de consommation dans le ménage

Dimension	Revenu bas (A)	Revenu élevé (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,07*	-0,08**	255/264	0,894
Âge utile : parler (mois)	-0,88	0,12	256/261	0,114
Âge utile : livres (mois)	-3,29***	-3,44***	252/259	0,894
Importance de lire aux tout-petits (≤ 6 mois)	0,15**	0,22***	252/259	0,461
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	1,45***	1,15***	263/265	0,533
dont : ne lit jamais	-0,08**	-0,07**	263/265	0,738
dont : lit tous les jours	0,22***	0,22***	263/265	0,974
Jeu (fois/sem.)	-0,13	-0,17	263/264	0,926
Chant (fois/sem.)	0,85**	0,47	262/265	0,432
Décrit ses actions (0/1)	0,07	0,24***	263/263	0,058*
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-77,66	-142,54***	240/247	0,356
Écrans : au moins 7h/sem. (0/1)	-0,11*	-0,10**	240/247	0,921
TV : durée (min/sem.)	-78,99*	-105,29***	256/262	0,650
Téléphone : durée (min/sem.)	1,49	-32,45**	245/251	0,119
TV : fréquence (fois/sem.)	-0,70*	-1,03***	262/266	0,537
TV laissée en fond (souvent/tjs)	-0,16**	-0,15**	261/265	0,895
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	1,68***	2,63***	263/266	0,262
Haut risque expression	-0,11*	-0,24***	263/266	0,148
Très haut risque expression	-0,09*	-0,14***	263/266	0,456
Âge langagier compréhension (mois)	1,08**	2,37***	263/266	0,077*
Haut risque compréhension	-0,11	-0,25***	263/266	0,103
Très haut risque compréhension	-0,07	-0,18***	263/266	0,103
Vocabulaire (é.-t.)	0,34***	0,47***	263/266	0,410
Syntaxe (é.-t.)	0,44***	0,51***	263/266	0,708
Comptines (é.-t.)	0,59***	0,63***	263/266	0,882
Socio-émotionnel (é.-t.)	0,29**	0,49***	263/266	0,254

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 1,45 dans « Revenu bas » (A) et 1,15 dans « Revenu élevé » (B), différence non significative ($p = 0,533$). *Note* : Séparation à la médiane du revenu par unité de consommation (981 €/mois) : (A) revenu bas, (B) revenu élevé. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interactif, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A4 : hétérogénéité en fonction du niveau de diplôme le plus élevé

Dimension	Sous bac (A)	Bac et plus (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,10**	-0,06*	127/392	0,438
Âge utile : parler (mois)	-1,04	-0,16	123/394	0,277
Âge utile : livres (mois)	-3,82***	-3,24***	128/383	0,713
Importance de lire aux tout-petits (≤ 6 mois)	0,16	0,20***	128/383	0,703
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	1,44***	1,23***	129/399	0,709
dont : ne lit jamais	-0,11*	-0,06**	129/399	0,480
dont : lit tous les jours	0,25***	0,20***	129/399	0,657
Jeu (fois/sem.)	0,07	-0,20	129/398	0,634
Chant (fois/sem.)	0,60	0,71***	128/399	0,862
Décrit ses actions (0/1)	0,16	0,16***	129/397	0,984
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-145,58***	-104,75**	119/368	0,531
Écrans : au moins 7h/sem. (0/1)	-0,19***	-0,08*	119/368	0,194
TV : durée (min/sem.)	-145,44***	-83,27**	126/392	0,253
Téléphone : durée (min/sem.)	-18,99	-16,53	121/375	0,909
TV : fréquence (fois/sem.)	-1,11*	-0,81***	129/399	0,649
TV laissée en fond (souvent/tjs)	-0,33***	-0,11**	128/398	0,029**
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	1,45*	2,51***	129/400	0,263
Haut risque expression	-0,14	-0,21***	129/400	0,483
Très haut risque expression	-0,18**	-0,10***	129/400	0,346
Âge langagier compréhension (mois)	1,44**	1,86***	129/400	0,589
Haut risque compréhension	-0,06	-0,22***	129/400	0,094*
Très haut risque compréhension	-0,12	-0,13***	129/400	0,958
Vocabulaire (é.-t.)	0,28*	0,40***	129/400	0,281
Syntaxe (é.-t.)	0,17	0,58***	129/400	0,025**
Comptines (é.-t.)	0,60***	0,62***	129/400	0,934
Socio-émotionnel (é.-t.)	0,28*	0,44***	129/400	0,399

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 1,44 dans « Sous bac » (A) et 1,23 dans « Bac et plus » (B), différence non significative ($p = 0,709$). *Note* : Séparation selon le diplôme du répondant : (A) sous le bac, (B) bac ou plus. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interactif, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A5 : hétérogénéité en fonction du nombre de langues parlées dans le ménage

Dimension	Français seul (A)	Bilingue (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,10*	-0,06*	117/402	0,511
Âge utile : parler (mois)	-0,60	-0,31	116/401	0,617
Âge utile : livres (mois)	-2,38**	-3,74***	112/399	0,276
Importance de lire aux tout-petits (≤ 6 mois)	0,16	0,21***	112/399	0,645
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	0,60	1,42***	117/411	0,156
dont : ne lit jamais	-0,05	-0,08***	117/411	0,641
dont : lit tous les jours	0,15	0,23***	117/411	0,459
Jeu (fois/sem.)	0,64	-0,30	116/411	0,084*
Chant (fois/sem.)	0,50	0,74***	116/411	0,694
Décrit ses actions (0/1)	0,19*	0,16***	115/411	0,773
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-108,91*	-112,63***	107/380	0,959
Écrans : au moins 7h/sem. (0/1)	-0,11*	-0,11**	107/380	0,927
TV : durée (min/sem.)	-61,86	-94,26***	114/404	0,567
Téléphone : durée (min/sem.)	-42,95*	-13,95	110/386	0,274
TV : fréquence (fois/sem.)	-0,69	-0,81***	117/411	0,860
TV laissée en fond (souvent/tjs)	-0,13	-0,15***	117/409	0,807
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	2,47***	2,41***	117/412	0,959
Haut risque expression	-0,31***	-0,17***	117/412	0,197
Très haut risque expression	-0,22***	-0,10***	117/412	0,145
Âge langagier compréhension (mois)	1,46*	2,06***	117/412	0,500
Haut risque compréhension	-0,27***	-0,18***	117/412	0,426
Très haut risque compréhension	-0,17**	-0,13***	117/412	0,673
Vocabulaire (é.-t.)	0,59***	0,43***	117/412	0,370
Syntaxe (é.-t.)	0,59***	0,51***	117/412	0,713
Comptines (é.-t.)	0,69**	0,63***	117/412	0,863
Socio-émotionnel (é.-t.)	0,24	0,48***	117/412	0,291

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 0,60 dans « Français seul » (A) et 1,42 dans « Bilingue » (B), différence non significative ($p = 0,156$). *Note* : Séparation selon les langues parlées à l'enfant : (A) français seul, (B) au moins deux langues. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interactif, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A6 : hétérogénéité en fonction du statut parental

Dimension	1 adulte (A)	2 adultes (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,01	-0,07**	152/367	0,369
Âge utile : parler (mois)	-0,92**	-0,25	154/363	0,207
Âge utile : livres (mois)	-1,80**	-3,80***	152/359	0,047**
Importance de lire aux tout-petits (≤ 6 mois)	0,18**	0,19***	152/359	0,902
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	0,73	1,31***	158/370	0,327
dont : ne lit jamais	-0,08	-0,07**	158/370	0,894
dont : lit tous les jours	0,11	0,22***	158/370	0,359
Jeu (fois/sem.)	0,53	-0,32	158/369	0,107
Chant (fois/sem.)	0,93*	0,61**	157/370	0,567
Décrit ses actions (0/1)	0,31***	0,12**	157/369	0,083*
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-126,29**	-120,97***	143/344	0,939
Écrans : au moins 7h/sem. (0/1)	-0,10	-0,11**	143/344	0,927
TV : durée (min/sem.)	-75,52	-105,39***	153/365	0,606
Téléphone : durée (min/sem.)	-35,37**	-12,40	148/348	0,250
TV : fréquence (fois/sem.)	-1,38**	-0,83***	157/371	0,411
TV laissée en fond (souvent/tjs)	-0,30***	-0,12**	157/369	0,072*
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	3,34***	1,94***	158/371	0,160
Haut risque expression	-0,26***	-0,16***	158/371	0,347
Très haut risque expression	-0,15***	-0,11***	158/371	0,443
Âge langagier compréhension (mois)	3,09***	1,45***	158/371	0,072*
Haut risque compréhension	-0,36***	-0,14***	158/371	0,025**
Très haut risque compréhension	-0,16***	-0,12***	158/371	0,579
Vocabulaire (é.-t.)	0,46***	0,42***	158/371	0,821
Syntaxe (é.-t.)	0,69***	0,44***	158/371	0,254
Comptines (é.-t.)	0,95***	0,53***	158/371	0,217
Socio-émotionnel (é.-t.)	0,38**	0,41***	158/371	0,866

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 0,73 dans « 1 adulte » (A) et 1,31 dans « 2 adultes » (B), différence non significative ($p = 0,327$). *Note* : Séparation selon la structure du foyer : (A) un adulte, (B) deux adultes. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interactif, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A7 : hétérogénéité en fonction de la structure de la famille

Dimension	Enfant unique (A)	Fratrie (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,10**	-0,05*	148/371	0,394
Âge utile : parler (mois)	0,70	-0,57*	146/371	0,132
Âge utile : livres (mois)	-2,51**	-3,54***	147/364	0,441
Importance de lire aux tout-petits (≤ 6 mois)	0,16*	0,19***	147/364	0,735
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	1,25***	1,18***	149/379	0,900
dont : ne lit jamais	-0,11**	-0,05*	149/379	0,348
dont : lit tous les jours	0,14	0,22***	149/379	0,411
Jeu (fois/sem.)	0,50	-0,43*	149/378	0,059*
Chant (fois/sem.)	0,06	0,88***	150/377	0,144
Décrit ses actions (0/1)	0,11	0,15***	150/376	0,679
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-183,82***	-97,32**	144/343	0,246
Écrans : au moins 7h/sem. (0/1)	-0,17**	-0,09**	144/343	0,369
TV : durée (min/sem.)	-145,31***	-90,75***	148/370	0,386
Téléphone : durée (min/sem.)	-34,62*	-8,53	146/350	0,275
TV : fréquence (fois/sem.)	-0,68	-1,02***	150/378	0,591
TV laissée en fond (souvent/tjs)	-0,23***	-0,14***	148/378	0,382
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	1,27*	2,41***	150/379	0,195
Haut risque expression	-0,12	-0,20***	150/379	0,459
Très haut risque expression	-0,00	-0,16***	150/379	0,021**
Âge langagier compréhension (mois)	2,14***	1,52***	150/379	0,444
Haut risque compréhension	-0,23***	-0,15***	150/379	0,454
Très haut risque compréhension	-0,09	-0,14***	150/379	0,522
Vocabulaire (é.-t.)	0,34**	0,41***	150/379	0,644
Syntaxe (é.-t.)	0,27*	0,53***	150/379	0,146
Comptines (é.-t.)	0,55**	0,66***	150/379	0,681
Socio-émotionnel (é.-t.)	0,48***	0,35***	150/379	0,480

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 1,25 dans « Enfant unique » (A) et 1,18 dans « Fratrie » (B), différence non significative ($p = 0,900$). *Note* : Séparation selon la fratrie : (A) enfant unique, (B) au moins un frère ou une sœur. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interagi, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A8 : hétérogénéité en fonction du lieu de naissance (France / à l'étranger)

Dimension	Né à l'étranger (A)	Né en France (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,10**	-0,03	293/226	0,227
Âge utile : parler (mois)	-0,58	-0,15	293/224	0,419
Âge utile : livres (mois)	-3,50***	-2,98***	292/219	0,661
Importance de lire aux tout-petits (≤ 6 mois)	0,20***	0,17**	292/219	0,796
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	1,72***	0,46	300/228	0,009***
dont : ne lit jamais	-0,12***	0,01	300/228	0,007***
dont : lit tous les jours	0,26***	0,12	300/228	0,124
Jeu (fois/sem.)	-0,33	-0,07	300/227	0,556
Chant (fois/sem.)	0,74**	0,47	299/228	0,587
Décrit ses actions (0/1)	0,13**	0,19**	300/226	0,514
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-120,51**	-108,03***	276/211	0,853
Écrans : au moins 7h/sem. (0/1)	-0,12**	-0,10*	276/211	0,752
TV : durée (min/sem.)	-107,41**	-85,22***	293/225	0,687
Téléphone : durée (min/sem.)	-9,55	-24,94*	281/215	0,450
TV : fréquence (fois/sem.)	-0,91**	-0,74*	299/229	0,752
TV laissée en fond (souvent/tjs)	-0,16***	-0,16**	298/228	0,949
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	2,16***	2,57***	300/229	0,632
Haut risque expression	-0,19***	-0,18***	300/229	0,953
Très haut risque expression	-0,13***	-0,12***	300/229	0,972
Âge langagier compréhension (mois)	1,82***	1,84***	300/229	0,976
Haut risque compréhension	-0,20***	-0,17**	300/229	0,772
Très haut risque compréhension	-0,11**	-0,16***	300/229	0,416
Vocabulaire (é.-t.)	0,42***	0,47***	300/229	0,777
Syntaxe (é.-t.)	0,43***	0,61***	300/229	0,328
Comptines (é.-t.)	0,57***	0,78***	300/229	0,392
Socio-émotionnel (é.-t.)	0,46***	0,34***	300/229	0,523

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 1,72 dans « Né à l'étranger » (A) et 0,46 dans « Né en France » (B), différence significative ($p = 0,009$). *Note* : Séparation selon le pays de naissance du parent répondant : (A) né à l'étranger, (B) né en France. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interagi, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A9 : hétérogénéité en fonction du moment de l'arrivée en France

Dimension	Arrivée récente (A)	Arrivée ancienne (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,17**	-0,04	106/184	0,144
Âge utile : parler (mois)	-1,13**	-0,37	109/181	0,353
Âge utile : livres (mois)	-5,43***	-2,95***	108/182	0,105
Importance de lire aux tout-petits (≤ 6 mois)	0,39***	0,09	108/182	0,022**
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	1,95***	1,84***	109/188	0,854
dont : ne lit jamais	-0,15***	-0,09**	109/188	0,445
dont : lit tous les jours	0,27***	0,32***	109/188	0,683
Jeu (fois/sem.)	0,22	-0,37	109/188	0,294
Chant (fois/sem.)	0,76	0,86*	109/187	0,881
Décrit ses actions (0/1)	0,14	0,14*	109/188	0,999
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-90,77	-148,81**	102/172	0,588
Écrans : au moins 7h/sem. (0/1)	-0,13	-0,11	102/172	0,867
TV : durée (min/sem.)	-87,07	-131,95***	107/183	0,621
Téléphone : durée (min/sem.)	-4,67	-12,98	104/175	0,783
TV : fréquence (fois/sem.)	-0,60	-1,16**	108/188	0,435
TV laissée en fond (souvent/tjs)	-0,22**	-0,14*	108/187	0,452
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	1,25	2,61***	109/188	0,249
Haut risque expression	-0,11	-0,23***	109/188	0,349
Très haut risque expression	-0,05	-0,16***	109/188	0,240
Âge langagier compréhension (mois)	1,73**	1,88***	109/188	0,879
Haut risque compréhension	-0,20**	-0,22***	109/188	0,868
Très haut risque compréhension	-0,10	-0,12**	109/188	0,902
Vocabulaire (é.-t.)	0,29*	0,49***	109/188	0,393
Syntaxe (é.-t.)	0,17	0,56***	109/188	0,118
Comptines (é.-t.)	0,47*	0,61***	109/188	0,678
Socio-émotionnel (é.-t.)	0,35*	0,46***	109/188	0,631

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 1,95 dans « Arrivée récente » (A) et 1,84 dans « Arrivée ancienne » (B), différence non significative ($p = 0,854$). *Note* : Sous-échantillon des parents nés à l'étranger (l'ancienneté d'arrivée n'est renseignée que pour eux). Séparation à la médiane de l'ancienneté d'arrivée en France (tranches : moins de 5, 5-10, 11-20, plus de 20 ans) : (A) arrivée plus récente, (B) arrivée plus ancienne. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interagi, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Tableau A10 : hétérogénéité en fonction du sexe de l'enfant

Dimension	Garçon (A)	Fille (B)	N (A/B)	Diff. (p)
<i>Croyances</i>				
Écrans aident à parler (0/1)	-0,06*	-0,07	283/236	0,810
Âge utile : parler (mois)	-0,04	-0,82**	286/231	0,170
Âge utile : livres (mois)	-3,37***	-3,32***	278/233	0,967
Importance de lire aux tout-petits (≤ 6 mois)	0,27***	0,12*	278/233	0,111
<i>Pratiques de stimulation</i>				
Lecture (fois/sem.)	1,15***	1,38***	287/241	0,616
dont : ne lit jamais	-0,03	-0,09***	287/241	0,211
dont : lit tous les jours	0,21***	0,24***	287/241	0,741
Jeu (fois/sem.)	0,08	-0,40	287/240	0,265
Chant (fois/sem.)	0,42	0,90***	287/240	0,326
Décrit ses actions (0/1)	0,22***	0,10	286/240	0,237
<i>Exposition aux écrans</i>				
Écrans : durée totale (min/sem.)	-173,04***	-50,10	266/221	0,076*
Écrans : au moins 7h/sem. (0/1)	-0,17***	-0,05	266/221	0,122
TV : durée (min/sem.)	-134,50***	-48,61	283/235	0,136
Téléphone : durée (min/sem.)	-33,75***	-3,35	270/226	0,155
TV : fréquence (fois/sem.)	-0,98**	-0,72*	286/242	0,624
TV laissée en fond (souvent/tjs)	-0,09	-0,24***	285/241	0,088*
<i>Développement de l'enfant</i>				
Âge langagier expression (mois)	2,30***	2,11***	287/242	0,826
Haut risque expression	-0,20***	-0,18***	287/242	0,825
Très haut risque expression	-0,12**	-0,10***	287/242	0,731
Âge langagier compréhension (mois)	1,89***	1,57***	287/242	0,666
Haut risque compréhension	-0,24***	-0,12*	287/242	0,195
Très haut risque compréhension	-0,13***	-0,12***	287/242	0,793
Vocabulaire (é.-t.)	0,46***	0,37***	287/242	0,614
Syntaxe (é.-t.)	0,46***	0,52***	287/242	0,727
Comptines (é.-t.)	0,61***	0,59***	287/242	0,936
Socio-émotionnel (é.-t.)	0,35***	0,46***	287/242	0,525

Lecture : chaque ligne donne l'effet du programme séparément dans les deux sous-groupes (A) et (B), et la colonne « Diff. (p) » teste si cet effet diffère entre eux. Par exemple, l'effet sur la fréquence de lecture est de 1,15 dans « Garçon » (A) et 1,38 dans « Fille » (B), différence non significative ($p = 0,616$). *Note* : Séparation selon le sexe de l'enfant : (A) garçon, (B) fille. Effet du traitement estimé dans chaque sous-groupe (OLS entièrement interagi, contrôles étendus). La colonne « Diff. (p) » donne la p-value du test d'interaction entre le traitement et la dimension. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Annexe B - Impact du programme sur chaque item de l'IDE

Le tableau ci-dessous présente, pour chaque item de l'Inventaire du Développement de l'Enfant version Langage, la différence entre la part d'enfants dont le parent déclare la capacité acquise (groupe traité moins groupe de comparaison, en points de pourcentage).

Item	Description	% Contr.	% Trait.	Diff. (pts)	N
C1	Babillage (actuel ou passé)	98,5	98,6	+0,1	561
C2	Montre des objets	96,5	97,8	+1,3	563
C3	Répond à son nom	99,1	99,1	-0,0	563
C4	Imite des sons	95,3	99,1	+3,8**	563
C5	Vient quand on l'appelle (la plupart du temps)	97,6	97,8	+0,1	563
C6	Fait au revoir de la main	99,1	98,7	-0,5	564
C7	Comprend « Non, non » (arrête un moment)	97,9	97,3	-0,6	564
C8	Dit « maman »/« papa » (ou équivalent)	97,4	100,0	+2,6**	563
C9	Dit ≥2 autres mots (hors maman/papa)	93,5	96,9	+3,4*	560
C10	Dit ≥10 mots	68,8	80,7	+11,9***	563
C11	Dit ≥5 noms d'objets usuels	51,2	64,3	+13,1***	564
C12	Utilise « mon/ma/le mien/la mienne »	15,9	26,3	+10,4***	563
C13	Demande plus (« plus », « encore », ...)	50,0	63,8	+13,8***	562
C14	Demande boire/manger avec mots ou sons	75,8	85,7	+9,8***	562
C15	Fait des phrases de 2-3 mots ou plus	42,4	61,2	+18,8***	564
C16	Nomme des objets dans un livre	44,1	64,7	+20,6***	564
C17	Montre ≥3 parties du corps sur demande	74,4	84,2	+9,8***	557
C18	Comprend « mettre » vs « enlever »	89,3	92,0	+2,6	562
C19	Répond correctement par oui/non	56,6	66,2	+9,6**	561
C20	Comprend des demandes simples	94,7	96,4	+1,7	563
C21	Suit des consignes en 2 parties	55,9	67,3	+11,4***	546
C22	Comprend « ouvre/ferme » et fait le geste	79,5	86,9	+7,5**	558
C23	Donne un jouet sur demande	92,6	96,4	+3,8*	564
C24	Raconte ce qui se passe dans un livre	15,3	28,1	+12,8***	564
C25	Utilise des phrases négatives	22,6	34,5	+11,9***	563
C26	Fait des phrases avec ≥4 mots	15,0	31,4	+16,4***	563
C27	Utilise au moins un pronom (moi/je/il/elle/tu/ça)	45,0	56,7	+11,7***	564
C28	Dit ≥20 mots	39,2	62,9	+23,7***	563
C29	Sait chuchoter	57,4	60,6	+3,2	557
C30	Utilise « s'il te plaît » et « merci » à bon escient	27,6	40,8	+13,2***	563
C31	Nomme ≥5 parties du corps	32,2	50,4	+18,3***	563
C32	Comprend « en haut/bas »	41,4	51,4	+10,0**	538
C33	Dit « grand » et « petit »	13,4	19,2	+5,8*	560
C34	Dit au moins son prénom	29,5	43,9	+14,4***	560
C35	Comprend ≥3 mots de position (dans/sur/sous/à côté ...)	27,2	39,2	+11,9***	551
C36	Dit quand quelque chose est lourd	22,6	33,5	+10,9***	558
C37	Répond aux sons d'animaux (chat/chien/canard ...)	43,0	64,9	+21,8***	559
C38	Exprime avec des mots ce qu'il aime/n'aime pas	14,4	23,2	+8,8***	564
C39	Utilise le mot « tu »	7,2	13,8	+6,6**	473
C40	Décrit des objets avec détails	7,6	18,3	+10,7***	475
C41	Dit « j'sais/peux/veux pas »	28,7	41,1	+12,4***	476
C42	Utilise articles/partitifs (« un/une/de/d'/le/la »)	11,8	20,5	+8,7***	474
C43	Utilise correctement « moi/mon/je »	13,6	19,3	+5,7*	476
C44	Chante des chansons (paroles articulées)	19,9	40,6	+20,8***	474
C45	Pose des questions « Que/Quoi/Quel/Où »	23,3	28,9	+5,6	476
C46	Donne des raisons avec « parce que »	1,4	5,1	+3,6**	476
C47	Parole claire (compréhensible la plupart du temps)	28,0	43,7	+15,7***	476
C48	Pose des questions « pourquoi/quand/comment »	5,7	14,7	+9,0***	476
C49	Dit ≥50 mots dans la vie quotidienne	23,7	39,6	+15,9***	475
C50	Pose des questions simples sans erreur	13,6	24,4	+10,7***	476
C51	Ordonne correctement les mots dans la phrase	25,2	41,1	+15,9***	475
C52	Utilise « garçon/fille » correctement	8,6	19,5	+10,9***	473
C53	Répond : que fait-on avec (biscuit/chapeau/verre...)	8,7	15,9	+7,2**	470
C54	Répond : que fais-tu quand (soif/faim/fatigué...)?	6,2	11,9	+5,6**	466
C55	Utilise « vite » et « lent/doucement » correctement	10,4	14,7	+4,3	475
C56	Comprend/utilise « plein » et « vide »	6,5	10,7	+4,2	474
C57	Dit si un bruit est fort ou pas	16,9	21,8	+4,9	475
C58	Dit nom et prénom sur demande	5,8	11,2	+5,5**	474
C59	Utilise « gentil »/« méchant » pour décrire	25,8	29,9	+4,1	476
C60	Comprend ≥6 mots de position	12,0	20,1	+8,1**	468
C61	Parle de ses émotions (0/1)	11,5	21,8	+10,4***	476
C62	Identifie correctement au moins une couleur	23,7	34,9	+11,2***	474
C63	Décrit avec adjectifs (petit/rouge/bon/rigolo ...)	12,2	20,3	+8,1**	476
C64	Fait des phrases de ≥6 mots	10,1	22,1	+11,9***	284
C65	Nomme des formes simples (rond/carré/triangle/étoile)	13,5	23,7	+10,2**	283
C66	Utilise le pluriel avec « les/des » correctement	8,2	14,8	+6,6*	281
C67	Récite des comptines (ex. 1,2,3, ...)	10,8	32,4	+21,5***	284
C68	Lie 2 phrases avec « et/ou/mais »	6,1	10,3	+4,2	284
C69	A beaucoup de vocabulaire (impression générale)	25,2	41,9	+16,7***	283
C70	Prononce correctement la plupart des mots	29,7	32,4	+2,6	284
C71	Répond au « pourquoi » avec explications	8,8	10,4	+1,7	282
C72	Définit/décrit « qu'est-ce qu'un...? »	8,2	10,4	+2,3	281
C73	Reconnait correctement ≥4 couleurs	14,2	22,8	+8,6*	284
C74	Dit correctement son âge	23,1	30,1	+7,0	283
C75	Comprend/utilise « facile » et « difficile/dur »	4,8	7,4	+2,6	282
C76	Participe aux conversations (0/1)	28,6	46,3	+17,8***	283
C77	Répète des histoires courtes dans l'ordre	5,3	7,6	+2,3	167
C78	Utilise correctement le passé (imparfait/voulions ...)	4,0	8,7	+4,7	167
C79	Utilise structures « Si..., alors... / Quand..., alors... »	2,7	4,3	+1,6	168
C80	Utilise pluriels irréguliers (chevaux/animaux ...)	2,7	2,2	-0,5	167

Item	Description	% Contr.	% Trait.	Diff. (pts)	N
C81	Évoque ce qui pourrait arriver (anticipation)	8,0	13,0	+5,0	167
C82	Exprime ses idées avec phrases complètes et claires	14,7	10,9	-3,8	167
C83	Fait des phrases longues/complexes (≥ 10 mots)	5,3	10,8	+5,4	168
C84	Raconte en détail des événements passés	1,3	5,4	+4,1	167
C85	Utilise pronoms/pluriels « nous/vous/ils/leurs/eux »	6,7	4,3	-2,3	167
C86	Demande la signification des mots	13,3	10,9	-2,5	167
C87	Dit les jours de la semaine dans l'ordre	0,0	3,2	+3,2	168
C88	Utilise « aujourd'hui/hier/demain » correctement	4,0	2,2	-1,8	167
C89	Montre droite/gauche correctement	10,8	22,8	+12,0**	166
C90	Comprend « pareil/différent » et justifie	2,7	2,2	-0,5	168
C91	Répond : que fais-tu avec (yeux/oreilles ...) ?	8,0	8,8	+0,8	166
C92	Répond aux questions « Si..., alors... ? »	5,3	5,4	+0,0	168
C93	Sais la catégorie des objets ?	5,3	5,4	+0,1	167
C94	Utilise superlatifs « le plus fort/meilleur/plus grand »	8,0	6,7	-1,3	165
C95	Dit en quoi sont faits certains objets	1,3	0,0	-1,3	168
C96	Dit le nom de sa ville	4,0	6,6	+2,6	166
C97	Répond : à quoi sert (four/livre ...) ?	4,0	7,6	+3,6	167
C98	Suit 3 consignes séquentielles dans l'ordre	18,1	15,7	-2,3	161
C99	Comprend/utilise « avant » et « après »	5,3	10,9	+5,5	167
C100	Parle du futur (au-delà de la journée)	2,7	7,6	+4,9	167
B0	Reste concentré 5 min+ (0/1)	54,9	78,0	+23,2***	544
B20	Dit avec mots quand mécontent (0/1)	15,0	26,8	+11,7***	563
B27	S'adapte en groupe (0/1)	62,0	69,1	+7,1*	544

Lecture : par exemple, pour l'item « Dit au moins 20 mots », 39,2 % des enfants du groupe de comparaison contre 62,9 % des bénéficiaires le réussissent, soit un écart de +23,7 points (significatif à 1 %). *Note* : différence, en points de pourcentage, entre la part d'enfants dont le parent déclare la capacité acquise (groupe traité moins groupe de comparaison). Test *t* de différence de moyennes, sans contrôles. * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.