

Série ETUDES
DOCUMENT DE TRAVAIL 2021-E01
Janvier 2021

**Évaluation multidimensionnelle de l'impact
de l'utilisation d'équipements numériques mobiles
sur les apprentissages des élèves**

Premiers résultats du Plan numérique de 2015

Version du 21/12/2020

Ghazala Azmat, Denis Fougère, Clémence Lobut

Évaluation multidimensionnelle de l'impact de l'utilisation d'équipements numériques mobiles sur les apprentissages des élèves : premiers résultats du Plan numérique de 2015

Décembre 2020

Auteurs : Ghazala Azmat¹, Denis Fougère² et Clémence Lobut³

1. Département d'Économie (Sciences Po, Paris)

2. CNRS, OSC et LIEPP (Sciences Po, Paris), IZA (Bonn) et CEPR (Londres)

3. DEPP-B4 (Ministère de l'Éducation nationale, de la Jeunesse et des Sports) et Département d'Économie (Sciences Po, Paris)

Remerciements :

Nous remercions les équipes de la DEPP pour leur aide précieuse tout au long du projet, notamment Sandra Andreu, Sylvia Antognarelli, Laurène Bocognano, Axelle Charpentier, Olivier Cosnefroy, Donna Fleury, Isabelle Lavieville, Alexis Lermite, Aïcha M'Bafoumou, Mitia Oberti, Christelle Raffaëlli, Thierry Rocher, Alexandre Touw et Ronan Vourc'h. Nous remercions également Philippe Wuillamier et Fabienne Rosenwald pour leur soutien. Nous remercions les autres équipes de recherche impliquées dans ce projet, notamment les équipes de Maud Besançon (LP3C, Université Rennes 2), de Philippe Dessus (LARAC, Université Grenoble-Alpes) et d'Yves Dutercq (CREN, Université de Nantes). Nous souhaitons rendre hommage à Yves Dutercq, décédé en mars 2020. Nous sommes très reconnaissants d'avoir eu la chance de travailler à ses côtés dans le cadre de ce projet et d'avoir pu bénéficier de sa grande connaissance du système éducatif français, ainsi que de sa bienveillance et sa chaleur.

Table des matières

1	Revue de littérature	6
1.1	Allocation d'un équipement individuel à chaque élève	6
1.2	Accès à un équipement numérique à domicile	7
1.3	Accès à un équipement numérique à l'école	8
1.4	Bilan de la revue de la littérature	9
2	Description du Plan numérique	10
3	Méthodologie d'évaluation	12
4	Échantillon et temps de mesure	16
4.1	Description des établissements de l'échantillon	16
4.2	Temps de mesure	18
4.3	Description des élèves de l'échantillon	18
4.4	Description des enseignants de l'échantillon	18
5	Matériel d'enquête et d'évaluation et construction des indicateurs de résultat	22
5.1	Élèves	22
5.1.1	Questionnaire sur l'utilisation du numérique	22
5.1.2	Évaluation des compétences disciplinaires (mathématiques et français) et numériques .	23
5.1.3	Évaluation des compétences sociocognitives (esprit critique, créativité et collaboration)	24
5.2	Enseignants	25
5.2.1	Partie du questionnaire relative aux pratiques d'enseignement	25
5.2.2	Partie du questionnaire relative à la créativité et à l'esprit critique	29
5.2.3	Mesure des compétences numériques des enseignants	30
6	Taux de participation	32
7	Résultats sur l'utilisation du numérique par les élèves	35
7.1	Accès à des équipements mobiles et utilisation en 5 ^e	35
7.2	Accès, utilisation et types d'usages des équipements numériques en 4 ^e	36
8	Résultats sur les indicateurs des enseignants	38
9	Résultats sur les scores des élèves	41
9.1	Scores des élèves en 5 ^e	41
9.2	Scores des élèves en 4 ^e	41
9.3	Tests de robustesse	42
10	Conclusion	44

Introduction

Lancé en 2015, le Plan numérique avait pour ambition majeure le déploiement à l'échelle nationale de ressources pédagogiques et d'équipements numériques mobiles dans les écoles et collèges français. Il s'inscrit dans la continuité d'une politique d'équipement des établissements, initiée en 1985 avec le plan « Informatique pour tous », et alors que le taux d'équipement est en hausse constante dans les établissements scolaires : il y avait un ordinateur pour 9,9 élèves en moyenne en France en 2007, contre un ordinateur pour 3,2 élèves en 2017⁴. La place centrale du numérique éducatif est également sensible à l'échelle internationale : l'OCDE indique que 72 % des élèves de 15 ans utilisent un ordinateur portable à l'école en 2012⁵.

L'omniprésence des équipements numériques mobiles dans le milieu scolaire et les importants investissements nécessaires pour leur déploiement incitent à interroger leur impact sur les pratiques d'enseignement et sur les apprentissages des élèves. Malgré un nombre important de recherches sur cette thématique, aucun consensus n'existe concernant ces effets. Leur mesure se heurte à l'important enjeu méthodologique du biais de sélection. En effet, les établissements équipés sont rarement comparables aux établissements non-équipés. C'est notamment le cas pour les établissements équipés dans le cadre du Plan numérique, qui ont été sélectionnés par appel à projets, en fonction d'un projet pédagogique et des infrastructures techniques de l'établissement, notamment en termes de connexion Internet⁶. Par ailleurs, il y a pu y avoir la volonté d'équiper en priorité les établissements les plus défavorisés. Ces caractéristiques risquent d'être corrélées avec les variables permettant la mesure des effets et notamment avec les résultats des élèves, faussant la mesure des effets du Plan numérique.

La Direction de l'Évaluation, de la Prospective et de la Performance (DEPP) a lancé en 2017 le programme d'évaluation ELAINE (Évaluation Longitudinale des ActIvités liées au Numérique Éducatif) avec l'objectif de répondre à un ensemble d'interrogations, tout en veillant à atténuer le biais de sélection : la disponibilité d'équipements numériques mobiles a-t-elle entraîné une modification des pratiques pédagogiques ? Permet-elle une amélioration des résultats disciplinaires, en mathématiques et en français, des élèves ? De leurs compétences numériques ? De leur créativité ? De leur esprit critique ? De leurs pratiques collaboratives ? Ces effets sont-ils homogènes ou varient-ils en fonction des caractéristiques des élèves (sexe et origine sociale), de leurs enseignants (sexe et ancienneté) ou de leur établissement scolaire (composition sociale, secteur et appartenance au réseau d'éducation prioritaire) ? Grâce au suivi d'un panel d'élèves de la 5^e à l'entrée en 2^{de}, ELAINE sera également en mesure d'apprécier l'impact du numérique sur l'orientation au lycée (voies professionnelle, générale ou technologique, enseignements optionnels suivis)⁷. Enfin, elle documente

4. *Repères et Références Statistiques*, Ministère de l'Éducation nationale (2007) ; *Repères et Références Statistiques*, Ministère de l'Éducation nationale (2017).

5. OCDE (2015), *Students, Computers and Learning: Making the Connection*, PISA, Éditions OCDE, Paris.

6. Les critères de sélection concernent notamment le débit Internet, les installations électriques et réseaux (bornes Wi-Fi, serveurs, etc.) ainsi que les dispositifs collectifs de visualisation interactive (vidéoprojecteurs, tableaux numériques interactifs, etc.).

7. Le suivi à plus long terme du panel permettra également de documenter leur parcours d'études dans l'enseignement

la place des outils numériques dans les pratiques d'enseignement, l'impact sur les compétences numériques des enseignants et fait le lien avec le pilotage pédagogique dans l'établissement.

Il est important de noter que l'ensemble des données présentées dans ce document ont été collectées avant la période de fermeture des établissements scolaires de mars à mai 2020 et que les résultats ne tiennent donc pas compte des effets potentiellement induits par la mise en œuvre de l'enseignement à distance dans le cadre du dispositif de continuité pédagogique.

Le présent document porte sur les effets du dispositif en classe de 5^e et de 4^e. Il dresse une revue de la littérature permettant d'identifier les limites des études antérieures et l'apport d'ELAINE (1), avant de décrire l'intervention évaluée, le Plan numérique, et le rôle de la DEPP dans cette évaluation (2) et de détailler la méthodologie d'évaluation (3). L'échantillon et le calendrier sont décrits dans la partie 4. La partie 5 présente le matériel d'évaluation et la construction des scores et indicateurs à partir des évaluations des élèves et des enquêtes des enseignants. Les taux de participation aux deux temps de mesure sont renseignés dans la partie 6. Les parties suivantes détaillent les résultats sur l'utilisation des équipements numériques par les élèves (7), les pratiques des enseignants en lien avec les outils numériques (8) et les compétences des élèves (9). La conclusion résume les principaux messages, décrit les futures étapes et propose des pistes pour de futures recherches.

1 Revue de littérature

L'évaluation de l'impact des outils numériques sur les apprentissages des élèves a été au centre de nombreuses études depuis son introduction dans les salles de classe dans les années 70, mais c'est à partir des années 90 et de l'entrée des écoles dans l'ère numérique que la littérature sur le sujet devient plus foisonnante. Aucun consensus précis n'émerge de cette littérature, mais on peut toutefois noter que peu d'études trouvent des effets négatifs statistiquement significatifs. La majorité des études n'identifient pas d'effet ou détectent des effets positifs, généralement faibles.

La littérature sur cette thématique et portant sur des élèves du primaire ou du secondaire peut être divisée en fonction de trois objets d'étude : l'allocation d'un équipement individuel à chaque élève, la disponibilité et/ou l'utilisation d'un équipement numérique à domicile, la disponibilité et/ou l'utilisation d'un équipement numérique à l'école.

1.1 Allocation d'un équipement individuel à chaque élève

De nombreux articles se sont intéressés à l'attribution d'équipements numériques individuels aux élèves, s'appuyant pour une large partie d'entre eux sur le programme « *One Laptop per Child* » (« Un Ordinateur par Enfant »)⁸. La majorité de ces études n'identifient pas d'impact statistiquement significatif ou trouvent des résultats contrastés. Ainsi, la revue de littérature menée par Sirajul Islam et Grönlund (2016) et basée sur 145 articles de recherche à l'international conclut à un impact ambigu de ce type de politique. De même, Beuermann *et al.* (2015) et Malamud *et al.* (2019) ont mené une étude randomisée contrôlée au Pérou auprès d'élèves d'écoles primaires et n'identifient pas d'effet statistiquement significatif sur les scores de mathématiques et de maîtrise de la langue. Meza-Cordero (2017) a réalisé une étude similaire au Costa-Rica et n'identifie pas non plus d'effet sur les résultats des élèves. Enfin De Melo *et al.* (2014) ne trouvent pas d'effet de l'allocation d'équipements numériques individuels sur les scores de mathématiques et de lecture chez près de 3 000 élèves en dernière année d'école primaire en Uruguay, via une étude dont le protocole d'identification des effets repose sur la méthode des doubles différences. On peut noter que l'ensemble de ces études ont été réalisées dans des pays en développement, la possibilité d'extrapolation de ces résultats au contexte français est donc limitée. De plus, la majeure partie de ces études portent sur des élèves de primaire et donc plus jeunes que ceux concernés par l'étude ELAINE dans le 2^d degré.

Quand des études identifient des effets positifs de ce type de dispositif, ceux-ci sont généralement modérés. C'est par exemple le cas de Hull et Duch (2017) qui utilisent des données portant sur plusieurs millions d'élèves du CP à la Terminale aux États-Unis avec une méthode des doubles différences et trouvent des effets positifs modérés en mathématiques (allant de 15 à 17 % d'un

8. Ce programme est financé par une association du même nom, basée au *Massachusetts Institute of Technology* aux États-Unis et dont l'objectif est de promouvoir l'accès à des ordinateurs peu coûteux au plus grand nombre d'enfants. Les pays d'intervention se situent principalement en Amérique latine et en Afrique. Des villes américaines ont également bénéficié du programme.

écart-type à moyen terme). Les études trouvant des effets négatifs sont rares. On peut citer Mora *et al.* (2018) dont l'étude porte sur une cohorte de plus de 170 000 élèves en Catalogne et qui identifient via une régression linéaire avec effet fixe un impact négatif sur les résultats en catalan, espagnol, anglais et mathématiques.

1.2 Accès à un équipement numérique à domicile

Par ailleurs, un autre pan de cette littérature s'intéresse aux effets de l'accès à un ordinateur à domicile et de son utilisation. À l'instar de Bulman et Fairlie dans une revue de littérature (2016), on peut noter une évolution dans les résultats de la recherche dans ce domaine. Si les premières études se basant sur des variables de contrôle ou instrumentales trouvaient de forts effets positifs, les études plus récentes avec randomisation concluent généralement à des effets faibles ou nuls, voire négatifs.

Parmi les premières études sans randomisation, celle de Fairlie (2005) estime via un modèle probit multivarié que le taux de scolarisation d'adolescents de 16 à 18 ans n'étant pas encore diplômés du Secondaire augmente lorsqu'ils disposent d'un ordinateur à leur domicile. Schmitt et Wassworth (2006) concluent eux, dans une étude longitudinale, à un effet positif de la possession d'un ordinateur à domicile sur les résultats d'une cohorte d'élèves au *GSCE* (diplôme de fin d'enseignement général) et au *A-level* (examen de fin d'études secondaires) en Angleterre. En revanche, Beltran *et al.* (2006) n'identifient pas d'impact positif de la possession d'un ordinateur à la maison sur la réussite scolaire.

Les résultats d'une expérience randomisée contrôlée en Californie portant sur 1 000 élèves inscrits dans 15 écoles indiquent que disposer d'un ordinateur à domicile n'a pas d'impact sur les notes ou sur les résultats à des tests standardisés, ou encore sur la présence en classe (Fairlie et Robinson, 2013), mais a un impact positif sur les interactions sociales des élèves, notamment sur le nombre d'interactions avec des amis ou des pairs (Fairlie et Kalil, 2017). On retrouve cette variation des effets selon les indicateurs étudiés dans les résultats de Malamud et Pop-Eleches (2011) qui estiment, grâce à une régression par discontinuité, un impact négatif d'un programme de bon d'achat d'ordinateurs pour les élèves sur leurs notes, mais une amélioration de leurs compétences numériques de 20 % à 40 % d'un écart-type et de leurs compétences cognitives mesurées par les matrices progressives de Raven (des tests d'intelligence à choix multiples) d'environ 33 % d'un écart-type. Agasisti *et al.* (2017) identifient eux un impact négatif du niveau d'utilisation des équipements numériques sur les scores des élèves à PISA 2015, en restreignant l'échantillon à 15 pays européens et en implémentant un appariement statistique par score de propension. Les conclusions de Fernandez Gutierrez *et al.* (2020) sont très proches et se basent sur ces mêmes données, mais uniquement dans les régions autonomes d'Espagne et via une régression linéaire multiniveaux.

1.3 Accès à un équipement numérique à l'école

Enfin, ces mêmes disparités dans les résultats émergent de la littérature portant sur l'impact de l'utilisation du numérique dans les établissements scolaires.

À partir des données de PISA 2012, l'OCDE met en avant une relation hétérogène entre l'utilisation des équipements numériques à l'école et les apprentissages, via l'étude des corrélations. Cette relation dépend de l'intensité de l'accès et de l'utilisation des outils numériques par les élèves à l'école : si le numérique a un effet positif à des niveaux modérés d'utilisation, une utilisation trop intense peut avoir des effets négatifs. De plus, toujours selon cette enquête, l'utilisation d'ordinateurs à l'école n'a pas d'effet sur la réduction des écarts entre les résultats des élèves issus de milieux défavorisés et ceux des élèves issus de milieux favorisés.

Récemment, plusieurs études randomisées contrôlées ont porté sur l'allocation d'équipements numériques dans les écoles. Barrow *et al.* (2009) trouvent un impact positif de la mise à disposition d'une salle d'ordinateurs pour les cours de mathématiques sur les résultats à un test standardisé d'algèbre. Dans une étude menée en Colombie, Barrera-Osorio et Linden (2009) ne trouvent pas d'effet de l'installation d'ordinateurs dans les salles de classe sur les scores des élèves à des tests standardisés, quelles que soient la classe ou la discipline. Cette absence d'effet est expliquée par la faible intégration des ordinateurs dans les pratiques des enseignants, malgré une formation à leur utilisation. Une conclusion proche est tirée par Bai *et al.* (2016). Dans leur étude, plus de 6 000 élèves de CM2 étudiant l'anglais dans des écoles rurales en Chine sont répartis en trois groupes : un groupe de comparaison, un groupe où les écoles sont équipées et un groupe où à la fois les écoles sont équipées et l'utilisation des équipements numériques est intégrée dans le programme d'enseignement. Ils ne mesurent d'effet positif que dans ce dernier cas, le simple équipement des écoles n'étant pas suffisant pour impacter les résultats des élèves. Enfin, on peut également citer les études de Bando *et al.* (2016) portant sur l'allocation d'ordinateurs portables dans des écoles élémentaires au Honduras et de Piper *et al.* (2016) portant sur l'allocation de tablettes aux élèves et aux enseignants d'écoles primaires au Kenya : aucune de ces deux études n'identifie d'effet sur les scores des élèves (en mathématiques et en anglais dans le premier cas et en maîtrise de la langue dans le second).

De plus, de nombreuses études ont mis en place des stratégies d'identification crédibles afin de mesurer l'impact causal de l'utilisation du numérique en classe, sans recourir à une attribution aléatoire des équipements. Ainsi, Cristia *et al.* (2014) mobilisent une méthodologie de doubles différences pour analyser les informations relatives à 8 000 élèves du secondaire au Pérou. Ils n'identifient pas d'effet de l'augmentation de la disponibilité d'ordinateurs sur les taux de redoublement et d'abandon. De même, la méthodologie quasi-expérimentale avec appariement statistique mise en place par De Witte et Rogge (2014) à partir des résultats de l'enquête TIMSS ne permet pas d'identifier un impact statistiquement significatif des différents niveaux d'utilisation du numérique sur les résultats en mathématiques.

1.4 Bilan de la revue de la littérature

Outre l'hétérogénéité des résultats évoquée en introduction de cette partie, on peut également noter que, lorsque les études identifient des effets positifs, la taille de ces effets varient majoritairement entre 15 % (Hull et Duch, 2017) et 55 % d'un écart-type (Chauhan, 2017). L'ensemble de nos résultats se situent dans cet intervalle.

Le premier constat de cette revue littérature est l'absence d'étude d'impact d'ampleur nationale en France. La majorité des études portent sur des données états-uniennes, ou collectées dans des pays concernés par le programme « *One Laptop Per Child* », souvent des pays en développement. Par ailleurs, outre l'enjeu de l'impact causal déjà détaillé plus haut, plusieurs autres limites peuvent être identifiées dans cette littérature. La première concerne la description des dispositifs évalués. Les interventions évaluées portent presque uniquement sur l'impact de la distribution d'équipements, et non sur leur utilisation. Pourtant la simple disponibilité d'équipements numériques ne préjuge pas de leur utilisation effective qui dépend de nombreux autres facteurs, notamment de leur adéquation avec les besoins des enseignants et des élèves. Le rôle de l'enseignant est souvent négligé et les études ne mobilisent généralement que des données relatives aux élèves, alors que l'intégration du numérique et l'évolution des pratiques d'enseignement liées à l'utilisation des nouvelles technologies sont des informations importantes à intégrer dans un modèle d'étude d'impact (Barrera-Osorio et Linden, 2009). Ensuite, la majorité des études repose sur une mesure ponctuelle n'identifiant que des effets à court terme et non les effets à moyen, voire long terme. Enfin, le champ des compétences scolaires évaluées se limite très fréquemment aux disciplines scolaires (en particulier aux mathématiques) alors que l'utilisation d'outils numériques peut impacter d'autres formes de compétences comme les compétences sociocognitives telles que la créativité, la collaboration et l'esprit critique.

La partie suivante détaille le dispositif évalué, le Plan numérique pour l'éducation de 2015. Nous expliciterons ensuite la méthodologie retenue, conçue pour répondre aux limites identifiées dans la littérature.

2 Description du Plan numérique

Le Plan numérique, lancé en 2015, a pour finalité d'accroître la part du numérique à l'école et au collège. Il est porté par l'action « Innovation Numérique pour l'Excellence Éducative » (INEE) du Programme d'Investissement d'Avenir et repose sur quatre piliers : « *la formation au numérique, le développement des ressources, l'équipement des acteurs dans un environnement de confiance et le développement des usages* ». Ces piliers sont déclinés en six volets stratégiques :

1. Ressources pédagogiques numériques et services innovants (mise en place d'une banque de ressources pédagogiques numériques, soutien de projets de contenus et partenariat d'innovation sur l'intelligence artificielle).
2. Développement de l'environnement d'accès aux ressources (plateforme sécurisée pour l'accès aux ressources et portail de présentation des ressources à destination des enseignants).
3. Équipements et ressources associées (déploiement de ressources pédagogiques et d'équipements mobiles dans des établissements sélectionnés au travers d'appels à projets, *via* une gouvernance partagée avec les collectivités territoriales et l'État).
4. Soutien à la filière des contenus et services pédagogiques numériques (transformation de la filière des contenus et services pédagogiques).
5. e-FRAN (soutien à des expérimentations innovantes proposées par des équipes pédagogiques).
6. ProFran (centré sur le numérique et l'enseignement professionnel).

Le présent document porte sur l'évaluation du volet 3 du Plan numérique concernant les équipements et les ressources, dans le 2^d degré. Entre 2015 et 2018, 3 069 collèges⁹ publics et privés sous contrat, répartis sur l'ensemble du territoire français métropolitain et d'Outre-mer, ont été sélectionnés après appel à projets. Lors des premières phases, la dotation en équipements a principalement pris la forme d'équipements individuels mobiles (EIM), des tablettes ou plus rarement des ordinateurs portables, attribués individuellement en classe de 5^e aux élèves ayant la possibilité d'utiliser ces équipements en classe et à domicile. Les élèves ne peuvent généralement installer que certaines applications approuvées (par le ministère de l'Éducation nationale, de la Jeunesse et des Sports, les académies, les collectivités territoriales ou les établissements scolaires) et ne peuvent pas utiliser l'équipement après une certaine heure. Ces conditions sont fixées par le gestionnaire de flotte, habituellement le Conseil départemental. Après la première année, les classes mobiles (CM) ont été la forme prédominante des équipements attribués dans le cadre du Plan. Il s'agit de tablettes ou d'ordinateurs portables disponibles sous la forme de valise ou de chariot et communs à tout l'établissement, de la 6^e à la 3^e. Les collèges ont mis en place des systèmes de réservation, formels ou non, pour permettre l'utilisation de ces équipements par tous les enseignants, pour un usage individuel ou collectif en classe avec les élèves.

Le 14 décembre 2016, le comité de pilotage INEE a demandé à la DEPP de mener, en collaboration

9. Ce chiffre a été transmis par la Direction du Numérique Éducatif, Ministère de l'Éducation nationale, de la Jeunesse et des Sports.

avec la Direction du Numérique Éducatif (DNE), une évaluation du volet 3 du Plan numérique, avec pour objectif de mesurer sur trois ans l'impact de la disponibilité des équipements mobiles sur les acquisitions des élèves et sur l'intégration pédagogique du numérique par les enseignants.

Le protocole d'évaluation proposé par la DEPP a été validé par le comité de pilotage INEE, ainsi que par un comité scientifique composé de chercheurs spécialisés en sciences de l'éducation, en psychologie et en évaluation des politiques publiques¹⁰. Ce protocole a été fortement contraint par le calendrier de déploiement des équipements dans les établissements. En effet, de nombreux collèges étant déjà équipés ou en attente d'équipements lors du lancement de l'évaluation, une méthodologie purement expérimentale avec sélection aléatoire des établissements équipés n'était pas envisageable. Pour les mêmes raisons, aucune mesure de référence, avant l'attribution des équipements, n'a pu être réalisée auprès de l'ensemble des élèves concernés. Il a donc été décidé de mettre en place une méthodologie quasi-expérimentale avec un appariement statistique, décrite ci-après.

10. Le comité scientifique est constitué de : Maud Besançon (Rennes 2), Philippe Dessus (LaRAC Grenoble Alpes), Denis Fougère (OSC), Nicolas Gauvrit (Paris-8), Todd Lubart (Paris V) et Françoise Tort (ENS Cachan).

3 Méthodologie d'évaluation

Le cadre conceptuel d'ELAINE a été mis en place afin de mesurer l'impact causal de la disponibilité et de l'utilisation des équipements numériques mobiles sur les apprentissages des élèves d'une part, et sur les pratiques des enseignants d'autre part. Il vise également à observer une potentielle variation de ces effets en fonction de l'origine sociale des élèves, de la composition sociale de l'établissement et du sexe des élèves et de l'enseignant et de l'ancienneté de l'enseignant.

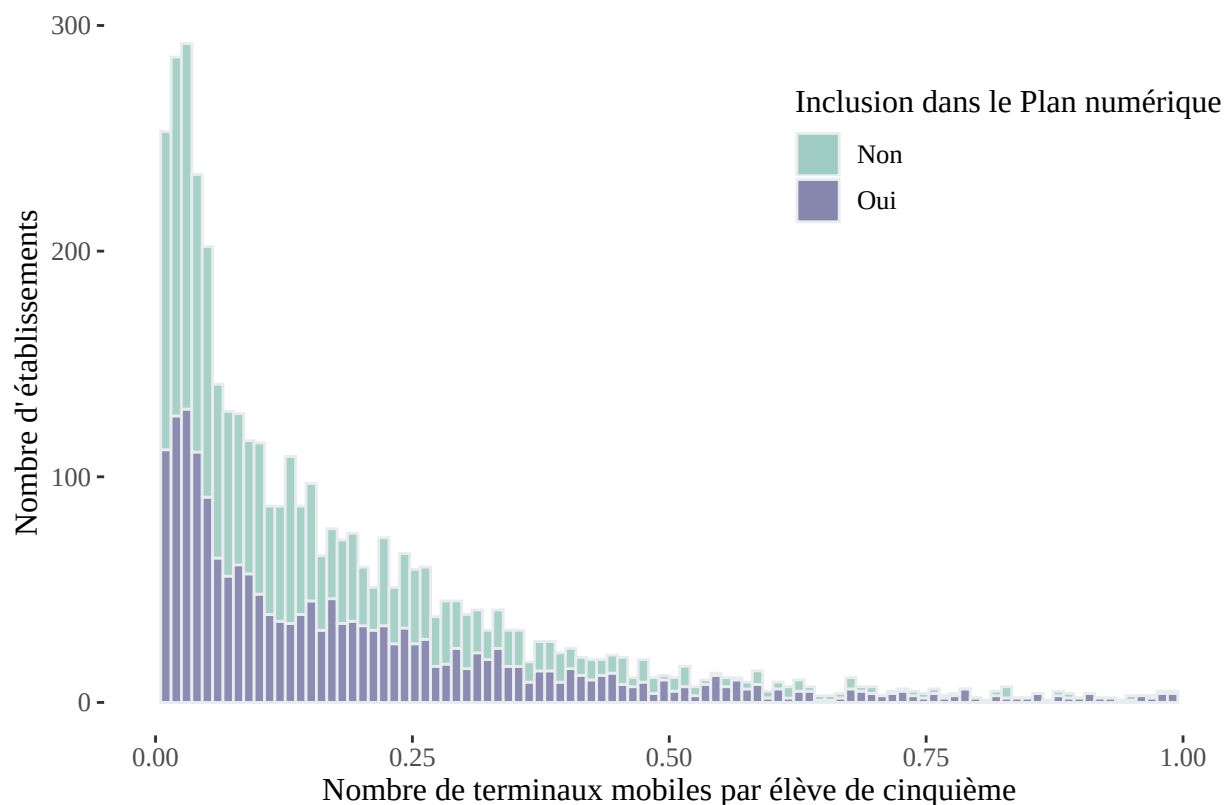
ELAINE est une étude sur échantillon à l'échelle nationale, mais non représentatif de l'ensemble des établissements scolaires. Elle permet le suivi d'élèves, d'enseignants et de chefs d'établissement pendant trois à quatre ans via des questionnaires *ad-hoc*. Disposer d'information sur les trois catégories d'acteurs à inclure dans les analyses permet de réduire le risque de biais statistique de variables omises tout en affinant la mesure de l'utilisation des outils numériques. De plus, les élèves ne sont pas uniquement évalués en mathématiques et en français, mais aussi sur leur créativité, leur esprit critique, leur degré de collaboration, ainsi que sur leurs compétences numériques. Ils sont suivis sur le long terme, ELAINE sera donc en mesure d'apprécier l'impact des équipements sur leur orientation après la classe de 3^e (voies professionnelle, générale ou technologique, enseignements optionnels suivis), leurs résultats au baccalauréat et leur choix d'orientation dans le supérieur, le cas échéant.

De nombreux établissements inclus dans le Plan numérique n'avaient pas encore reçu les équipements numériques mobiles lors du lancement d'ELAINE, tandis que des établissements non-inclus dans le Plan numérique ont été équipés par d'autres biais. Cela se traduit par le chevauchement des distributions du taux d'équipement en équipements numériques mobiles dans les établissements du second degré selon leur inclusion ou non dans le Plan numérique (cf. figure 1). Pour cette raison, la construction des groupes de comparaison ne pouvait reposer sur la seule information de leur inclusion ou non dans le Plan, mais devait aussi reposer sur la disponibilité des différents types d'équipements mobiles (individuels ou en classe mobile). L'échantillon inclut donc des collèges pouvant être répartis dans les cinq groupes suivants :

1. équipements individuels mobiles disponibles (financés par le Plan numérique ou non) ;
2. équipements en classe mobile disponibles et financés par le Plan numérique ;
3. pas d'équipement mobile disponible mais établissement sélectionné dans le Plan numérique ;
4. équipements en classe mobile disponibles et financés par un autre biais que le Plan numérique ;
5. pas d'équipement mobile disponible et établissement non sélectionné dans le Plan numérique.

Ces cinq groupes constituent les groupes bénéficiaires ou de comparaison pour l'ensemble de l'étude d'impact.

Le risque de biais de sélection est ici particulièrement élevé car les établissements du Plan numérique ont été sélectionnés *via* des appels à projet. Le Tableau 1 reporte les résultats d'une régression



Source : ETIC 2017 (DEPP) et DNE

FIGURE 1 – Distribution du taux d'équipement en terminaux mobiles (nombre de terminaux mobiles par élève) selon l'inclusion ou non de l'établissement dans le Plan numérique

logistique sur l'ensemble des collèges en France métropolitaine et DROM, identifiant les caractéristiques des établissements influençant leur inclusion dans le Plan numérique. Les coefficients estimés montrent que les établissements publics, notamment en éducation prioritaire, sont plus fréquemment inclus dans le Plan. Ces variables étant fortement corrélées aux résultats des élèves, la mise en évidence d'un effet causal du numérique sur les apprentissages risque de ce fait d'être soumise à un fort biais de sélection. Par exemple, les élèves des établissements en éducation prioritaire ont en moyenne des résultats plus bas au DNB (Diplôme National du Brevet)¹¹. Les deux groupes d'établissements n'étant pas comparables, la différence entre les pratiques des enseignants et entre les compétences des élèves dans les établissements équipés et ceux non équipés pourrait être liée à d'autres facteurs que la disponibilité des équipements. Par ailleurs, le risque de biais de sélection existe également pour les établissements équipés *via* d'autres sources de financement, la disponibilité de ces équipements n'étant pas aléatoire.

Une méthodologie quasi-expérimentale a été mise en place, mobilisant les méthodes d'appariement d'échantillons (*matching*) introduites par Rubin (1977). La définition d'établissement bénéficiaire

11. Stéfano (2018), *L'éducation prioritaire*, Note d'information de la DEPP.

pourra varier selon l’objectif des analyses et sera précisée au fil de l’analyse : collèges disposant d’équipements individuels mobiles, dans le Plan numérique ou non, collèges disposant d’équipements en classe mobile, dans le Plan numérique ou non. On parlera d’élèves bénéficiaires ou d’enseignants bénéficiaires pour les élèves ou les enseignants appartenant aux collèges bénéficiaires. En faisant l’hypothèse que la sélection des établissements dans le Plan numérique est faite uniquement sur la base de leurs caractéristiques observables¹², ce qui peut bien évidemment être questionné, il est possible de mettre en œuvre la méthode statistique de l’appariement dans une analyse en coupe, c’est-à-dire pour chaque temps de mesure, et dans une analyse de l’évolution des scores des élèves entre chaque temps de mesure. Cette méthode repose sur l’hypothèse que l’inclusion dans le Plan numérique et les résultats des élèves sont indépendants, conditionnellement à un ensemble de variables observables. Ces variables observables agissent donc comme un processus de randomisation qui affecterait la distribution d’équipements numériques aléatoirement à certains collèges et pas à d’autres à caractéristiques d’établissement identiques. Cette méthode permet ainsi de comparer deux élèves, l’un bénéficiaire et l’autre non, scolarisés dans des collèges similaires (à l’exception du fait que l’un de ces collèges est doté en équipements numériques mobiles, l’autre non) et présentant des caractéristiques individuelles (telles que le genre, l’âge, l’origine sociale, etc.) identiques. Plusieurs méthodologies d’appariement statistiques sont envisageables. La méthodologie de l’*entropy balancing* a été retenue, car elle permet le meilleur équilibrage entre les variables observables utilisées lors de l’appariement (voir en annexe pour une description détaillée de cette méthode). Une analyse complémentaire par *Inverse Probability Weighting* (IPW) est également réalisée¹³. Dans l’ensemble du document, seuls les résultats statistiquement significatifs sont commentés. Une mention est ajoutée lorsque les résultats issus de l’IPW ne concordent pas avec ceux de l’*entropy balancing*.

12. Sexe de l’élève, indicatrice selon le lieu de naissance (en France ou hors de France), indicatrice selon le retard scolaire, indicatrice selon l’avance scolaire, PCS la plus élevée des représentants de l’élève, secteur et réseau du collège, taille de l’unité urbaine du collège, effectif total d’élèves du collège, pourcentage de filles dans le collège, composition sociale du collège, taux de retard à l’entrée en 6^e du collège, âge moyen des enseignants, note moyenne du collège à l’épreuve terminale du DNB, écart au niveau du collège entre le taux de réussite global au Diplôme national du brevet et celui des élèves de PCS défavorisées.

13. Chaque unité d’observation reçoit une pondération égale à l’inverse de la probabilité d’être bénéficiaire. Cette probabilité est estimée par le score de propension. Voir Hirano *et al.* (2003).

TABLE 1 – Paramètres estimés du modèle logit visant à expliquer l’inclusion d’un établissement dans le Plan Numérique

	Inclusion dans le Plan numérique
Constante	0.75 (0.82)
Macro-strate (réf. : Etablissement public hors éducation prioritaire)	
Etablissement privé sous contrat	−1.24*** (0.10)
Etablissement public en éducation prioritaire	0.38*** (0.12)
Composition sociale de l’établissement (%) (réf. : défavorisée)	
Très favorisée	−0.02*** (0.00)
Favorisée	0.01 (0.01)
Moyenne	0.01 (0.00)
Effectif d’élèves en classe de 5 ^e	0.00 (0.00)
Effectif enseignants (ETP)	0.02*** (0.01)
Taux de redoublement en 5 ^e	−0.07** (0.03)
Taux de retard en 6 ^e	−0.01 (0.01)
Note moyenne à l’épreuve terminale au Diplôme National du Brevet (DNB)	0.01 (0.04)
Pourcentage de filles	0.01 (0.01)
Âge moyen des enseignants	−0.02 (0.01)
Nombre moyen d’élèves par division	0.01 (0.02)
Population moyenne de la commune	0.00*** (0.00)
Taux de pauvreté moyen dans la zone économique	−0.02*** (0.01)
Chômage moyen de la commune en 2017	−0.06*** (0.02)
AIC	7396.76
BIC	7731.77
Log Likelihood	−3649.38
Deviance	7298.76
Nombre d’obs.	6884

Remarques : L’échantillon porte sur l’ensemble des collèges publics et privés en France métropolitaine et DROM. Le modèle inclut des effets fixes pour les académies. Les données concernent l’année scolaire 2017-2018, et sont donc postérieures à la sélection dans le Plan Numérique, mais on peut supposer que l’inclusion dans le Plan Numérique n’a pas eu d’influence sur ces variables.

Sources : DNE (2017), APAE (2017-2018).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4 Échantillon et temps de mesure

L'échantillonnage des établissements a été réalisé en deux temps :

1. Dans un premier temps, 1 500 collèges publics et privés sous contrat en France métropolitaine et DROM faisant partie du Plan numérique et 1 500 autres n'en faisant pas partie ont été sélectionnés *via* un échantillonnage stratifié avec équilibrage en fonction de la taille de l'établissement. Les strates utilisées étaient le secteur, le réseau et la taille de la commune ou de l'unité urbaine du collège. Ces établissements ont été interrogés en décembre 2017 à propos des équipements numériques mobiles disponibles au sein de l'établissement. Les 2 347 établissements répondants (soit 78 %) constituent la base de sondage pour la seconde phase de l'échantillonnage. L'étude d'impact se fonde sur les informations transmises à ce stade (et ayant potentiellement été révisées par les chefs d'établissement lors de l'envoi du courrier d'information relatif à l'évaluation en mai 2018 puis en mai 2019).
2. Dans un second temps, parmi les établissements répondants et n'ayant pas participé aux pré-tests du matériel d'évaluation, 240 collèges ont été sélectionnés en fonction de leur participation ou non au Plan numérique, et de la présence ou non d'équipements numériques mobiles (individuels et partagés), soit 40 collèges dans chacun des groupes définis dans la section 3, sauf dans le premier groupe où 80 collèges ont été sélectionnés. Parmi les 240 collèges sélectionnés, 217 collèges (soit 90 %) ont transmis les listes des élèves et des enseignants d'une classe de 5^e et sont donc considérés comme appartenant à l'échantillon ELAINE.

4.1 Description des établissements de l'échantillon

Le tableau 2 décrit et situe les 217 établissements de l'échantillon par rapport à la population nationale des collèges. Les établissements du Plan numérique sont surreprésentés. Cela explique également la surreprésentation des établissements publics, plus nombreux dans le Plan numérique. De plus, les établissements de l'échantillon ont en moyenne des effectifs d'élèves et d'enseignants légèrement plus importants que la moyenne nationale, car la probabilité de sélection dans la procédure d'échantillonnage dépendait de l'effectif d'élèves de 5^e. En revanche, les compositions sociales des établissements sont très proches (à l'exception de la catégorie "favorisée"), tout comme l'âge moyen des enseignants et les moyennes au DNB. D'autres variables sont statistiquement différentes entre l'échantillon et la population nationale (notamment le taux de retard en 6^e, le pourcentage de filles et le nombre d'élèves par division) mais ces écarts sont toujours faibles.

Au sein des 217 collèges, une classe de 5^e a été sélectionnée (selon une règle aléatoire basée sur l'initiale du nom de famille du professeur principal¹⁴). Dans cette classe, tous les élèves et les

14. Le chef d'établissement était invité à ordonner les classes de 5^e par ordre alphabétique selon l'initiale du nom de famille du professeur principal et à sélectionner la première classe de cet ordonnancement.

enseignants des 11 disciplines principales¹⁵ ont été intégrés à l'échantillon. Ces listes ont été établies par les chefs d'établissement. En 4^e, pour le deuxième temps de mesure, les enseignants des élèves ont été ajoutés à l'échantillon, en utilisant les bases RELAIS, des bases administratives recensant l'ensemble des enseignants et leur affectation.

TABLE 2 – Comparaison des collèges de l'échantillon ELAINE avec la population nationale de collèges

	Collèges en France métropolitaine et DROM	Collèges de l'échantillon	Ecart entre l'échantillon et la population nationale
Dans le Plan numérique	43.02	61.75	18.73***
Public	75.97	82.03	6.06**
Réseau (%)			
Hors éducation prioritaire	84.38	82.49	-1.9
REP	10.42	12.9	2.48
REP+	5.2	4.61	-0.59
Composition sociale de l'établissement (%)			
Très favorisée	20.55	20.4	-0.15
Favorisée	11.9	12.53	0.64*
Moyenne	26.69	26.91	0.22
Défavorisée	35.56	35.08	-0.49
Effectif moyen d'élèves de cinquième	115.12	123.92	8.8***
Effectif moyen d'enseignants (ETP)	31.6	33.72	2.12***
Taux de redoublement des élèves de cinquième	0.42	0.38	-0.04
Taux de retard en sixième	8.3	7.47	-0.82***
Note moyenne à l'épreuve terminale du DNB	10.9	10.88	-0.02
Pourcentage de filles	48.83	49.29	0.45**
Âge moyen des enseignants	42.45	42.66	0.21
Nombre d'élèves par division	24.96	25.44	0.48***
Population moyenne de la commune	38155.44	31552.59	-6602.85*
Taux de pauvreté moyen dans la zone économique	17.63	17.61	-0.03
Chômage moyen de la commune en 2017	9.37	9.65	0.28*
Nombre d'observations	7025	217	

Note : Les niveaux de significativité statistique sont indiqués par les symboles suivants : * pour 10 %, ** pour 5 %, *** pour 1 %. La significativité des écarts est mesuré par un test de Chi2 pour les variables catégorielles et par un test de Student pour les variables numériques.

Champ : Collèges publics et privés sous contrat en France métropolitaine et DROM en 2017

Source : DNE, APAE 2017-2018 (DEPP), INSEE

15. Français, Mathématiques, Langues vivantes 1 et 2, Histoire-Géographie-Education Morale et Civique, Sciences Physiques, Sciences de la Vie et de la Terre, Technologie, Arts Plastiques, Musique et Éducation Physique et Sportive.

4.2 Temps de mesure

Les élèves de l'échantillon sont observés de la 5^e à la 2^{de}. Leurs enseignants de 5^e des 11 disciplines principales sont observés pendant deux ans. Leurs enseignants de 4^e sont interrogés lors du 2^e temps de mesure. Les évaluations des élèves et les enquêtes auprès des enseignants ont eu lieu en mai-juin 2018 et en mai-juin 2019. De plus, les chefs d'établissement ont été interrogés en mai-juin 2019. Les résultats des élèves aux évaluations exhaustives à l'entrée en 2^{de} ainsi que les données relatives à leur orientation après la 3^e seront collectés lors de l'année scolaire 2020-2021. Les élèves, les enseignants et les chefs d'établissement n'ont pas pu être interrogés pour le troisième temps de mesure prévu en mai 2020 en raison de la crise sanitaire et de la fermeture des établissements scolaires. La DEPP prévoit également un suivi à plus long terme des élèves (poursuite d'études dans le supérieur).

4.3 Description des élèves de l'échantillon

L'échantillon est initialement constitué de 5 630 élèves de 5^e. Parmi ces élèves, 5 203 sont suivis en 4^e (les 427 restants ont soit redoublé, soit changé d'établissement et ne sont pas suivis). L'échantillon de 4^e est décrit en fonction du groupe de leur établissement dans le tableau 3. La proportion de filles est quasiment identique dans chacun des cinq groupes d'établissements. En revanche, la composition sociale est plus inégale entre les groupes, les groupes 4 et 5 (hors Plan numérique, respectivement avec des équipements mobiles partagés et sans équipements mobiles) étant les plus favorisés. Cela s'explique par les caractéristiques d'établissement, notamment le secteur (la proportion d'élèves dans un établissement public est plus élevée dans les groupes 2 et 3) et le réseau (les groupes 1 et 2 réunissent une proportion d'élèves en REP et REP+ plus importante que la moyenne de l'échantillon). De même, les élèves des établissements des groupes 1 à 3 ont des résultats moyens au Diplôme National du Brevet (DNB) inférieurs au reste de l'échantillon. Enfin, les élèves dans des établissements implantés dans des unités urbaines de grande taille sont surreprésentés dans le groupe 1.

4.4 Description des enseignants de l'échantillon

Au premier temps de mesure, 2 541 enseignants de 5^e font partie de l'échantillon. Cet échantillon est décrit en fonction du groupe de leur établissement dans le tableau 4. La proportion de femmes et d'hommes est relativement stable entre les groupes, tout comme les répartitions par tranches d'âge (avec cependant une proportion de jeunes enseignants plus élevée dans les groupes 2 et 3). Les variations entre les groupes sont plus importantes concernant les variables au niveau établissement : la proportion d'enseignants exerçant dans un établissement public est nettement supérieure dans les groupes 2 et 3, en comparaison aux groupes 4 et 5 tandis que le groupe 1 se situe dans la moyenne. La répartition des enseignants en fonction du réseau d'éducation de l'établissement n'est pas homogène non plus : la proportion d'enseignants rattachés à des établissements en réseau

d'éducation prioritaire est nettement plus importante dans les groupes 1 et 2 et plus faibles dans les groupes 4 et 5, résultat qui se reflète dans la distribution des enseignants en fonction des résultats au Brevet, les établissements en réseau d'éducation prioritaire ayant en moyenne des notes plus basses au DNB. Enfin, les enseignants classés dans le groupe 1 exercent dans des établissements situés dans des grandes unités urbaines (52 % dans une unité urbaine de plus de 200 000 habitants ou en agglomération parisienne, contre 38 % pour l'ensemble des enseignants de l'échantillon).

TABLE 3 – Description des élèves de l'échantillon selon le groupe de leur établissement

	Groupe 1 Equipements individuels		Groupe 2 Equipements partagés dans le cadre du PN		Groupe 3 Pas d'équipement mais sélectionné dans le PN		Groupe 4 Equipements partagés en dehors du PN		Groupe 5 Pas d'équipement et pas sélectionné dans le PN		Total	
	%	N	%	N	%	N	%	N	%	N	%	N
Fille	49.09	783	50.75	542	51.74	387	48.37	356	49.43	522	49.78	2590
PCS de l'élève												
Très favorisée	28.99	456	30.97	328	25.03	185	32.97	239	33.69	350	30.34	1558
Favorisée	17.86	281	14.54	154	16.10	119	19.17	139	18.09	188	17.16	881
Moyenne	32.87	517	32.48	344	35.05	259	32.55	236	31.09	323	32.70	1679
Défavorisée	20.28	319	22.00	233	23.82	176	15.31	111	17.13	178	19.81	1017
Secteur de l'établissement												
Secteur privé sous contrat	20.25	323	10.86	116	3.21	24	27.85	205	30.49	322	19.03	990
Secteur public en éducation prioritaire	18.62	297	19.10	204	15.64	117	11.14	82	10.42	110	15.57	810
Secteur public hors éducation prioritaire	61.13	975	70.04	748	81.15	607	61.01	449	59.09	624	65.40	3403
Taille de l'unité urbaine												
Plus de 200 001 hab. ou Agglomération parisienne	50.72	809	42.88	458	26.47	198	29.62	218	32.01	338	38.84	2021
Entre 20 001 et 200 000 hab.	23.64	377	23.50	251	30.21	226	12.77	94	19.60	207	22.20	1155
Entre 2 001 et 20 000 hab.	19.12	305	26.03	278	31.28	234	45.52	335	32.20	340	28.68	1492
Commune rurale	6.52	104	7.58	81	12.03	90	12.09	89	16.19	171	10.28	535
Quartile de la note moyenne à l'épreuve terminale du DNB												
Inférieure à 10	33.48	534	26.50	283	25.53	191	17.80	131	17.64	182	25.51	1321
Strict. sup. à 10 et inf. ou égale à 11	23.07	368	22.10	236	45.86	343	21.88	161	21.32	220	25.64	1328
Strict. sup. à 11 et inf. ou égale à 12	19.25	307	17.88	191	18.45	138	31.93	235	37.21	384	24.23	1255
Strict. sup. à 12	24.20	386	33.52	358	10.16	76	28.40	209	23.84	246	24.62	1275
Quartile de l'effectif global d'élèves												
Inférieur à 390	21.63	345	23.22	248	21.79	163	25.14	185	34.38	363	25.06	1304
Supérieur à 391 et inférieur à 520	26.21	418	33.90	362	12.70	95	28.80	212	22.63	239	25.49	1326
Supérieur à 521 et inférieur à 645	27.96	446	23.78	254	32.35	242	26.63	196	14.68	155	24.85	1293
Supérieur à 646	24.20	386	19.10	204	33.16	248	19.43	143	28.31	299	24.60	1280

Sources : APAE 2017-2018, BCE 2017 (DEPP)

TABLE 4 – Description des enseignants de l'échantillon selon le groupe de leur établissement

	Groupe 1 Equipements individuels		Groupe 2 Equipements partagés dans le cadre du PN		Groupe 3 Pas d'équipement mais sélectionné dans le PN		Groupe 4 Equipements partagés en dehors du PN		Groupe 5 Pas d'équipement et pas sélectionné dans le PN		Total	
	%	N	%	N	%	N	%	N	%	N	%	N
Femme	63.14	466	59.23	292	62.61	216	61.19	216	62.26	292	61.80	1482
Quartile d'âge												
Moins de 39 ans	28.32	209	31.50	155	31.88	110	24.15	85	25.37	119	28.30	678
De 39 à 44 ans	21.95	162	23.37	115	22.32	77	26.70	94	19.62	92	22.54	540
De 45 à 51 ans	26.02	192	24.19	119	20.87	72	26.70	94	27.72	130	25.33	607
Plus de 51 ans	23.71	175	20.93	103	24.93	86	22.44	79	27.29	128	23.83	571
Secteur de l'établissement												
Secteur privé sous contrat	18.92	150	8.92	47	3.57	13	24.45	89	28.60	141	17.32	440
Secteur public en éducation prioritaire	20.18	160	21.06	111	15.93	58	12.91	47	11.76	58	17.08	434
Secteur public hors éducation prioritaire	60.91	483	70.02	369	80.49	293	62.64	228	59.63	294	65.60	1667
Taille de l'unité urbaine												
Plus de 200 001 hab. ou Agglomération parisienne	50.44	400	42.31	223	25.27	92	29.67	108	30.43	150	38.29	973
Entre 20 001 et 200 000 hab.	24.09	191	22.58	119	29.12	106	13.74	50	20.69	102	22.35	568
Entre 2 001 et 20 000 hab.	18.92	150	26.76	141	33.24	121	44.23	161	32.66	161	28.89	734
Commune rurale	6.56	52	8.35	44	12.36	45	12.36	45	16.23	80	10.47	266
Quartile de la note moyenne à l'épreuve terminale au DNB												
Inférieure à 10	34.80	276	28.27	149	25.82	94	19.78	72	19.92	96	27.15	687
Strict. sup. à 10 et inf. ou égale à 11	23.96	190	21.44	113	45.05	164	22.80	83	22.41	108	26.01	658
Strict. sup. à 11 et inf. ou égale à 12	19.17	152	18.79	99	18.96	69	31.87	116	36.93	178	24.27	614
Strict. sup. à 12	22.07	175	31.50	166	10.16	37	25.55	93	20.75	100	22.57	571
Quartile de l'effectif global d'élèves												
Inférieur à 385 élèves	22.07	175	22.05	114	22.25	81	25.27	92	36.31	179	25.33	641
Supérieur à 386 et inférieur à 510 élèves	24.21	192	33.08	171	12.64	46	28.02	102	22.11	109	24.50	620
Supérieur à 511 et inférieur 640 élèves	30.01	238	25.73	133	29.12	106	27.20	99	14.60	72	25.60	648
Supérieur à 641 élèves	23.71	188	19.15	99	35.99	131	19.51	71	26.98	133	24.58	622

Sources : APAE 2017-2018, BCE 2017 (DEPP)

5 Matériel d'enquête et d'évaluation et construction des indicateurs de résultat

L'évaluation multidimensionnelle du Plan numérique conçue par la DEPP, en lien avec le conseil scientifique, a conduit cette dernière à s'associer à différentes équipes de recherche afin de construire les différents instruments d'évaluation et d'enquête. Ces outils ont été testés auprès d'élèves, d'enseignants et de chefs d'établissement (hors échantillon ELAINE) afin de s'assurer de la qualité des items et questions proposés dans le cadre de l'évaluation. Un effort a été mené par la DEPP afin d'uniformiser autant que possible les procédures de validation des items et questions ainsi que les procédures de construction des indicateurs bâtis à partir des réponses. Le lecteur comprendra néanmoins que l'implication d'équipes de chercheurs de champs disciplinaires différents a pu conduire à des critères différents s'agissant de la construction de ces indicateurs.

5.1 Élèves

Les compétences des élèves ont été mesurées via une évaluation sur support numérique, divisée en deux séquences de 50 minutes, en fin de 5^e et de 4^e. Les contenus ont été réalisés par le bureau B2 de la DEPP, qui était également chargé de l'organisation et du suivi de l'administration de ces évaluations.

5.1.1 Questionnaire sur l'utilisation du numérique

Les élèves ont été interrogés sur leur fréquence et type d'utilisation, chez eux ou au collège, des équipements numériques, que ces équipements aient été fournis dans le cadre du Plan numérique ou non. Ils étaient par exemple interrogés sur la fréquence d'utilisation des outils numériques pour l'envoi de message, la recherche d'information, etc. Ce questionnaire n'a été soumis aux élèves qu'en 4^e, lors du deuxième temps de mesure. Il incluait des questions portant également sur l'année de 5^e ¹⁶.

Deux types d'analyses sont réalisées sur la base de ces informations. La première, descriptive, vise à identifier les effets de la distribution des équipements numériques mobiles sur l'accès aux outils numériques et leur utilisation (cf. partie 7). La seconde, visant à tester la robustesse des résultats de l'étude d'impact, utilise ces informations pour construire les groupes de comparaison indépendamment des informations transmises par le chef d'établissement en décembre 2017 (cf. partie 9).

16. Pour des raisons techniques, le questionnaire n'avait pu être soumis aux élèves en 5^e lors du premier temps de mesure.

5.1.2 Évaluation des compétences disciplinaires (mathématiques et français) et numériques

5.1.2.1 Contenu des évaluations

Les évaluations de mathématiques se composaient de 19 items en 5^e et de 17 items en 4^e (dont 10 items identiques dits d’ancrage) et portant respectivement sur l’ensemble du programme de mathématiques de 5^e et de 4^e. Les évaluations de français se composaient d’un exercice en compréhension écrite (contenant 13 items, identiques en 5^e et 4^e) et d’un exercice de compréhension orale (11 items en 5^e et 12 items en 4^e, tous différents).

L’évaluation des compétences numériques a été élaborée par un groupe de travail, piloté par la DEPP, constitué de professeurs de différentes disciplines (lettres modernes, mathématiques, technologie, sciences de la vie et de la terre) et documentalistes. Elle comptait 52 items, identiques en 5^e et 4^e et portant sur les champs suivants : la communication avec le numérique, la connaissance et le respect de la loi, la gestion de son espace de travail, l’identité numérique, les matériels et réseaux, la pensée algorithmique et la recherche d’information. Pour définir ces champs, le groupe de travail a consulté des rapports d’universitaires et le référentiel européen (DigiComp 2.1 de 2017) sur lequel s’appuie la plateforme PIX, mais aussi le cadre de référence des compétences numériques (CRCN) élaboré par les ministères de l’Éducation nationale, de la Jeunesse et des Sports et de l’Enseignement supérieur, de la Recherche et de l’Innovation.

5.1.2.2 Expérimentation et sélection des items

Les évaluations de mathématiques et de français avaient déjà été expérimentées dans le cadre de précédentes opérations conduites par la DEPP. La sélection des items avait donc été réalisée en amont de l’étude ELAINE.

Les items de l’évaluation des compétences numériques ont été sélectionnés après expérimentation de deux modules en février 2018 auprès d’élèves de 5^e dans des collèges ne faisant pas partie de l’échantillon ELAINE : le premier contenait 54 items et a été expérimenté auprès de 459 élèves, le second contenait 57 items et a été expérimenté auprès de 451 élèves du même niveau d’enseignement. Le premier critère de sélection était le pourcentage de réussite : les items ayant un taux de réussite inférieur à 20 % étaient exclus. Le second était une mesure de la difficulté, estimée avec la fonction « *tam.mml.2pl* » du package *tam.mml* sur R, avec exclusion des items ayant une difficulté supérieure à 4 en valeur absolue. Enfin, on a calculé le coefficient de corrélation point bisériale, mesurant la corrélation entre le score à l’item (1 si l’item est correct, 0 sinon) et le score total (la somme des scores de chaque item). Les items dont les coefficients de corrélation point bisériale étaient inférieurs à 0,2 n’ont pas été retenus (Laveault et Grégoire, 2002). Parmi l’ensemble des items répondant à ces critères, une sélection finale a pu être effectuée pour obtenir un test d’une durée inférieure à 30 minutes et contenant des items relatifs à l’ensemble des domaines souhaités.

5.1.2.3 Calcul des scores des élèves aux évaluations de mathématiques, de français et de compétences numériques

Les réponses des élèves aux évaluations ont été collectées en mai-juin 2018 pour la 5^e et en mai-juin 2019 pour la 4^e. Le processus de nettoyage et de calcul des scores est identique pour les deux temps de mesure et pour les quatre domaines : mathématiques, compréhension écrite, compréhension orale et compétences numériques. Le coefficient de corrélation point bisériale a été à nouveau calculé et on a retiré de l'analyse les items pour lesquels ce coefficient était inférieur à 0,2 (en 5^e, suppression de 4 items de l'évaluation de compétences numériques et de 5 items de l'évaluation de mathématiques ; en 4^e, suppression de 3 items de l'évaluation de compétences numériques). On a également vérifié que la mesure de la difficulté des items était bien inférieure à 4 en valeur absolue, ce qui est toujours le cas.

Le score des élèves est estimé selon un modèle de réponse à l'item à deux paramètres. Ces paramètres sont la difficulté de l'item et sa discrimination (plus un item est discriminant, plus la probabilité de le réussir est faible en-deçà d'un niveau de compétence et plus elle sera élevée au-dessus de ce même niveau de compétence). Le score est égal au paramètre de compétence de ce modèle, qui est estimé par maximum de vraisemblance.

5.1.3 Évaluation des compétences sociocognitives (esprit critique, créativité et collaboration)

Les élèves de l'échantillon ELAINE ont également été invités en 5^e et en 4^e à répondre à des échelles visant à mesurer leur esprit critique, leur créativité et leurs pratiques collaboratives. Ces dimensions ont été étudiées car le développement de ces compétences sociocognitives est primordial pour permettre aux élèves d'évoluer dans le monde contemporain et devenir des citoyens au sein d'une société de la connaissance (Ananiadou et Claro, 2009).

5.1.3.1 Contenu des évaluations

Le questionnaire portant sur l'esprit critique a été conçu par l'équipe de recherche de Nicolas Gauvrit (Paris-8). Il distingue les dispositions individuelles telles que des habitudes ou attitudes générales de pensée, des compétences cognitives (soit la capacité d'un élève à penser de façon critique). Ces dernières sont divisées en cinq dimensions : l'interprétation (capacité à décoder), l'analyse, l'évaluation (capacité à juger de la pertinence d'un argument ou de la fiabilité d'une source), l'inférence (le raisonnement déductif et probabiliste, l'induction) et l'explication (capacité à formuler et justifier un raisonnement).

La créativité et la collaboration des élèves sont évaluées au moyen d'un questionnaire conçu par l'équipe de Todd Lubart (Université de Paris). La créativité est définie comme « *la capacité à réaliser une production qui soit à la fois nouvelle et adaptée au contexte dans lequel elle se manifeste* »

(Sternberg et Lubart, 1995). L'approche multivariée identifie quatre types de ressources nécessaires pour l'émergence de la créativité : les facteurs cognitifs (intelligence et connaissances), conatifs (personnalité et motivation), émotionnels, et l'environnement (Lubart et *al.*, 2003). La créativité est mesurée sur la base de deux échelles : l'échelle de Karwowski (ou « *Short Scale of Creative Self* ») qui invite les élèves à se positionner sur une échelle de Likert contenant onze affirmations liées à la créativité (par exemple : « Je suis doué(e) pour proposer des solutions originales aux problèmes. »), et une échelle d'auto-positionnement concernant la créativité, contenant également onze affirmations (par exemple : « la créativité, c'est imaginer des choses qui n'existent pas. »). La collaboration est mesurée par une échelle d'auto-évaluation, constituée de dix affirmations à propos desquelles les élèves doivent indiquer leur niveau d'accord (par exemple : « je discute le pour et le contre des idées proposées. »). Ces deux dernières échelles ont été spécifiquement créées pour l'étude ELAINE.

5.1.3.2 Expérimentation des items

Ces échelles ont été expérimentées en février 2018 auprès de 910 élèves de 5^e scolarisés dans des établissements ne faisant pas partie de l'échantillon ELAINE.

5.1.3.2.1 Calcul des scores des élèves

Le calcul des scores des élèves est en cours, en collaboration avec les équipes de recherche concernées.

5.2 Enseignants

5.2.1 Partie du questionnaire relative aux pratiques d'enseignement

5.2.1.1 Contenu de la partie

La partie du questionnaire développée par l'équipe de recherche de Philippe Dessus (Université Grenoble-Alpes) couvre un large éventail de dimensions : les équipements disponibles des enseignants (au domicile et au collège), la formation au numérique, l'intégration du numérique dans les pratiques professionnelles, les freins à l'usage du numérique, les pratiques collaboratives entre enseignants, les pratiques visant à favoriser la collaboration entre élèves et le sentiment d'auto-efficacité des enseignants en matière d'intégration du numérique. Cette dernière dimension est mesurée par l'échelle TPACK (*Technology, Pedagogy and Content Knowledge*) créée par Mishra et Koehler (2006). Elle permet de documenter cette dimension en les interrogeant sur des aspects liés aux savoirs didactiques, ainsi qu'aux connaissances pédagogiques et technologiques. Les enseignants étaient invités à préciser leur degré d'adhésion à des propositions telles que « *Je suis capable de dispenser des cours qui combinent de façon pertinente la discipline que j'enseigne, les technologies numériques et la pédagogie* » ou « *Je suis capable de choisir des outils numériques qui illustrent le contenu de mon cours et optimisent mon enseignement* ». Ainsi, la notion d'intégration du numérique ne se

définit pas ici par la simple utilisation des outils numériques en classe mais prend en compte les connaissances nécessaires à l'intégration des outils dans l'enseignement (Harris *et al.*, 2009).

5.2.1.2 Expérimentation des items

Le questionnaire a été expérimenté auprès de 385 enseignants de 5^e des 11 disciplines principales répartis dans 45 collèges ne faisant pas partie de l'échantillon ELAINE. Les items pour lesquels une modalité a été sélectionnée par moins de 5 % des répondants ou ne saturant pas dans le cadre de l'analyse factorielle exploratoire sont retirés du questionnaire final ou amendés.

5.2.1.3 Construction des indicateurs sur les enseignants

Les réponses des enseignants ont été collectées en mai-juin 2018 puis en mai-juin 2019. Tout d'abord, les modalités retenues pour le questionnaire sont regroupées en fonction de la distribution des répondants pour chaque question, quand cela avait du sens, pour qu'il ne reste aucune modalité ayant été sélectionnée par moins de 5 % de l'échantillon. Lorsque cela n'est pas possible, la question est retirée de l'analyse (15 items supprimés pour cette raison).

La sous-partie « Recours en cas de difficulté » porte sur les recours les plus fréquents des enseignants en cas de difficultés techniques ou pédagogiques avec le numérique. Elle est constituée de questions sous un format différent et n'était pas destinée à une analyse factorielle similaire aux autres sous-parties. Elle n'est donc pas incluse dans ces analyses.

La mesure de la consistance interne vise à évaluer l'homogénéité d'un sous-ensemble de questions, en indiquant le degré auquel les réponses aux items de ce sous-ensemble sont similaires. Plusieurs mesures sont utilisées pour déterminer la consistance interne de chacune des sous-parties de l'enquête : la corrélation moyenne inter-items (la corrélation d'un item avec chacun des autres items du sous-ensemble), la corrélation moyenne item-total (la corrélation de chaque item à la moyenne des items du sous-ensemble) et l'alpha de Cronbach. Est également reporté ici le nombre de variables ayant une corrélation item-total inférieure à 0,2 et qui sont exclues de l'analyse.

Les seuils considérés comme acceptables dans la littérature pour la corrélation inter-item moyenne se situent entre 0,15 et 0,50, tandis que la corrélation item-total moyenne doit être la plus élevée possible, et la corrélation item-total de chaque item doit être supérieure à 0,2 (Clark et Watson, 1995).

Les mesures de consistance interne sont détaillées dans les tables 5 et 6 et sont globalement satisfaisantes. Il faut toutefois noter que :

- un item a été supprimé dans la sous-échelle « Connaissance de la technologie » : « *J'ai une bonne connaissance générale des technologies* » car sa corrélation inter-item était élevée (0,66). Pour les sous-échelles « Connaissance du contenu » et « Connaissance de la pédagogie », les corrélations inter-item moyennes restent élevées, mais le nombre d'items étant bas et

les alphas de Cronbach et les corrélations item-total étant satisfaisants, on conserve ces sous-parties en l'état ;

- dans la sous-partie « Formation au numérique », un item relatif à la formation initiale a été supprimé car il avait une corrélation inter-item faible (0,11), ce qui s'explique par la thématique relativement éloignée des autres items, qui portent tous sur la formation continue ;
- les mesures de la consistance interne de la sous-partie « Équipements disponibles » ne sont pas entièrement satisfaisantes. Les mêmes mesures ont été effectuées en divisant cette sous-partie en deux, distinguant « Équipements disponibles au domicile » et « Équipements accessibles dans l'établissement » sans qu'elles en soient améliorées. La suppression des deux items « Ordinateur fixe disponible à domicile » et « Ordinateurs fixes disponibles dans l'établissement » (deux items ayant des corrélations inter-item faibles) améliorent l'alpha de Cronbach qui passe de 0,53 à 0,56. On retire donc ces deux items et on mesure deux indicateurs distinguant les équipements à domicile et ceux dans l'établissement (ordinateur fixe, ordinateur portable, connexion internet, tablette électronique, liseuse numérique, smartphone, vidéoprojecteur, tableau numérique interactif).

L'analyse en correspondances multiples est une méthode statistique utilisée dans l'analyse de données de volume important, où un grand nombre d'individus sont décrits par un grand nombre de variables qualitatives (Duval, 2017). L'information contenue dans ces données est synthétisée via des graphiques où les individus et leurs caractéristiques observées sont projetés sur des plans factoriels. Les individus aux caractéristiques les plus proches et les modalités de réponse partagées par ces individus sont à une distance faible sur ces plans. Au contraire, plus les individus diffèrent, plus la distance augmente. Les indices retenus pour synthétiser l'information sont les coordonnées de chaque individu sur le plan.

Des analyses de correspondances multiples ont été réalisées avec le package *FactoMineR* sur R, au sein de chacune des sous-parties listées précédemment, avec les ajustements précisés, et ce à l'exception des sous-parties « Connaissance pédagogique du contenu » et « Recours en cas de difficultés et maîtrise ».

Pour chaque sous-partie, on ne retient qu'une seule dimension suite à l'analyse de l'histogramme des valeurs propres, à partir du critère du coude. Ainsi, le nombre de dimensions est choisi au point de cassure de la courbe du pourcentage de variance expliquée en fonction du nombre de dimensions. Augmenter le nombre de dimensions n'apporterait qu'un gain marginal faible en termes de variance expliquée. Pour certaines sous-parties, une seconde dimension pourrait être retenue, mais cette seconde dimension mesure toujours la dispersion des réponses, qui n'est pas une mesure pertinente dans le cadre de cette étude.

TABLE 5 – Mesures de la consistance interne des indicateurs du questionnaire à destination des enseignants - première partie

Sous-parties	Nombre d'items	Alpha de Cronbach	Corrélation moyenne inter-item	Corrélation moyenne item-total	Nombre de variables avec corrélation item-total <0,2
Formation au numérique	14	0.76	0.23	0.51	1
Equipements disponibles à domicile	5	0.76	0.32	0.63	0
Equipements disponibles dans l'établissement	7	0.76	0.25	0.56	0
Intégration du numérique dans les pratiques d'enseignement	18	0.86	0.29	0.56	0
Freins	15	0.86	0.34	0.61	0
Pratiques collaboratives entre les enseignants	12	0.81	0.33	0.61	0
Pratiques collaboratives entre les élèves	15	0.82	0.28	0.56	0

TABLE 6 – Mesures de la consistance interne des indicateurs du questionnaire à destination des enseignants - deuxième partie

Sous-parties	Nombre d'items	Alpha de Cronbach	Corrélation moyenne inter-item	Corrélation moyenne item-total	Nombre de variables avec corrélation item-total <0,2
Connaissance du contenu	3	0.76	0.68	0.87	0
Connaissance de la pédagogie	4	0.79	0.61	0.83	0
Connaissance de la technologie	8	0.87	0.54	0.76	0
Connaissance pédagogique du contenu	1				
Connaissance technologique du contenu	3	0.84	0.76	0.9	0
Connaissance technologique du contenu	3	0.84	0.76	0.9	0
Connaissance techno-pédagogique	5	0.84	0.6	0.81	0
Connaissance techno-pédagogique du contenu	6	0.89	0.64	0.83	0

5.2.2 Partie du questionnaire relative à la créativité et à l'esprit critique

5.2.2.1 Contenu du questionnaire

L'enquête auprès des enseignants porte également sur la créativité. L'objectif de ces questions était de déterminer si la créativité des enseignants pouvait expliquer les effets différenciés de la distribution des équipements sur les pratiques et ainsi de mettre en lumière un des mécanismes possibles par lesquels les pratiques d'enseignement peuvent influencer les scores des élèves. Le questionnaire a été conçu par l'équipe de recherche de Maud Besançon (Université Rennes 2). Il se compose de 3 échelles :

- l'échelle de Karwowski, identique à celle utilisée pour les élèves ;

- l'échelle « *Use of Creative Cognition Scale* » (Rogaten et Moneta, 2013) portant sur les techniques et stratégies permettant aux personnes d'être créatives ;
- l'échelle « *Self-Report of Creative Traits* » (Runco, 2017) permettant de mesurer la personnalité créative des individus à travers des descriptifs de personnalité. Pour des questions de calendrier, ces échelles n'ont pu être expérimentées dans le cadre d'ELAINE. Les analyses sont donc réalisées directement à partir des données recueillies auprès des enseignants de l'échantillon ELAINE, en mai-juin 2018 puis mai-juin 2019.

5.2.2.2 Construction des indicateurs

Pour chacune de ces échelles, trois tests sont effectués afin de vérifier la sensibilité des items (validation graphique de la distribution des items), la structure factorielle du test (par analyse factorielle) et la consistance interne (mesure de l'alpha de Cronbach et des corrélations des dimensions du test). Le tableau 7 résume les résultats de ces tests.

TABLE 7 – Mesures de la consistance interne des indicateurs du questionnaire à destination des enseignants - partie sur la créativité

Échelle	Sensibilité des items	Structure factorielle	Consistance interne
Échelle de Karwowski	Satisfaisante	1 facteur	0,91
Use of Creative Cognition Scale	Acceptable, la moyenne des items est très élevée	1 facteur	0,84
Self-Report of Creative Traits	Acceptable, la moyenne des items est élevée	Incertaine	0,65

Les trois échelles possèdent les qualités psychométriques suffisantes pour être exploités et seront donc utilisées pour les analyses.

5.2.3 Mesure des compétences numériques des enseignants

5.2.3.1 Contenu de l'enquête

L'enquête portant sur les compétences numériques des enseignants est composée de 22 items au premier temps de mesure (auprès des enseignants de 5^e en juin 2018) et de 22 items distincts au deuxième (auprès des enseignants de 5^e et de 4^e en juin 2019). Les items ont été sélectionnés dans le contenu de l'évaluation des élèves, en choisissant les items les plus généraux. L'évaluation couvre les

champs suivants : la communication avec le numérique, la gestion de son espace de travail, l'identité numérique et les matériels et réseaux.

5.2.3.2 Expérimentation des items

Soixante-seize items ont été expérimentés en février 2018. Ils ont été répartis en trois modules de 32 items chacun (certains items sont dans plusieurs modules). Les modules ont été soumis à 340 enseignants (entre 99 et 125 enseignants par module). Au sein de ces 76 items, les items finaux ont été sélectionnés selon les critères suivants. Les items dont le pourcentage de bonnes réponses est inférieur à 60 % et l'indice de discrimination inférieur à 0,3 sont supprimés. Le coefficient de corrélation point bisériale doit être supérieur à 0,2 et la mesure de la difficulté, à 4 en valeur absolue. Enfin, l'*infit* se trouve entre 0,8 et 1,2 et l'*outfit* entre 0,7 et 1,3. Ces deux mesures indiquent le degré de précision des données par rapport au modèle. Parmi les items répondant à l'ensemble de ces critères, les items finaux du premier et deuxième temps de mesure ont été sélectionnés de telle sorte que le temps de passation soit inférieur à 30 minutes.

5.2.3.3 Construction des indicateurs sur les enseignants

La méthodologie suivie est proche de celle utilisée pour la construction des scores des élèves. La consistance interne du test est validée par la mesure de l'alpha de Cronbach, ici égale à 0,77, soit un niveau acceptable. Sur la base du critère du coude, la structure unidimensionnelle du test est validée.

Pour chaque item, on mesure ensuite un indice de discrimination (suppression des items si cet indice est inférieur à 0,2) et le coefficient de corrélation point bisériale (suppression de items si le coefficient est inférieur à 0,2, nous avons été ici moins stricts que lors de l'expérimentation). Quatre items sont supprimés sur la base de ces critères. On mesure ensuite l'*infit* et l'*outfit* (suppression des items dont l'*infit* et l'*outfit* ne sont pas compris en 0,5 et 1,5), aucun item n'est supprimé sur cette base. Enfin, on vérifie que la mesure de la difficulté est toujours inférieure à 4 en valeur absolue.

Pour chaque enseignant, un indicateur est estimé selon un modèle de réponse à l'item à deux paramètres (la difficulté et la discrimination, comme pour les élèves). L'indicateur est égal au paramètre de compétence de ce modèle, qui est estimé par maximum de vraisemblance.

6 Taux de participation

Les taux de participation aux différentes enquêtes et évaluations sont détaillés dans le tableau 8. Ils sont calculés après le nettoyage des données brutes (suppression des observations ayant un taux de réponse trop faible, un temps de passation trop court, etc.).

Des modèles logistiques sont construits afin de déterminer les variables observables influençant la participation au premier temps de mesure. Les variables incluses dans le modèle sont des variables individuelles le cas échéant (âge, sexe et discipline pour l’enseignant ; sexe, lieu de naissance et PCS du représentant légal 1 pour l’élève), ainsi que des variables relatives à l’établissement ou à la zone géographique : le type d’équipements numériques mobiles disponibles, le secteur, le réseau et l’académie, ainsi que la taille de l’unité urbaine, la composition sociale du collège, le nombre d’enseignants, le taux de redoublement en 5^e, la note moyenne à l’épreuve terminale du DNB, le taux de retard en 6^e, l’effectif de 5^e, le nombre d’élèves par division, l’âge moyen des enseignants, le taux de chômage de la commune en 2017 et le taux de pauvreté de la zone économique en 2015.

Pour les enseignants, les résultats indiquent que ce sont avant tout les variables individuelles qui influencent la participation des enseignants, notamment leur discipline et l’âge. Les enseignants des disciplines scientifiques et les plus jeunes participent davantage que les autres. On note également que l’appartenance au groupe 2, à la PCS moyenne dans l’établissement, le taux de retard en 6^e et le taux de pauvreté de la zone économique sont corrélés négativement avec la probabilité de participer des enseignants.

Concernant les collèges, peu de variables observées incluses dans la régression logistique n’ont d’impact sur la participation de l’établissement aux évaluations des élèves, hormis la composition sociale de l’établissement (PCS très favorisée et moyenne), l’appartenance au secteur public hors éducation prioritaire et le taux de pauvreté de la zone économique, qui diminuent la probabilité de participation.

Chez les élèves, la variable expliquée est la participation à l’épreuve de mathématiques. Le sexe et la PCS du représentant légal 1 ont un impact significatif, les garçons et les élèves issus de milieu défavorisé ayant une probabilité plus faible de participer. De plus, les élèves des collèges en secteur public hors EP ont également une probabilité plus faible de participer à l’évaluation. On observe par ailleurs que les élèves des établissements des groupes 1 et 3 répondent moins souvent. La taille de l’unité urbaine ainsi que la taille de l’établissement sont également corrélées à la participation des élèves.

Lors du premier temps de mesure, un souci technique a empêché entre 800 et 1 000 élèves de 5^e de répondre à tout ou partie des évaluations sur la compréhension écrite, la créativité, la collaboration et l’esprit critique. Les élèves étaient orientés en fonction de leur réponse à la première question de l’évaluation, portant sur les équipements numériques : les élèves déclarant disposer d’EIM étaient aléatoirement redirigés vers la suite de l’évaluation ou à sa toute fin, sans pouvoir répondre aux

évaluations citées ci-dessus. Les élèves des collèges du groupe 1 (avec EIM) ont donc une probabilité plus faible d’avoir répondu entièrement à ces évaluations. Toutefois, l’erreur technique ayant été aléatoire, elle n’impacte pas en moyenne les estimations sur les scores des élèves.

Les taux de non-réponse élevés sont potentiellement problématiques. Dans tous les cas, cette non-réponse devrait affecter la précision des coefficients estimés. De plus, si celle-ci n’est pas aléatoire, elle risque même de biaiser ces derniers et donc la validité interne et externe de nos résultats. Pour les enseignants, il est probable que la non-réponse ne soit pas aléatoire mais déterminée par une variable non-observée. En effet, la non-réponse est plus faible chez les enseignants jeunes et dans les disciplines scientifiques. On peut donc supposer que l’intérêt pour la thématique enquêtée, soit les technologies numériques, influence à la fois la participation des enseignants et leurs réponses. Le risque semble plus faible pour les élèves, en raison des modalités de passation des évaluations : si les enseignants reçoivent un lien de connexion sur leur adresse électronique utilisée à des fins professionnelles et choisissent de répondre ou non à l’enquête, les élèves sont eux interrogés dans l’établissement, pendant les heures de cours et encadrés par un personnel du collège. Leur possibilité de ne pas répondre est donc limitée.

De prochaines analyses pourront se concentrer sur cette question de la non-réponse, en étudiant notamment les solutions élaborées dans la littérature pour réduire les biais éventuellement induits.

TABLE 8 – Taux de participation

		Premier temps de mesure			Deuxième temps de mesure		
		Effectifs concernés	Nombre de répondants	%	Effectifs concernés	Nombre de répondants	%
Enseignants		2553	1767	69	5548	2692	49
Élèves	Collèges (organisation d'au moins une épreuve)	217	165	76	217	191	88
	Mathématiques et compétences numériques		3482	81		4051	89
	Compréhension orale, esprit critique, créativité et collaboration		3653	85			
	Compréhension écrite	5630 dont 4287 dans les collèges ayant organisé au moins une épreuve	2786	65	5203 dont 4570 dans les collèges ayant organisé au moins une épreuve	4117	90

TABLE 9 – Répartition des taux de participation en fonction des groupes ELAINE

	Groupe 1 Equipements individuels		Groupe 2 Equipements partagés dans le cadre du PN		Groupe 3 Pas d'équipement mais sélectionné dans le PN		Groupe 4 Equipements partagés en dehors du PN		Groupe 5 Pas d'équipement et pas sélectionné dans le PN	
	Non-répondants	Répondants	Non-répondants	Répondants	Non-répondants	Répondants	Non-répondants	Répondants	Non-répondants	Répondants
Etablissements										
Evaluation des élèves	28	72	22	78	16	84	26	74	23	77
Enquête des enseignants	0	100	0	100	0	100	0	100	2	98
Elèves	11	89	8	92	18	82	6	94	7	93
Enseignants	32	68	34	66	28	72	32	68	25	75

Note :

Les établissements sont considérés comme répondants à l'évaluation des élèves ou à l'enquête des enseignants quand au moins un élève ou un enseignant, respectivement, a répondu. Un élève est considéré comme répondant quand on dispose de son score à au moins une épreuve. Le taux de participation des élèves est calculé sur la base des établissements considérés comme répondants à l'évaluation des élèves.

7 Résultats sur l'utilisation du numérique par les élèves

7.1 Accès à des équipements mobiles et utilisation en 5^e

Les élèves de 5^e scolarisés dans des collèges dotés d'équipements numériques mobiles (groupes 1, 2 et 4) sont nettement plus nombreux à déclarer avoir accès à une tablette au collège ou en dehors du collège, qu'elle leur appartienne ou non (92 % pour les élèves des collèges dotés d'EIM et 74 % pour ceux des collèges dotés de CM contre 69 % pour ceux des collèges non équipés). Par ailleurs, respectivement 82 %, 61 % et 55 % des élèves ont accès à une tablette et l'utilise. Cet écart est associé, au moins pour partie, à des différences dans la nature des équipements disponibles dans les collèges : 49 % des élèves scolarisés dans des établissements disposant d'EIM (groupe 1) déclarent avoir accès à une tablette appartenant au collège, alors qu'ils ne sont que 15 % dans les collèges dotés de CM (groupes 2 et 4) et moins de 5 % dans les établissements non équipés (groupes 3 et 5).

En revanche, on ne constate pas d'écart concernant l'accès à un ordinateur : 94 % des élèves scolarisés dans des établissements disposant d'EIM (groupe 1) ou de CM (groupes 2 et 4) déclarent avoir accès à un ordinateur, contre 93 % des élèves scolarisés dans des établissements non équipés (groupes 3 et 5). Enfin, l'écart concernant l'accès à un smartphone est faible et n'est significatif que pour les élèves disposant d'EIM comparativement aux élèves scolarisés dans les établissements non équipés (respectivement 84 % et 82 %). Ces écarts ne sont pas surprenants, au regard des équipements distribués dans le cadre du Plan numérique.

TABLE 10 – Accès aux équipements numériques des élèves de cinquième au collège ou en dehors du collège, en fonction des équipements disponibles dans leur collège

	% ayant accès à une tablette	Écart avec les élèves des collèges sans équipement numérique mobile	% ayant accès à un ordinateur	Écart avec les élèves des collèges sans équipement numérique mobile	% ayant accès à un smartphone	Écart avec les élèves des collèges sans équipement numérique mobile
EIM	91.8	23.1 ***	93.6	0.6	84.4	2.8 *
CM	73.7	5 ***	94.1	1.0	80.8	-0.8

Note : Les niveaux de significativité statistique sont indiqués par les symboles suivants : * pour 10 %, ** pour 5 %, *** pour 1 %.

Champ : Les élèves de quatrième interrogés sur leur accès aux équipements numériques et leur utilisation en cinquième lors de l'enquête ELAINE 2019.

Source : MENJS-DEPP

7.2 Accès, utilisation et types d'usages des équipements numériques en 4^e

En classe de 4^e, ces mêmes élèves ont été interrogés à propos de leurs équipements numériques mobiles, ainsi que la fréquence et le type d'utilisation de ces équipements, au collège ou en dehors du collège. Ainsi, 74 % des élèves scolarisés en 4^e dans les établissements disposant d'équipements individuels mobiles déclarent s'être vus attribuer un ordinateur portable ou une tablette à usage individuel. Parmi ces élèves, la part d'élèves utilisant ces équipements en dehors du collège au moins une fois par semaine est relativement faible (43 %). De même, la part d'élèves utilisant en classe un équipement numérique souvent ou très souvent est plutôt faible, mais elle est statistiquement plus élevée pour les élèves scolarisés dans des collèges disposant d'équipements individuels mobiles ou de classe mobile, comparativement aux élèves scolarisés dans des établissements sans équipement numérique mobile, que cet usage en classe soit individuel ou collectif.

TABLE 11 – Utilisation des équipements numériques par les élèves de quatrième au collège, en fonction des équipements disponibles dans leur collège

	% utilisant souvent ou très souvent un équipement numérique individuellement en classe	Écart avec les élèves des collèges sans équipement numérique mobile	% utilisant souvent ou très souvent un équipement numérique collectivement en classe	Écart avec les élèves des collèges sans équipement numérique mobile
EIM	27.3	17.8 ***	18.8	9.7 ***
CM	12.9	3.4 ***	13.5	4.4 ***

Note : Les niveaux de significativité statistique sont indiqués par les symboles suivants : * pour 10 %, ** pour 5 %, *** pour 1 %.

Champ : Les élèves de quatrième interrogés lors de l'enquête ELAINE 2019.

Source : MENJS-DEPP

L'analyse par régression logistique multivariée indique que les élèves scolarisés dans des établissements disposant d'équipements individuels mobiles ou de classe mobile ne déclarent pas utiliser plus fréquemment des équipements numériques en dehors du collège (qu'il s'agisse de smartphone, d'ordinateur ou de tablette et que cet équipement appartienne à l'élève, à sa famille ou à un ami). De même, il n'y a pas d'écart significatif concernant la fréquence de certains types d'activités, au collège ou en dehors du collège, que ces activités soient scolaires et non, à l'exception de la consultation de l'ENT (Espace Numérique de Travail) plus fréquente chez les élèves scolarisés dans les établissements disposant d'équipements individuels mobiles, notamment pour obtenir des informations à propos de l'emploi du temps, et dans une moindre mesure communiquer avec l'enseignant à propos du travail scolaire.

Si ces premiers éléments sont descriptifs, la suite des analyses vise à établir un lien causal entre la disponibilité des équipements numériques mobiles et les pratiques d'enseignement, d'une part, et

entre la disponibilité de ces équipements et les résultats des élèves, d'autre part.

8 Résultats sur les indicateurs des enseignants

Encadré : la standardisation

La standardisation est une pratique usuelle dans la littérature scientifique, consistant à centrer et réduire les données. Les données, ici chaque score ou indicateur, sont transformées par la soustraction d'une moyenne de référence et la division par l'écart-type de référence. Ici, la référence est le groupe de comparaison. Ainsi, la moyenne dans le groupe de comparaison est égale à 0 et l'écart-type à 1, ce qui permet ensuite d'exprimer les effets estimés en pourcentage d'écart-type du score ou indicateur du groupe de comparaison. Cette transformation a pour but de rendre la comparaison des distributions des scores ou indicateurs plus aisée.

Pour l'ensemble des parties suivantes, les indicateurs des enseignants et les scores des élèves sont standardisés en suivant la procédure décrite ci-dessus. Pour faciliter l'interprétation des effets dits "standardisés" sur les scores des élèves, ils peuvent être convertis en termes de progression de l'élève médian dans une classe : à titre d'illustration, en moyenne, un effet de 20% d'un écart-type signifie qu'un élève initialement médian (classé 13^e dans une classe de 25 élèves) atteint en moyenne le niveau de celui qui était classé 10^e.

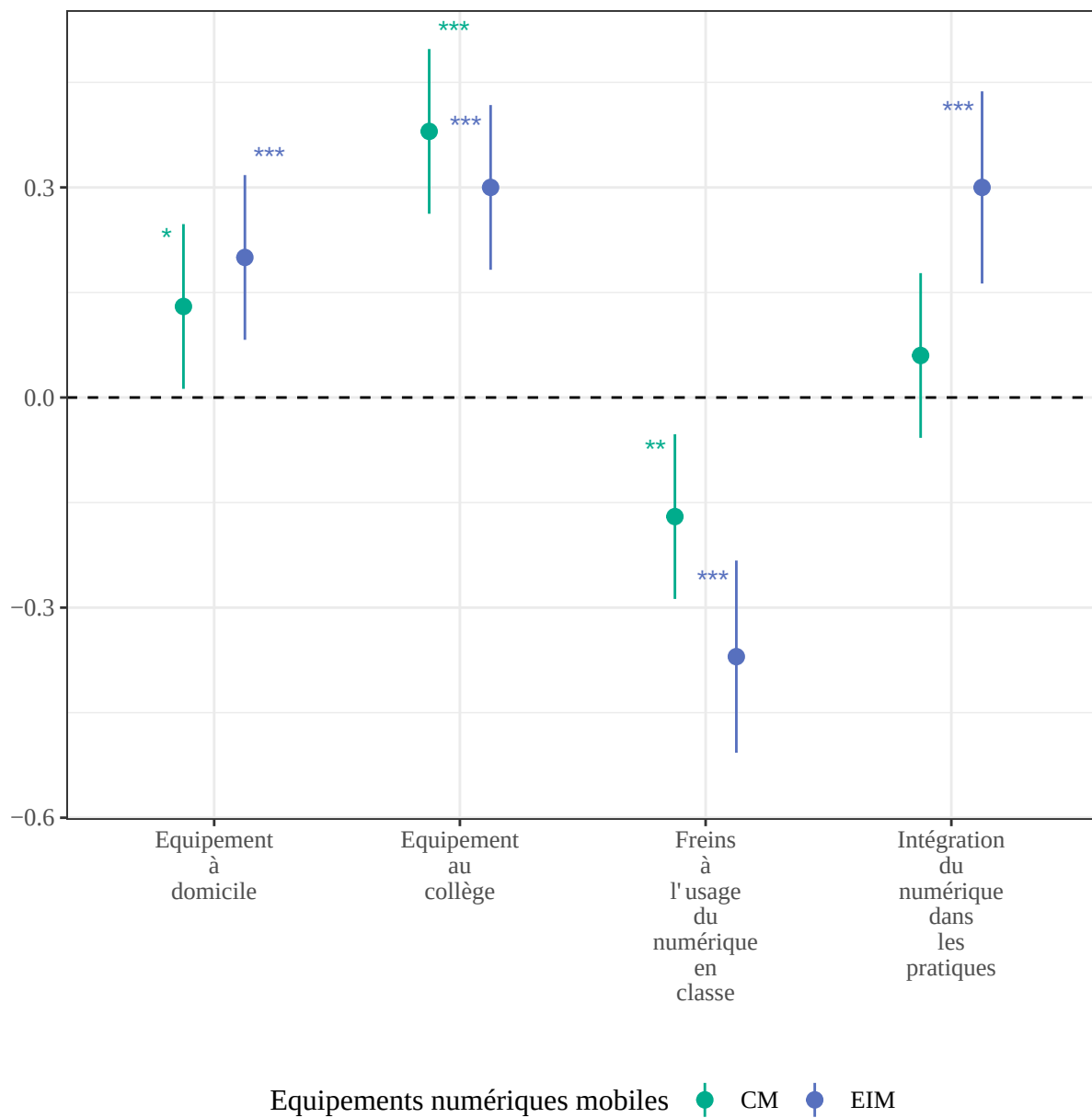
Les enseignants exerçant dans les établissements disposant d'EIM (groupe 1) et de CM (groupes 2 et 4) indiquent des niveaux d'équipement numérique au collège, d'une part, et à domicile, d'autre part, supérieurs à ceux des autres enseignants. Cela est toujours vrai, que les EIM et les CM aient été financés par le Plan numérique ou non. Ainsi, 72 % des enseignants des collèges disposant d'EIM ont accès à une tablette à domicile, et 77 % au collège, contre respectivement 53 % et 37 % des enseignants des collèges sans équipement numérique mobile. Cela pourrait s'expliquer d'une part par la volonté des établissements d'équiper également les enseignants, quand les élèves recevaient des équipements, et d'autre part par un souci des enseignants de disposer à domicile d'outils proches de ceux qu'ils utilisent avec les élèves en classe.

La disponibilité d'EIM en classe diminue les freins perçus par les enseignants aux usages pédagogiques du numérique : ceux-ci sont significativement moins importants pour ceux exerçant dans des collèges disposant d'EIM (groupe 1), que pour leurs pairs des collèges des groupes 3 et 5, sans équipement numérique mobile (effets de 37 % à 39 % d'écart-type). Par exemple, les enseignants du groupe 1 sont moins nombreux que ceux des groupes 3 et 5 à considérer que les contraintes horaires de la discipline enseignée constituent un frein à l'utilisation du numérique dans leur enseignement (52 % contre 59 %). Cet effet s'observe également, mais dans une moindre mesure, pour les enseignants exerçant dans des collèges disposant de CM (effets de 17 à 19 % d'écart-type). À nouveau, cela est toujours vrai, que l'on compare les établissements disposant d'équipements financés par le Plan numérique ou non, comparativement aux établissements sans équipement. Par ailleurs, on peut noter que cet effet est plus important pour les enseignants avec le plus d'ancienneté.

La disponibilité d'EIM (groupe 1) conduit les enseignants à intégrer davantage le numérique dans leurs pratiques professionnelles, incluant l'usage du numérique en classe (par exemple, en faisant travailler les élèves en autonomie avec les outils numériques). Cet effet est significatif quel que soit le groupe de comparaison : établissements disposant de classe mobile ou sans équipement, faisant partie du Plan numérique ou non. La taille de l'effet varie de 30 % à 37 % d'écart-type de l'indicateur de pratiques numériques du groupe de comparaison. À titre d'illustration, les enseignants des collèges où les élèves disposent d'EIM (groupe 1) sont 33 % à monter au moins une fois par semaine de séquences d'activités en classe avec manipulation de matériel numérique par les élèves, contre 22 % dans les collèges sans équipement numérique mobile (groupes 3 et 5). La disponibilité de CM reste en revanche sans effet sur l'usage du numérique dans les pratiques enseignantes documentées par ELAINE.

La disponibilité d'EIM (groupe 1) n'impacte pas le sentiment de compétences des enseignants renseigné par l'échelle TPACK au travers des six indicateurs détaillés dans la section 5.2.1. Les enseignants se sentent majoritairement efficaces, qu'ils enseignent dans des établissements dotés d'équipements numériques mobiles ou non. Par exemple, 58 % des enseignants des collèges disposant d'EIM (groupe 1) indiquent être d'accord ou très d'accord avec la proposition suivante : « *Je sais choisir des technologies numériques pour améliorer le contenu d'un cours* ». Ils sont 55 % parmi les enseignants des collèges disposant de CM (groupes 2 et 4) et 57 % dans les collèges sans équipement numérique mobile (groupes 3 et 5). Ces écarts ne sont pas significatifs.

À cet égard, il convient de noter que les enseignants exerçant dans des collèges dotés d'EIM n'ont pas bénéficié d'un surcroît significatif de formation au numérique. Dans le groupe 1, ils sont 65 % à déclarer avoir suivi au moins une formation continue sur le numérique et l'informatique en lien avec l'enseignement. Cette proportion est identique pour leurs collègues des collèges sans équipement numérique mobile (groupes 3 et 5). Il faut toutefois noter que nous ne disposons pas d'information sur la temporalité de ces formations, qui peuvent dater d'avant le déploiement du Plan numérique de 2015. La relation causale entre ce plan et les formations ne peut donc pas être établie. De plus, on ne note pas d'effet des EIM ou des CM sur les compétences numériques évaluées chez les enseignants (maîtrise technique des outils, communication, pensée algorithmique, recherche d'information, etc.), à l'exception d'un léger effet positif des EIM lorsque l'on restreint l'échantillon aux hommes. Enfin, les enseignants ont été interrogés sur la fréquence et la nature de la collaboration entre enseignants et sur leurs pratiques visant à favoriser la collaboration entre élèves. Alors que la disponibilité des outils numériques aurait pu inciter les enseignants à échanger davantage entre eux et à développer le travail collaboratif entre élèves, on ne distingue aucun effet. De même, on ne relève pas d'impact significatif de la distribution d'équipements numériques mobiles sur le sentiment de créativité déclaré par les enseignants.



Note : la barre représente l'intervalle de confiance à 95 %.

Les niveaux de significativité statistique sont indiqués par les symboles suivants :

* pour 10 %, ** pour 5 % et *** pour 1 %.

FIGURE 2 – Effet moyen estimé de la disponibilité des EIM et des CM sur les indicateurs des enseignants (en % d'écart-type)

9 Résultats sur les scores des élèves

9.1 Scores des élèves en 5^e

La mise à disposition d'EIM (groupe 1), via le Plan numérique ou non, a un effet positif sur le niveau moyen des élèves de 5^e en compréhension orale du français. Cet effet est compris entre 17 % et 31 % d'écart-type selon le groupe de comparaison, ce qui correspond à une progression de 2 à 3 rangs dans la classe. La mise à disposition de CM (groupes 2 et 4) a également un effet positif de 7 % d'écart-type sur la compréhension orale (progression d'un rang), comparativement aux élèves des établissements sans équipement numérique mobile (groupes 3 et 5)¹⁷.

Des analyses complémentaires ont porté sur l'identification d'effets différenciés selon le sexe des élèves, leur origine sociale et la composition sociale et scolaire des établissements (regroupés en fonction de la PCS des familles, du taux de redoublement en 6^e, du résultat moyen à l'écrit de l'épreuve terminale du DNB, etc.). On observe des effets différenciés des EIM sur la compréhension orale du français selon le sexe (écart de 24 points de pourcentage d'écart-type en faveur des filles). On n'observe pas d'écart significatif selon l'origine sociale des élèves ou la composition sociale et scolaire des collèges. Pour les CM, aucun effet différencié n'est identifié.

Les EIM ont un faible effet sur les compétences numériques des élèves (9 % d'écart type, soit 1 rang) en 5^e¹⁸. Enfin, on ne mesure aucun effet des EIM (groupe 1) ou des CM (groupes 2 et 4) sur le niveau moyen des élèves en mathématiques et en compréhension écrite du français en 5^e.

9.2 Scores des élèves en 4^e

L'effet moyen des EIM (groupe 1) sur la compréhension orale semble perdurer en 4^e mais de manière plus modeste (14 % d'écart-type, soit un gain d'un rang). Cependant, on n'identifie plus de manière systématique un effet différencié selon le sexe de l'élève ou la composition de l'établissement. Par ailleurs, la distribution de CM n'a plus d'effet en 4^e sur la compréhension orale.

En 4^e, les EIM et les CM ont un impact positif modéré sur le score moyen en mathématiques (respectivement 14 % et 13 % d'écart-type, soit 1 rang) et sur les compétences numériques (respectivement 8 % et 9 %, soit 1 rang). Par exemple, 46 % des élèves des collèges disposant d'EIM (groupe 1) identifient correctement les informations pouvant figurer dans un courriel, contre 39 % des élèves des collèges sans équipement numérique mobile (groupes 3 et 5). De même, ceux du groupe 1 sont 42 % à identifier la bonne définition du droit à l'opposition (contre 37 % pour ceux des groupes 3 et 5)¹⁹.

Pour les EIM, les effets varient très fortement en fonction de la composition sociale et scolaire des

17. Toutefois, on n'identifie pas d'effet des CM lors des analyses par IPW.

18. Toutefois, il faut noter que les résultats des analyses par IPW sont différents.

19. Toutefois, il faut noter que l'effet des EIM sur le score de mathématiques et de compétences numériques n'est pas statistiquement significatif lors de l'analyse par IPW.

établissements (écarts de 40 à 64 points de pourcentage en faveur des élèves des collèges défavorisés pour les mathématiques et de 61 à 71 points de pourcentage pour les compétences numériques). Les EIM ont également un effet différencié selon l'origine sociale individuelle des élèves sur les compétences numériques (entre 37 et 45 points de pourcentage en faveur des élèves issus de familles de PCS défavorisées), mais pas sur les compétences en mathématiques. Enfin, il n'y a pas d'écart en fonction du sexe.

Par ailleurs, seuls les EIM ont un impact modéré mais positif sur la compréhension écrite (12 % d'écart-type, soit 1 rang). Cet effet ne varie pas selon les différentes catégories d'élèves décrites précédemment.

9.3 Tests de robustesse

Pour tester la robustesse de nos résultats, nous avons dupliqué ces analyses en construisant les groupes de comparaison non plus à l'aide des déclarations des chefs d'établissement, mais en utilisant les déclarations des élèves eux-mêmes, décrites précédemment de ce document. De plus, l'ensemble des analyses portant sur les compétences des élèves ont été réalisées en utilisant non plus les scores standardisés des élèves mais leurs scores bruts (ici, le pourcentage de bonnes réponses à chaque évaluation). Dans les deux cas, les conclusions sont identiques.

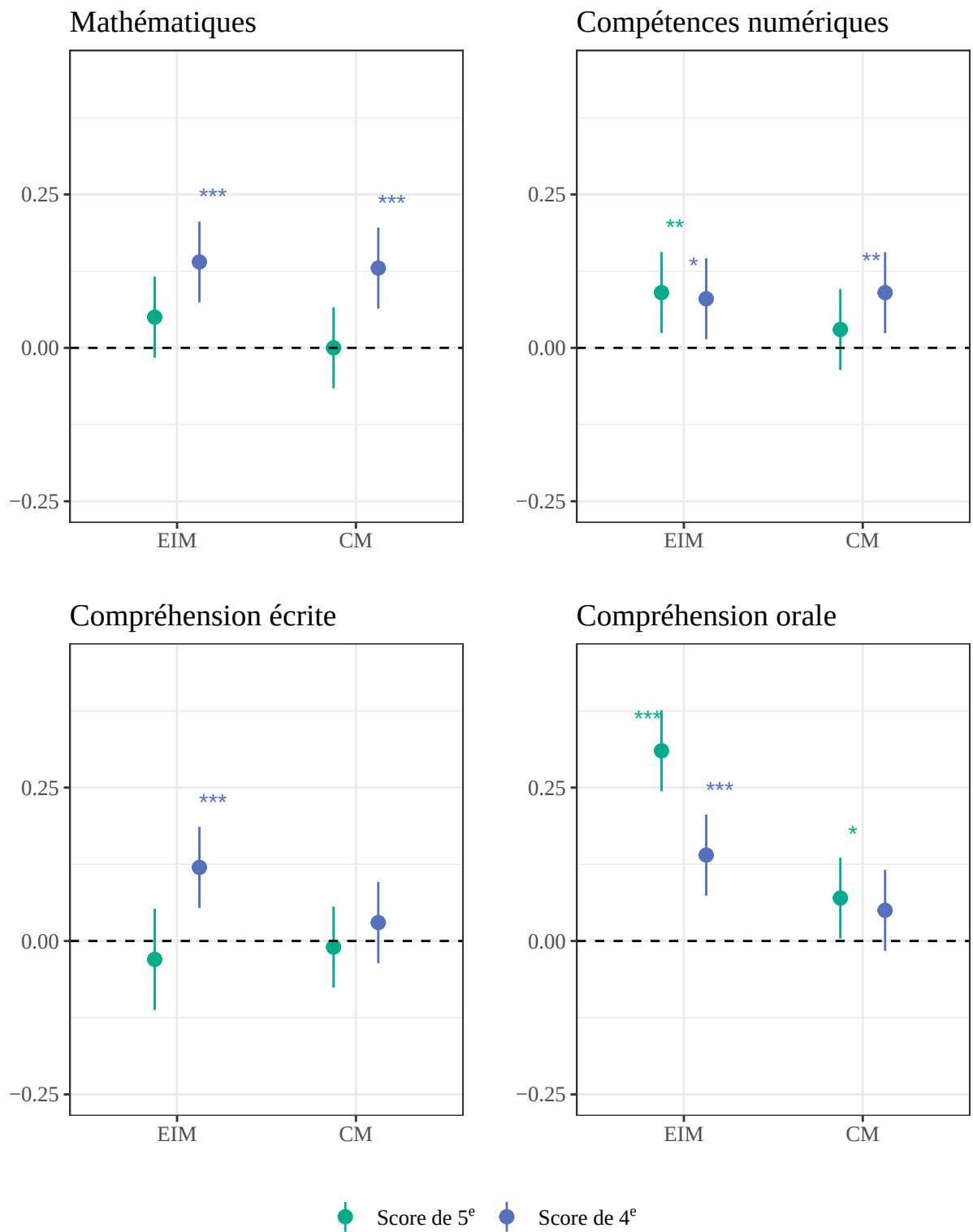


FIGURE 3 – Effet moyen estimé de la disponibilité des EIM et des CM sur les compétences disciplinaires et numériques des élèves (en % d'écart-type)

10 Conclusion

Les effets estimés de la disponibilité d'EIM et de CM sur les apprentissages des élèves sont globalement positifs mais modérés. Cela peut s'expliquer par les changements modestes intervenus dans les pratiques enseignantes et dans l'utilisation des outils par les élèves en dehors du collège. On peut supposer que l'impact le plus fort concerne la compréhension orale du français car les possibilités offertes par les outils numériques avec l'accès à des ressources difficilement mobilisables par d'autres biais (comme par exemple les vidéos), tandis qu'on peut supposer que les outils numériques n'ont qu'une plus-value faible pour les documents écrits, en comparaison aux outils classiques comme les manuels scolaires. Cette hypothèse est étayée par la revue de littérature réalisée par Tricot (2020) qui indique que l'usage du numérique produit des effets plus positifs pour les activités pédagogiques telles « regarder une vidéo » que pour les activités pédagogiques telles que « lire et comprendre un texte ». Par ailleurs, les faibles effets mesurés sur les compétences numériques peuvent sembler contre-intuitifs. Une analyse des résultats de l'enquête ICILS en France (sur un autre échantillon que l'étude ELAINE) pourra apporter des pistes de compréhension, en comparant les résultats en littératie numérique et pensée informatique des élèves des établissements faisant partie du Plan numérique aux autres élèves.

Les effets particulièrement faibles des CM sur les scores des élèves peuvent être mis en parallèle avec les taux peu élevés d'accès et d'utilisation des tablettes au collège pour les établissements des groupes 2 et 4. La seule dotation en équipements matériels partagés du Plan numérique de 2015 ne semble pas suffisante pour impacter sensiblement les usages numériques des enseignants et des élèves, ainsi que les apprentissages scolaires.

Le Plan numérique a également concerné 6 807 écoles primaires publiques en France métropolitaine et DROM²⁰, avec la distribution d'équipements sous la forme de classe mobile. Or, les effets de la distribution d'équipements numériques mobiles sur les compétences des élèves peuvent varier selon le stade de leur scolarité. C'est pourquoi un autre volet de l'étude ELAINE portera sur le suivi d'élèves du CM1 à la 6^e, de novembre 2020 à novembre 2022. Ces élèves de CM1 sont répartis dans 300 écoles primaires publiques en France métropolitaine et DROM. L'étude intégrera également leurs enseignants, ainsi que les directeurs et les IEN des écoles concernées. S'ajouteront à ces panels des enseignants, des directeurs d'écoles et des IEN de 2 000 autres écoles, afin de disposer d'échantillons suffisamment larges pour identifier de potentiels effets. Le suivi des élèves jusqu'à leur entrée en 6^e permettra d'étudier plus précisément la période charnière du passage de l'école primaire au collège.

Enfin, l'étude ELAINE est dotée d'un volet qualitatif complémentaire des analyses quantitatives. Les monographies se dérouleront en novembre et décembre 2020 dans 4 collèges de l'échantillon. Par des observations en classe, des *focus groups* et des entretiens individuels, ces monographies permettront notamment d'éclairer les mécanismes par lesquels les élèves et les enseignants peuvent s'appropriier ou non les outils numériques, pour les intégrer ou non dans les apprentissages et les enseignements.

20. ETIC (DEPP)

Elles s'intéresseront également à l'organisation et au fonctionnement de l'établissement pouvant influencer la culture, les usages et les compétences numériques des personnels et des élèves.

Si l'étude ELAINE a été élaborée avec l'objectif de répondre à un vaste ensemble de questions de recherche relatives à l'impact du numérique sur les apprentissages des élèves, de nombreuses thématiques sont encore à explorer. Tout d'abord, la relation causale entre pratiques d'enseignement et effet des outils sur les scores reste à étudier. Est-ce par l'appropriation de ces outils par les enseignants que les compétences des élèves sont impactées ? Un autre enjeu central est celui de la motivation des élèves : l'utilisation des outils numériques en classe est-il un moyen de mobiliser l'attention des élèves en classe ? Les incite-t-elle à travailler d'avantage, à la maison ou à l'école ? Ces outils sont-ils une source d'attrait pour les élèves ? Ou au contraire sont-ils propices à susciter la déconcentration et une perte de motivation ?

Ces deux champs pourront permettre de mieux comprendre les mécanismes à l'œuvre dans le cadre de la distribution et de l'utilisation d'équipements numériques à l'école.

Annexes

L'équilibrage par *entropie balancing* : principe et méthodes

L'équilibrage par entropie est une procédure préalable (*preprocessing*) de traitement des données qui permet de produire des échantillons équilibrés avant même de mettre en œuvre des techniques d'estimation d'impact causal, notamment des techniques d'appariement d'échantillons (*matching*). Cette procédure consiste à repondérer les observations du groupe de contrôle en leur affectant des poids scalaires de sorte qu'après repondération, les premiers moments (moyenne, variance, coefficient d'asymétrie, etc.) de la distribution des covariables dans le groupe de contrôle soient les mêmes (tout au moins, le plus proche possible) que ceux de la distribution de ces mêmes covariables dans le groupe de traitement²¹. Les poids qui résultent de l'équilibrage par entropie peuvent être utilisés pour d'autres approches expérimentales ou quasi-expérimentales, par exemple dans le cadre d'une expérimentation randomisée, d'une analyse par doubles différences ou par régression avec discontinuité.

Notons $Y(1)$ la variable de résultat (ici, le score à un test standardisé) lorsqu'un ou une élève bénéficie d'un équipement numérique grâce au Plan numérique, et $Y(0)$ la même variable de résultat lorsqu'il ou elle n'en bénéficie pas. Notons ensuite D la variable binaire qui indique si l'élève est dans un établissement équipé grâce au Plan : $D = 1$ si c'est le cas, $D = 0$ sinon. L'objectif est d'estimer au mieux l'effet moyen du traitement (i.e., le fait de disposer d'un équipement numérique dans l'établissement) pour ceux des élèves qui en bénéficient effectivement (i.e., pour les traités). Cet effet, noté ATT (comme « *Average Treatment effect for the Treated* ») est formellement défini par l'espérance conditionnelle suivante :

$$ATT = E[Y(1)|D = 1] - E[Y(0)|D = 1]$$

L'espérance conditionnelle $E[Y(0)|D = 1]$ n'est pas identifiable puisque, pour les élèves du groupe de traitement, le score qu'ils auraient obtenu sans pouvoir utiliser d'équipement numérique dans le cadre scolaire (« le contrefactuel ») n'est pas observé. Le principe est d'estimer ce contrefactuel par la moyenne des scores des élèves du groupe de contrôle dont les caractéristiques, représentées par les covariables observables dans l'échantillon (sexe, PCS des représentants légaux, caractéristiques des établissements, etc.), sont « proches ». Cette estimation est donnée par la formule suivante :

$$E[Y(0)|\hat{D} = 1] = \frac{\sum_{i|D=0} Y_i w_i}{\sum_{i|D=0} w_i}$$

Dans cette expression, w_i est la pondération appliquée à l'élève i du groupe de contrôle représenté par la notation $\{i|D = 0\}$, et Y_i est le score obtenu par l'élève i du groupe de contrôle. Les pondérations

21. Voir Hainmueller, J. (2012) "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies", *Political Analysis*, 20(1), 25-46, et Hainmueller, J. and Xu, Y. (2013), "Ebalance: A Stata Package for Entropy Balancing", *Journal of Statistical Software*, 54(7), 1-18.

sont obtenues en minimisant la fonction $H(w)$ définie ci-après :

$$\min_{w_i} H(w) = \sum_{i|D=0} h(w_i)$$

sous les contraintes d'équilibrage et de normalisation suivantes :

$$\sum_{i|D=0} w_i c_{ri}(X_i) = m_r$$

avec $r \in 1, \dots, R$

$$\sum_{i|D=0} w_i = 1$$

et

$$w_i \geq 0$$

pour tout i tel que $D = 0$.

Dans les équations ci-dessus, $h(\cdot)$ est une fonction de perte et $c_{ri}(X_i) = m_r$ représente un ensemble de R contraintes d'équilibrage que l'on impose aux moments des distributions des covariables des élèves du groupe de contrôle, repondérées de la façon suivante :

1. La fonction de perte $h(\cdot)$ est la divergence de Kullback-Leibler définie par $h(w_i) = w_i \log(\frac{w_i}{q_i})$, où w_i est la pondération estimée pour l'élève i du groupe de contrôle et q_i est une pondération de base, ici choisie comme étant égale à $q_i = \frac{1}{n_0}$, n_0 étant le nombre d'élèves dans le groupe de contrôle. La fonction de perte $h(\cdot)$ mesure la distance entre la distribution des pondérations de contrôle estimées, définie par le vecteur $W = [w_1, \dots, w_{n_0}]'$, et la distribution des pondérations de base, spécifiée par le vecteur $Q = [q_1, \dots, q_{n_0}]'$, avec $q_i \geq 0$ pour tout i tel que $D = 0$ et $\sum_{i|D=0} q_i = 1$. Cette fonction de perte est non-négative et décroît lorsque W se rapproche de Q . Elle est égale à 1 lorsque $W = Q$.
2. Les contraintes d'équilibrage exprimées par la première contrainte sont choisies par l'analyste de sorte à égaliser les moments des distributions des covariables au sein des deux groupes, traitement et contrôle. Le terme m_r peut par exemple correspondre au moment d'ordre r de la covariable X_j caractérisant les élèves. En pratique, ces moments sont spécifiés pour les élèves du groupe de contrôle de telle sorte que $c_{ri}(X_{ij}) = X_{ij}^r$ ou bien que $c_{ri}(X_{ij}) = (X_{ij} - \mu_j)^r$ où μ_j est la moyenne de la variable X_j .
3. Les deux contraintes suivantes sont des contraintes de normalisation. La première signifie que

la somme des pondérations doit être égale à 1²². La seconde signifie que chaque pondération doit être positive ou nulle. Ces pondérations ne peuvent être négatives car, si tel était le cas, la fonction de perte $h(\cdot)$, qui est une métrique de distance, ne serait pas définie.

L'équilibrage par entropie peut être considéré comme une extension de la procédure de pondération par le score de propension (*propensity score*) qui est souvent estimé par un modèle de type Logit dans la première étape d'une approche par appariement d'échantillons (*matching*). Toutefois, l'appariement par score de propension impose que le score estimé en première étape satisfasse a posteriori la condition d'équilibrage (*balancing condition*). Or, ce n'est pas toujours le cas : il faut alors spécifier et estimer un nouveau score de propension, et ce jusqu'à ce que celui-ci vérifie la propriété d'équilibrage. Cette procédure d'ajustement ne converge pas nécessairement, c'est là une des difficultés majeures de la mise en œuvre de la technique statistique de *matching*. L'équilibrage par entropie n'a pas ce défaut, car le calcul des pondérations est fait ici préalablement à toute inférence causale, en mettant en œuvre une procédure permettant d'égaliser les moments (le plus souvent, la moyenne, la variance et éventuellement le coefficient d'asymétrie) de la distribution des covariables dans les groupes de contrôle et de traitement. L'approche par appariement sur le score de propension ne garantit pas cette égalisation des premiers moments des distributions des covariables dont les valeurs sont observées dans l'échantillon.

22. Cette somme peut être aussi égale au nombre n_1 d'élèves faisant partie du groupe de traitement.

Régressions logistiques - Participation

TABLE 12: Paramètres estimés du modèle logit visant à expliquer la participation des enseignants au premier temps de mesure de l'enquête ELAINE

	Participation
Constante	3.60** (1.70)
Sexe (réf. : Femme)	
Homme	-0.17 (0.11)
Discipline (réf. : Arts plastiques)	
Education musicale	0.42* (0.22)
Education physique et sportive	-0.04 (0.21)
Français	0.41* (0.22)
Histoire-géographie-EMC	0.32 (0.21)
Langue vivante ou éteinte	0.32* (0.18)
Mathématiques	1.35*** (0.25)
Physique-chimie	0.66*** (0.22)
SVT	0.48** (0.22)
Technologie	1.16*** (0.24)
Âge	-0.03*** (0.01)
Groupe de l'établissement (réf. : Groupe 1 - Equipements individuels)	
Groupe 2 - Equipements partagés et pas sélectionné dans le PN	-0.32* (0.17)
Groupe 3 - Pas d'équipement mais sélectionné dans le PN	-0.06 (0.18)
Groupe 4 - Equipements partagés et pas sélectionné dans le PN	-0.26 (0.18)
Groupe 5 - Pas d'équipement et pas sélectionné dans le PN	0.09 (0.18)

	Participation
Macro-strate (réf. : Etablissement privé)	
Secteur public en éducation prioritaire	−0.09 (0.34)
Secteur public hors éducation prioritaire	−0.26 (0.22)
Académie (réf. : Aix Marseille)	
Amiens	0.57 (0.40)
Besançon	13.65 (251.33)
Bordeaux	−0.45 (0.29)
Caen	0.29 (0.56)
Clermont-Ferrand	0.31 (0.41)
Corse	−0.94* (0.51)
Créteil	−0.17 (0.32)
Dijon	−0.25 (0.36)
Grenoble	−0.38 (0.34)
La Réunion	1.68** (0.68)
Lille	0.12 (0.31)
Limoges	1.42 (1.09)
Lyon	−0.11 (0.31)
Martinique	0.06 (0.80)
Montpellier	0.13 (0.29)
Nancy-Metz	0.39 (0.44)
Nantes	0.09 (0.36)
Nice	−0.16

	Participation
	(0.38)
Orléans-Tours	0.18
	(0.37)
Paris	−0.94**
	(0.48)
Poitiers	0.12
	(0.41)
Reims	0.20
	(0.41)
Rennes	0.04
	(0.33)
Rouen	−0.30
	(0.35)
Strasbourg	−0.27
	(0.40)
Toulouse	0.10
	(0.35)
Versailles	−0.47
	(0.31)
Taille de l'unité urbaine (réf. : Commune rurale)	
Plus de 200 001 hab. ou Agglomération parisienne	−0.10
	(0.31)
Entre 20 001 et 200 000 hab.	0.01
	(0.31)
Entre 2 001 et 20 000 hab.	−0.04
	(0.26)
Pourcentage d'élèves issus de la PCS (réf. : défavorisée)	
Très favorisée	−0.01
	(0.01)
Favorisée	−0.00
	(0.01)
Moyenne	−0.02*
	(0.01)
Effectif ETP enseignants	−0.02
	(0.01)
Taux de redoublement des élèves de cinquième	−0.08
	(0.06)
Note moyenne à l'écrit au brevet	−0.04
	(0.07)
Pourcentage d'élèves en retard à l'entrée en sixième	−0.03**
	(0.02)

	Participation
Effectif d'élèves de cinquième	0.00 (0.00)
Nombre d'élèves par division	0.01 (0.04)
Age moyen des enseignants	0.00 (0.03)
Taux de chômage de la commune en 2017	−0.00 (0.04)
Taux de pauvreté de la zone économique en 2015	−0.03*** (0.01)
AIC	2849.05
BIC	3213.34
Log Likelihood	-1361.53
Deviance	2723.05
Nombre d'obs.	2398
Remarques : Un enseignant est considéré comme participant s'il a répondu à au moins une des trois parties de l'enquête, telles que définies ci-dessus. Les mêmes analyses ont été menées en fonction de la participation à chacune des parties du questionnaire, il n'y a pas de différence significative dans les résultats, la participation s'expliquant principalement par l'âge et la discipline des enseignants.	
Sources : DNE (2017), APAE (2017-2018).	
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	

TABLE 13: Paramètres estimés du modèle logit visant à expliquer la participation des établissements à l'évaluation des élèves au premier temps de mesure d'ELAINE

	Participation
Constante	10.52 (7.97)
Groupe de l'établissement (réf. : Groupe 1 - équipements individuels)	
Groupe 2 - Equipements partagés et pas sélectionné dans le PN	0.03 (0.72)
Groupe 3 - Pas d'équipement mais sélectionné dans le PN	1.07 (0.87)
Groupe 4 - Equipements partagés et pas sélectionné dans le PN	0.04 (0.79)
Groupe 5 - Pas d'équipement et pas sélectionné dans le PN	-0.46 (0.83)
Macro-strate (réf.: Etablissement privé)	
Secteur public en éducation prioritaire	-1.52 (1.53)
Secteur public hors éducation prioritaire	-1.74* (1.00)
Académie (réf. : Aix Marseille)	
Amiens	0.73 (1.49)
Besançon	20.35 (10754.01)
Bordeaux	-0.17 (1.14)
Caen	19.84 (5597.97)
Clermont-Ferrand	0.40 (1.50)
Corse	21.05 (7486.26)
Créteil	1.82 (1.53)
Dijon	20.14 (3942.58)
Grenoble	1.63 (1.75)
La Réunion	22.25

	Participation
	(5255.66)
Lille	2.56
	(1.61)
Limoges	19.44
	(10754.61)
Lyon	0.62
	(1.33)
Martinique	23.20
	(10753.99)
Montpellier	20.42
	(2626.73)
Nancy-Metz	1.03
	(1.69)
Nantes	0.53
	(1.43)
Nice	−0.24
	(1.35)
Orléans-Tours	0.62
	(1.34)
Paris	21.00
	(6046.75)
Poitiers	19.51
	(3896.62)
Reims	20.55
	(4060.13)
Rennes	1.79
	(1.47)
Rouen	1.75
	(1.41)
Strasbourg	20.73
	(4606.11)
Toulouse	4.90**
	(2.23)
Versailles	0.02
	(1.39)
Taille de l'unité urbaine (réf. : Commune rurale)	
Plus de 200 001 hab. ou Agglomération parisienne	0.60
	(1.53)
Entre 20 001 et 200 000 hab.	1.06
	(1.51)
Entre 2 001 et 20 000 hab.	−0.37
	(1.36)

	Participation
Pourcentage d'élèves issus de la PCS (réf. : défavorisée)	
Très favorisée	−0.08* (0.04)
Favorisée	−0.08 (0.05)
Moyenne	−0.10** (0.05)
Effectif ETP enseignants	−0.01 (0.07)
Taux de redoublement des élèves de cinquième	−0.46 (0.30)
Note moyenne à l'écrit au brevet	0.30 (0.29)
Pourcentage d'élèves en retard à l'entrée en sixième	−0.06 (0.08)
Effectif d'élèves de cinquième	0.01 (0.02)
Nombre d'élèves par division	−0.05 (0.16)
Âge moyen des enseignants	−0.11 (0.12)
Population totale de la commune	−0.00 (0.00)
Taux de chômage de la commune en 2017	0.05 (0.21)
Taux de pauvreté de la zone économique en 2015	−0.13*** (0.05)
AIC	254.84
BIC	433.97
Log Likelihood	-74.42
Deviance	148.84
Nombre d'obs.	217

Remarques : Un établissement est considéré comme participant à l'évaluation des élèves si on dispose du score d'au moins un élève pour au moins une évaluation.

Sources : DNE (2017), APAE (2017-2018).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE 14: Paramètres estimés du modèle logit visant à expliquer la participation des élèves à l'évaluation de mathématiques au premier temps de mesure d'ELAINE

	Participation
Constante	1.77*** (0.31)
Sexe (réf. : Fille)	
Garçon	-0.27*** (0.09)
Lieu de naissance (réf. : Né hors de France)	
Né en France	0.28 (0.19)
PCS d'origine du représentant 1 (réf. : Défavorisée)	
Très favorisée	1.04*** (0.16)
Favorisée	0.88*** (0.17)
Moyenne	0.31*** (0.11)
Groupe de l'établissement (réf. : Groupe 1 - Equipements individuels)	
Groupe 2 - Equipements partagés et pas sélectionné dans le PN	0.27** (0.13)
Groupe 3 - Pas d'équipement mais sélectionné dans le PN	-0.43*** (0.13)
Groupe 4 - Equipements partagés et pas sélectionné dans le PN	0.37** (0.17)
Groupe 5 - Pas d'équipement et pas sélectionné dans le PN	0.27* (0.15)
Macro-strate (réf. : Etablissement privé)	
Secteur public en éducation prioritaire	0.14 (0.23)
Secteur public hors éducation prioritaire	-0.30* (0.18)
Taille de l'unité urbaine (réf. : Commune rurale)	
Plus de 200 001 hab. ou Agglomération parisienne	-0.41** (0.20)
Entre 20 001 et 200 000 hab.	-0.29

	Participation
	(0.20)
Entre 2 001 et 20 000 hab.	0.25
	(0.19)
Effectif d'élèves de cinquième	0.01***
	(0.00)
Effectif d'enseignants (ETP)	−0.05***
	(0.01)
AIC	3161.56
BIC	3267.52
Log Likelihood	-1563.78
Deviance	3127.56
Nombre d'obs.	3763
Remarques : Un élève est considéré comme répondant si nous disposons de son score en mathématiques (il a répondu à un nombre suffisant de questions lors de l'évaluation).	
Sources : DNE (2017), APAE (2017-2018), Sysca (2018-2019).	
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	

Bibliographie

- Angrist, Joshua, et Victor Lavy. 2002. « New evidence on classroom computers and pupil learning ». *The Economic Journal* 112 (482): 735-65.
- Barrera-Osorio, Felipe, et Leigh L Linden. 2009. *The use and misuse of computers in education: evidence from a randomized experiment in Colombia*. The World Bank.
- Barrow, Lisa, Lisa Markman, et Cecilia Elena Rouse. 2009. « Technology's edge: The educational benefits of computer-aided instruction ». *American Economic Journal: Economic Policy* 1 (1): 52-74.
- Beltran, Daniel, Kuntal Das, et Robert W Fairlie. 2006. « Do home computers improve educational outcomes? Evidence from matched current population surveys and the national longitudinal survey of youth 1997 ».
- Carter, Susan Payne, Kyle Greenberg, et Michael S Walker. 2017. « The impact of computer usage on academic performance: Evidence from a randomized trial at the United States Military Academy ». *Economics of Education Review* 56: 118-32.
- Chauhan, Sumedha. 2017. « A meta-analysis of the impact of technology on learning effectiveness of elementary students ». *Computers & Education* 105: 14-30.
- Cheung, Alan CK, et Robert E Slavin. 2013. « The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis ». *Educational research review* 9: 88-113.
- Comi, Simona Lorena, Gianluca Argentin, Marco Gui, Federica Origo, et Laura Pagani. 2017. « Is it the way they use it? Teachers, ICT and student achievement ». *Economics of Education Review* 56: 24-39.
- De Witte, Kristof, et Nicky Rogge. 2014. « Does ICT matter for effectiveness and efficiency in mathematics education? » *Computers & Education* 75: 173-84.
- Faber, Benjamin, Rosa Sanchis-Guarner, et Felix Weinhardt. 2015. « ICT and education: Evidence from student home addresses ». National Bureau of Economic Research.
- Fairlie, Robert W, et Peter Riley Bahr. 2018. « The effects of computers and acquired skills on earnings, employment and college enrollment: Evidence from a field experiment and California UI earnings records ». *Economics of Education Review* 63: 51-63.
- Fairlie, Robert W, et Ariel Kalil. 2017. « The effects of computers on children's social development and school participation: Evidence from a randomized control experiment ». *Economics of Education Review* 57: 10-19.
- Fairlie, Robert W, et Rebecca A London. 2012. « The effects of home computers on educational outcomes: Evidence from a field experiment with community college students ». *The Economic Journal* 122 (561): 727-53.

- Fairlie, Robert W, et Jonathan Robinson. 2013. « Experimental evidence on the effects of home computers on academic achievement among schoolchildren ». *American Economic Journal: Applied Economics* 5 (3): 211-40.
- Falck, Oliver, Constantin Mang, et Ludger Woessmann. 2018. « Virtually no effect? Different uses of classroom computers and their effect on student achievement ». *Oxford Bulletin of Economics and Statistics* 80 (1): 1-38.
- Fiorini, Mario. 2010. « The effect of home computer use on children's cognitive and non-cognitive skills ». *Economics of Education review* 29 (1): 55-72.
- Hull, Marie, et Katherine Duch. 2017. « One-to-one technology and student outcomes ».
- Lai, Fang, Renfu Luo, Linxiu Zhang, Xinzhe Huang, et Scott Rozelle. 2015. « Does computer-assisted learning improve learning outcomes? Evidence from a randomized experiment in migrant schools in Beijing ». *Economics of Education Review* 47: 34-48.
- Malamud, Ofer, Santiago Cueto, Julian Cristia, et Diether W Beuermann. 2019. « Do children benefit from internet access? Experimental evidence from Peru ». *Journal of Development Economics* 138: 41-56.
- Malamud, Ofer, et Cristian Pop-Eleches. 2011. « Home computer use and the development of human capital ». *The Quarterly journal of economics* 126 (2): 987-1027.
- Mora, Toni, Josep Oriol Escardibul, et Giorgio Di Pietro. 2018. « Computer and students achievement: An analysis of the One Laptop per Child program in Catalonia] ». *International Journal of Educational Research* 92: 145-57.
- Muralidharan, Karthik, Abhijeet Singh, et Alejandro J Ganimian. 2019. « Disrupting education? Experimental evidence on technology-aided instruction in India ». *American Economic Review* 109 (4): 1426-60.
- Naik, Gopal, Chetan Chitre, Manaswini Bhalla, et Jothisna Rajan. 2020. « Impact of use of technology on student learning outcomes: Evidence from a large-scale experiment in India ». *World Development* 127: 104736.
- Schmitt, John, et Jonathan Wadsworth. 2006. « Is there an impact of household computer ownership on children's educational attainment in Britain? » *Economics of Education review* 25 (6): 659-73.

