



HAL
open science

Étude de validation de la plateforme en ligne TACIT dédiée à l'évaluation et l'apprentissage différencié de la compréhension de l'implicite des textes

Olivier Le Bohec, Christophe Quaireau, Stéphane Deline, Yvonnick Noël,
Jérémy Nogues, Karine Lavandier, Fanny de La Haye-Nicolas

► To cite this version:

Olivier Le Bohec, Christophe Quaireau, Stéphane Deline, Yvonnick Noël, Jérémy Nogues, et al.. Étude de validation de la plateforme en ligne TACIT dédiée à l'évaluation et l'apprentissage différencié de la compréhension de l'implicite des textes. Presses universitaires de Provence. Actualité de la psychologie différentielle, Presses universitaires de Provence, 2024, 979-10-32à-0511-8. hal-04716633

HAL Id: hal-04716633

<https://hal.science/hal-04716633v1>

Submitted on 1 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Actualité de la psychologie différentielle



sous la direction de
Basilie Chevrier, Bruno Dauvier, Isabelle Fort

PSY 

ACTUALITÉ DE LA PSYCHOLOGIE DIFFÉRENTIELLE

PSY

explore
les mondes
psychiques et
les comportements
à travers des
approches
diverses et
complémentaires.

Cet ouvrage présente des travaux de recherche présentés lors des XXIV^e Journées internationales de psychologie différentielle qui se sont tenues à Aix Marseille Université en juin 2022. Depuis 1975, de colloque en colloque, les travaux présentés et discutés portent sur l'étude des différences interindividuelles et de la variabilité intra-individuelle, dans une perspective délibérément large. Cette ouverture naturelle, compte tenu des liens entretenus avec de nombreuses branches de la psychologie, offre l'occasion d'interagir avec les collègues qui, traitant initialement de questions développementales, cognitives, sociales ou cliniques, rencontrent dans leurs propres recherches ou pratique des phénomènes touchant aux différences individuelles, ou sont intéressés par celles-ci. La psychologie différentielle aborde aussi, depuis ses débuts, des questions dites de terrain. Ainsi, l'émergence de nouvelles demandes sociétales dans les domaines de l'éducation, de la santé et du travail a impulsé un nouvel élan reliant étroitement les concepts, les méthodes et les pratiques. Cet ouvrage éclaire cinq thématiques : intelligence, éducation, orientation et insertion professionnelle, émotions, santé. Pour chacune d'entre elles sont présentés de nouveaux modèles touchant aux sphères cognitives, émotionnelles et motivationnelles, des méthodes de recueil et de traitement de données incluant de nouveaux outils à usage des praticiens, et des recherches trouvant des applications dans les champs de la formation, de la psychopathologie et des environnements professionnels.

Couverture

© Thomas Arciszewski

Basilie Chevrier, maître de conférences en psychologie différentielle, étudie le développement psychosocial des adolescents et des jeunes adultes.

Bruno Davier, professeur de psychologie différentielle, est spécialisé dans la modélisation des différences individuelles.

Isabelle Fort, maître de conférences en psychologie différentielle, étudie l'influence des dimensions constitutives de la métamémoire sur l'évolution des aptitudes mnésiques au cours du vieillissement ainsi que le rôle de l'auto-efficacité dans les processus d'orientation et d'accompagnement liés à l'élaboration d'un projet professionnel.

Tous trois enseignent à Aix Marseille Université.



29 €

Table des matières

Introduction	5
Pierre-Yves Gilles	

Intelligence

Analyse psychométrique en réseaux des sous-tests de l'Échelle d'Intelligence de Wechsler pour enfants et adolescents. 5 ^e édition (WISC-V)	11
Mathilde Bastien, Jean-Luc Kop, Salome Döll, Thierry Lecerf	
Étude de la stabilité à long terme des scores de la WISC-V	31
Salome Döll, Mathilde Bastien, Thierry Lecerf	
Demain tous crétins, aujourd'hui tous crédules Pourquoi il n'y a pas de déclin de l'intelligence en France	47
Corentin Gonthier, Jacques Grégoire, Maud Besançon	

Éducation

Stéréotype, conseils en orientation et aides associés aux élèves en situation de handicap et aux élèves tout-venant	63
Marjorie Chavenon, Claire Enea-Drapeau, Thomas Arciszewski, Pierre-Yves Gilles	
Un modèle systémique et dynamique de la réussite étudiante	75
Alice Constans, Raphaël Mizzi, Marie-Laure Barbier	
Étude de validation de la plateforme en ligne TACIT dédiée à l'évaluation et l'apprentissage différencié de la compréhension de l'implicite des textes	91
Olivier Le Bohec, Christophe Quaireau, Stéphane Deline, Yvonnick Noël, Jérémy Nagues, Karine Lavandier, Fanny De La Haye	

Étude de validation de la plateforme en ligne TACIT dédiée à l'évaluation et l'apprentissage différencié de la compréhension de l'implicite des textes

Olivier Le Bohec*, Christophe Quaireau*, Stéphane Deline*, Yvonnick Noël*, Karine Lavandier**, Jérémie Nogues* et Fanny De La Haye***

* Université Rennes 2, **orthophoniste, ***INSPÉ de Nouméa

1 - Etat actuel des difficultés en compréhension de texte chez les élèves français

Depuis plus de vingt-cinq ans, la France connaît un accroissement du nombre de ses élèves en difficulté face à la compréhension écrite. Ces difficultés sont objectivées par les résultats de la récente étude PISA 2018 (OECD, 2019). En 2011, les difficultés touchaient déjà environ un élève sur cinq en début de 6^{ème} (Daussin, Keskaik et Rocher, 2011). Ces deux études font état d'un écart considérable entre les élèves, en fonction de leur milieu socio-économique et d'un accroissement des inégalités entre les élèves les plus avancés et ceux qui le sont moins. Les difficultés de compréhension peuvent notamment s'expliquer par une baisse des performances dans le domaine des compétences langagières (orthographe, vocabulaire, syntaxe) et non au niveau des mécanismes de base de la lecture qui, eux, ont eu tendance à rester stables dans le temps (Daussin, Keskaik et Rocher, 2011).

La compréhension de texte est une compétence cruciale à l'école mais aussi dans la vie de tous les jours. Comprendre un texte est une habileté complexe qui nécessite d'une part d'identifier les informations explicites pertinentes, et d'autre part d'élaborer les liens implicites entre ces informations. Concernant cet implicite des textes, diverses taxonomies d'inférence ont été proposées mais aucun consensus quant à une taxonomie définitive n'est apparu (Hall, 2016). De nombreux chercheurs ont trouvé utile de distinguer a minima l'établissement de liens entre les informations explicites (par exemple, la personne derrière le pronom) et la génération de nouvelles informations (par exemple l'état interne des personnages, Kintsch et Rawson, 2005). Lefebvre et collaborateurs (2012) ont proposé d'aller au-delà de ces deux grandes catégories d'inférences assez consensuelles et de distinguer trois dimensions (non-exclusives) des inférences : i) la source d'information des inférences, ii) la contribution des inférences à la compréhension et iii) la direction des inférences. Ces trois axes permettent à ces chercheurs de positionner la plupart des types d'inférences proposées dans la littérature dans un tableau de synthèse cohérent et lisible (Lefebvre et al., 2012 ; p. 545).

De manière générale, cette compétence inférentielle, cruciale dans la compréhension de textes (Cain et Oakhill, 1999 ; Cain et al., 2001) est acquise spontanément chez les bons lecteurs. Chez les lecteurs en difficulté, cette acquisition leur fait défaut de manière récurrente et peut nécessiter un apprentissage explicite (Oakhill et al., 2019). Plusieurs auteurs ont, en effet, démontré que l'entraînement des stratégies inférentielles chez les élèves peut produire des effets positifs (Yuill et Joscelyne, 1988 ; Snow, 2002 ; Fritschmann et al., 2007). On peut noter, en France, les travaux de recherche de Potocki et al. (2015a, 2015b) qui ont permis de mettre en évidence l'efficacité de l'outil numérique LoCoTex (Ecalte et al., 2013), sur le

développement des compétences en compréhension de textes écrits, avec des élèves lecteurs en difficulté.

Jusqu'à aujourd'hui, encore peu d'outils en français permettent d'évaluer le niveau de compréhension de textes et, simultanément, de remédier aux difficultés identifiées. Seuls quelques outils papier-crayon ou en partie numérique tels que "Je lis, je comprends" (Académie d'Orléans-Tour ; cf. aussi Cèbe, Goigoux et Thomazet, 2004) et le ROLL (Réseau des Observatoires Locaux de la Lecture, développé par A. Bentolila) permettent d'évaluer le niveau de compréhension de textes et offrent aux enseignants la possibilité de proposer des activités de remédiation et/ou d'entraînement. Ces deux outils, bien qu'ayant l'avantage d'exister, ne sont pas étalonnés et n'ont pas encore fait, à notre connaissance, l'objet d'une expérimentation permettant de valider leur efficacité. D'autres outils pédagogiques permettent de proposer aux élèves un enseignement de la compréhension, ou d'apprendre à inférer, mais sans pouvoir au préalable évaluer leur niveau. "Lector-Lectrix" (Goigoux et Cèbe, 2009), "Lectorino et lectorinette" (Goigoux et Cèbe, 2013), "Stratégies pour lire au quotidien" (Gorzegno et al., 2010) font partie de ces outils qui ne permettent pas directement de construire une pédagogie différenciée. Même si des niveaux de difficulté différents sont proposés aux enseignants, ceux-ci sont obligés de se fier à leur intuition pédagogique pour attribuer des exercices adaptés au niveau de leurs élèves.

C'est précisément l'objectif de la plateforme TACIT Rennes 2 que de permettre aux enseignants de faire face à l'hétérogénéité des élèves au sein d'une même classe, en disposant d'un système d'évaluation diagnostique et d'entraînement suggérant aux enseignants des séquences pédagogiques adaptées pour les élèves du CE1 à la classe de 3ème. La plateforme TACIT, dont la conception a démarré en 2006, a été ouverte aux enseignants en 2012. Le dispositif est basé sur un modèle psychométrique explicite de la compétence en compréhension. On peut également noter l'arrivée en 2019 de la start-up Lalilo, qui propose un outil de pédagogie différenciée, prioritairement focalisé sur les élèves de CP et de CE1, dont l'algorithme n'est cependant pas explicite.

2 - Développement d'une plateforme numérique d'évaluation et d'apprentissage

2.1 Construction d'un test adaptatif

La méthodologie de développement de test utilisée dans TACIT est celle des modèles de réponse à l'item, spécifiquement le modèle de Rasch (Rasch, 1960). Ces modèles évaluent conjointement, sous la forme de paramètres estimés dans un modèle de probabilité, la compétence des élèves testés et la difficulté des épreuves, à partir d'une étude préalable sur un échantillon test de grande taille. Le module TACIT de compréhension de l'implicite des textes comporte actuellement 3738 items, dont 3658 calibrés et chaînés sur la période 2007-2021, sur un échantillon de N=15468 élèves. C'est l'un des avantages des modèles de réponse à l'item que de pouvoir rapporter les difficultés d'item à un même repère (chaînage), même lorsque les sessions d'évaluation ont été passées par des sous-groupe de sujets distincts, à des périodes différentes, à partir de sous-ensembles d'items partiellement différents. On s'appuie pour cela sur un noyau de 100 items communs à toutes les sessions.

Dans les modèles de réponse à l'item, les compétences d'élèves et les difficultés d'exercices sont projetées comme des coordonnées sur une seule dimension latente, et la probabilité de réussite à un exercice est une fonction croissante de la distance entre la position d'un élève et la position d'un exercice : deux élèves positionnés au-dessus ou en-dessous de la position de l'exercice ont respectivement une forte et une faible probabilité de le réussir. La mise en correspondance du niveau de compétence des enfants avec le niveau de difficulté des exercices permet aux enseignants de facilement sélectionner, via l'interface, le niveau de difficulté des items. Ces items seront par exemple choisis dans la zone de compétence d'un élève ou dans une zone légèrement plus élevée (cf. la notion classique de zone proximale de développement), ou encore dans une zone plus basse, pour mettre l'élève en situation motivante de réussite. Ces possibilités de choix permettent concrètement des remédiations individualisées et facilitent la gestion de l'hétérogénéité des compétences au sein d'une même classe. La définition d'un modèle unidimensionnel explicite de la réussite, et la sélection préalable d'une banque d'exercices qui s'y conforment, permettent de contrôler précisément les probabilités de réussite pour chaque couple enfant-exercice. Par défaut, TACIT règle à 0.67 cette probabilité de réussite. Mais les enseignants ont toute latitude pour choisir de mettre au défi certains élèves avancés, ou d'encourager certains élèves en difficulté, en fixant librement des niveaux plus difficiles ou plus faciles d'exercices.

L'utilisation de TACIT par les enseignants suppose donc une évaluation préalable du niveau des élèves, à partir de laquelle le choix d'une session d'exercices adaptés est possible. Les enseignants disposent de 5 formulaires de 20 exercices, pour procéder à ces évaluations potentiellement cinq fois dans l'année. La construction des échelles par le modèle de Rasch garantit le parallélisme de ces cinq formes : c'est en effet une propriété de ce modèle que de rendre possible des évaluations directement comparables, bien qu'obtenues avec des items différents. A cet égard, les modèles de réponses à l'item sont la méthodologie de choix pour évaluer l'évolution globale d'une cohorte ou d'une classe, potentiellement sur plusieurs années, avec des tests constitués de sous-ensembles d'items différents, pourvu que la calibration ait montré qu'ils appartiennent bien à la même échelle homogène, ce qui est le cas dans TACIT.

2.2 Développement d'une interface utilisateur

L'interface utilisateur de TACIT comporte une rubrique dédiée aux enseignants et un accès spécifique pour les élèves. L'enseignant va toujours débiter le travail, sur la plateforme, par une évaluation qui va permettre de positionner chacun de ses élèves sur l'échelle de compétences en compréhension implicite. L'enseignant peut ensuite construire des séances d'apprentissage ou d'entraînement, en choisissant le niveau de difficulté des exercices proposés à chaque élève ou groupe d'élèves. Des options d'aides permettent de mettre en saillance des mots indices, d'introduire une question préparatoire avant la question inférentielle ou d'accéder à la lecture audio des exercices. L'enseignant peut choisir de faire travailler ses élèves en autonomie mais aussi en petit groupe, dans le cadre de séances tutorées qui permettent à l'enseignant de les accompagner et de travailler l'explicitation des stratégies de compréhension. Ce mode tutoré permet aussi de faire travailler les élèves en mode collaboratif. L'enseignant peut également proposer des exercices à réaliser en dehors du temps de classe pour être retravaillés ensemble dans un second temps (classe inversée).

Sur les exercices, deux modalités de réponse sont possibles : QCM ou réponses ouvertes. L'interface rend possible le retour sur les exercices échoués, individuellement ou collectivement. La page de résultats permet un suivi individualisé et l'enseignant visualise aisément les réussites, les échecs et l'évolution du travail.

Côté élève, l'interface est épurée et sans éléments distrayeurs. Sur la page d'accueil, l'élève accède aux différentes séances proposées par son enseignant. Il peut visualiser la barre

d'avancement au cours d'une séance et son score final. De nombreuses options de mise en forme sont accessibles : police de caractères, taille de la police, interligne, interlettre, options « Lirecouleur ». Ces nombreuses options sont particulièrement utiles pour éviter la surcharge visuelle, problématique dans certaines formes de dyslexies (e.g. Perea et al., 2012) et, de manière plus générale, pour répondre aux difficultés des élèves à besoins éducatifs particuliers. Elles permettent aussi de contribuer au respect des principes de l'École inclusive (Bataille et Midelet, 2018).

TACIT peut être utilisé, sur ordinateur ou tablette, en présentiel mais aussi à distance et le dispositif intègre une messagerie instantanée.

3 - Présentation de l'étude

3.1 Objectifs de l'étude

Cette étude a deux objectifs principaux : i) montrer que l'entraînement en classe avec TACIT-enseignement accélère l'acquisition de la compréhension de l'implicite des textes écrits des élèves d'âge scolaire et ii) identifier les facteurs permettant d'expliquer la réussite finale. Le niveau des élèves a été évalué avant et après une phase d'entraînement.

3.2 Méthode

3.2.1 Matériel

Les épreuves de pré-test et post-test font partie des cinq formes parallèles disponibles sur la plateforme. La version du test d'évaluation en pré-test était la même pour tous les individus. Le post-test était constitué de la dernière évaluation de l'année effectuée par les élèves. Les évaluations et la phase d'entraînement ont été menées par les enseignants en situation réelle de classe.

Les items TACIT sont construits sur un format standardisé. Ils sont constitués d'un court texte, suivi d'une question inférentielle et de quatre réponses possibles, incluant une réponse « Je ne sais pas » (voir Figure 1).

3.2.2 Participants

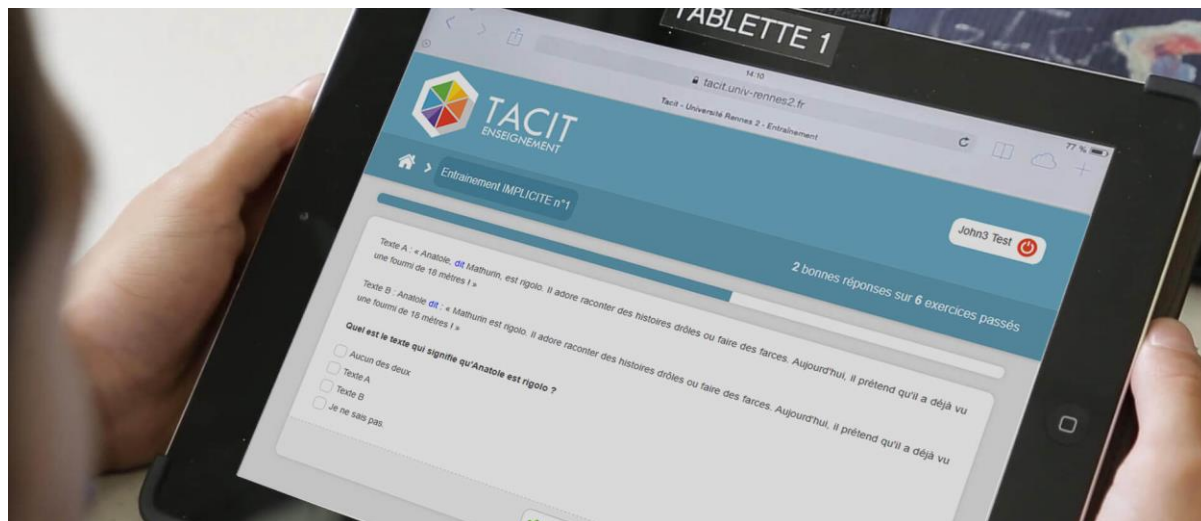
L'échantillon est constitué d'élèves d'écoles métropolitaines, du CE1 à la 6^{ème}, ayant fait leur première inscription et évaluation sur l'outil TACIT entre septembre 2017 et mars 2018. Ces élèves étaient donc des primo-utilisateurs de la plateforme.

La plateforme d'entraînement étant réglée, pour rappel, pour assurer un taux de succès de 67 % quand l'élève travaille dans son niveau (identifié à l'aide de l'évaluation), et quel que soit son niveau initial, les taux de réussite très faibles (c'est-à-dire à deux écart-types en-dessous de la moyenne) ont été considérés comme des passations exploratoires. Ce diagnostic est conforté par un temps moyen de réponse chez ces sujets étonnamment faible. Les observations correspondantes ont donc été écartées de l'échantillon. De même, les séries très longues

d'exercices d'entraînement (entre 500 et 1000), ont été considérées comme des passations exploratoires des enseignants eux-mêmes, et ont aussi été écartées des analyses.

Figure 1

L'interface TACIT



L'échantillon final est constitué de $N=3467$ élèves qui ont réalisé une moyenne de 157,5 exercices, sur une durée moyenne de 25 semaines dans l'année, avec un taux de réussite moyen observé de 69 %. Le tableau 1 présente certaines caractéristiques des individus de chaque niveau académique et quelques informations générales relatives aux résultats des évaluations réalisées (âge, score de Rasch, durée d'entraînement, nombres d'exercices réalisés, et réussis).

Tableau 1

Statistiques descriptives - moyenne (écart-type) - de l'échantillon par niveau académique.

Niveau	n	Première évaluation		Dernière évaluation		Durée de l'entraînement	Nombre d'exercices	% items réussis
		Âge	Score de Rasch	Âge	Score de Rasch			
CE1	396	7,396 (0,38)	-1,86 (0,63)	7,809 (0,378)	-1,32 (0,76)	150,8 (55,6)	171,5 (107,7)	67,98 (11,94)
CE2	689	8,376 (0,382)	-1,53 (0,69)	8,911 (0,37)	-0,81 (0,78)	195,2 (58,2)	148 (96,2)	68,08 (11,18)
CM1	749	9,372 (0,404)	-1,11 (0,74)	9,885 (0,414)	-0,49 (0,92)	187,5 (58,8)	140 (90,7)	67,86 (11,64)

CM2	885	10,384 (0,439)	-0,79 (0,76)	10,893 (0,451)	-0,13 (0,83)	186 (61,8)	155 (111,6)	70,03 (11,17)
Sixième	748	11,529 (0,475)	-0,62 (0,79)	11,914 (0,482)	-0,02 (0,89)	140,4 (46,8)	179,2 (112,5)	70,14 (10,8)
Total	3467	9,672 (1,413)	-1,09 (0,85)	10,149 (1,393)	-0,46 (0,95)	174,3 (60,7)	157,5 (105,1)	68,97 (11,33)

Note : % items réussis : pourcentage d'items réussis sur l'ensemble des items passés en entraînement (entre parenthèses : probabilité de réussite prévue par le système dans le cas d'une stricte conformité aux préconisations de la plateforme : 67 %)

3.2.3 Procédure

Les élèves ont réalisé au moins une séance d'entraînement (20 exercices), et ont été évalués à deux occasions espacées de 90 jours minimum, soit l'équivalent d'un trimestre d'entraînement, pour garantir un effet d'entraînement moyen tangible (Bianco, 2010).

4 - Principaux résultats

4.1 Mesure de la valeur ajoutée du dispositif adaptatif

Deux analyses statistiques ont été réalisées à l'aide du logiciel R (R Core Team, 2019) :

1. Une analyse intergroupe de la différence des scores d'un niveau académique au niveau académique supérieur, sans entraînement : l'objet de cette analyse est de quantifier la progression de base des élèves, hors entraînement TACIT,
2. Une comparaison des scores entre niveau scolaire n , après entraînement TACIT, et niveau scolaire $n+1$, sans entraînement (groupes contrôles) : l'objet de cette analyse est de montrer que l'usage de la plateforme amène une plus-value, par rapport aux différences de base mises en évidence au point précédent.

On notera qu'il n'est pas possible, en situation pédagogique concrète, de disposer d'un groupe contrôle naturel : les utilisateurs de la plateforme en fin d'année ont tous bénéficié d'un entraînement. Les utilisateurs en fin d'année sont donc contrastés avec des primo-utilisateurs de la rentrée suivante.

Ces analyses impliquent des comparaisons multiples. Une correction de Bonferroni a donc été appliquée systématiquement. Sur l'ensemble de l'étude, le nombre planifié de comparaisons est de 20 (seuil de décision corrigé $\alpha=0,0025$). Lorsque les conditions des tests paramétriques de comparaisons de moyenne ne tiennent pas (normalité, homoscedasticité), les statistiques non paramétriques équivalentes ont été utilisées.

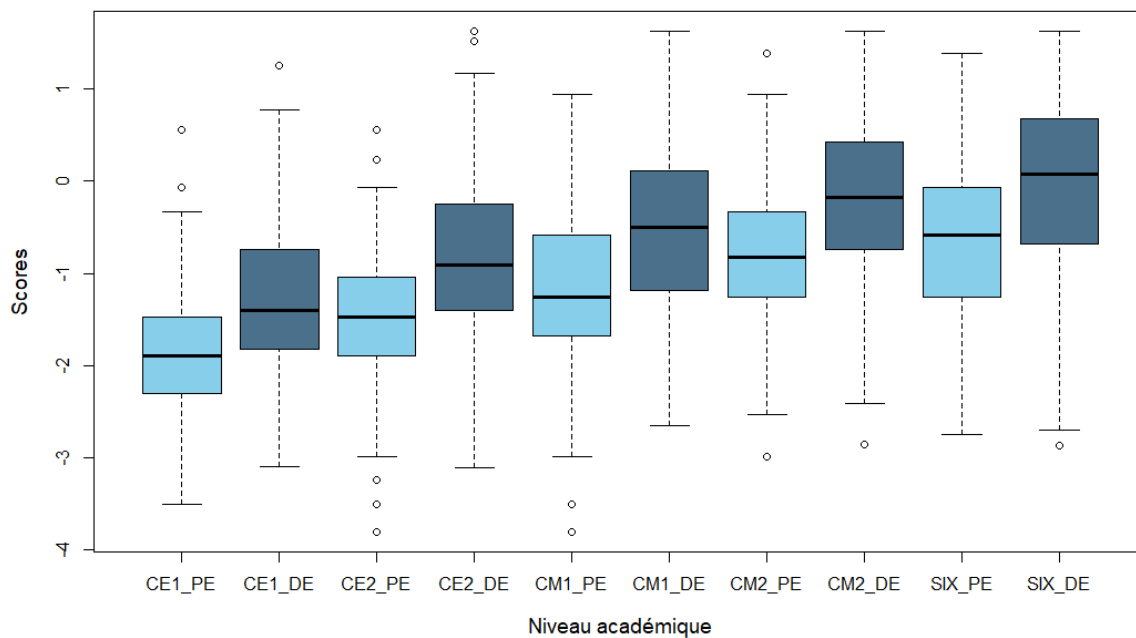
4.1.1 Analyse intergroupe entre niveaux scolaires, sans entraînement

L'objet de cette analyse est d'établir une forme de ligne de base de la progression en compréhension de l'implicite, en dehors de l'utilisation de TACIT, pour préparer l'analyse comparative qui suit.

La figure 2 représente la distribution des scores des individus selon le niveau académique. La comparaison des performances obtenues pour tous les groupes de niveau, du CE1 à la 6ème, lors de la première évaluation, avant tout entraînement, montre une progression attendue de la performance entre chaque groupe de niveau scolaire n et le groupe du niveau $n+1$ (tableau 2). Ainsi, la compréhension de l'implicite augmente avec le niveau scolaire des élèves.

« Insérer ici la figure 2 »

Figure 2. Scores de Rasch à la première (PE) et dernière (DE) évaluation, par niveau académique (SIX = Sixième).



« Insérer ici le tableau 2 »

Tableau 2

Comparaison inter-niveaux : participants non-entraînés, niveau n vs niveau $n+1$

Comparaison	<i>U de Mann-Whitney</i>	<i>Z</i>	<i>p</i>	<i>D de Cohen</i>

CE1_PE vs CE2_PE	96685,5	-8,03	$p < 0,0025$	0,49
CE2_PE vs CM1_PE	176363,0	-10,43	$p < 0,0025$	0,59
CM1_PE vs CM2_PE	256878,0	-7,88	$p < 0,0025$	0,43
CM2_PE vs SIX._PE	677046,5	-4,87	$p < 0,0025$	0,22

Note : PE= Première évaluation. SIX. = Sixième. Un seuil de décision corrigé $\alpha = 0,0025$ a été appliqué.

4.1.2 Analyse de la valeur ajoutée de l'entraînement

Pour la deuxième analyse, nous comparons les performances des élèves entraînés du niveau n , et celles des élèves non-entraînés du niveau $n+1$ (tableau 3) : le groupe des élèves non-entraînés au niveau académique $n+1$ fait office de groupe contrôle (voir tableau 1). Chronologiquement, ces nouveaux utilisateurs ont réalisé leur première évaluation en début d'année scolaire (date médiane : 4 novembre 2017). Les élèves entraînés ont, quant à eux, réalisé leur dernière évaluation en général en fin d'année scolaire (date médiane : 28 avril 2018). En conséquence, les nouveaux utilisateurs sont plus âgés que les sujets entraînés, en moyenne de 5,88 mois (de 5,52 à 7,68 mois), soit environ six mois de différence. On cherche, dans l'analyse suivante, à observer si, en dépit de cette différence d'âge, on observe un avantage chez les élèves entraînés sur TACIT, en moyenne plus jeunes.

« Insérer ici le tableau 3 »

Cette analyse est effectuée avec des tests non-paramétriques, les conditions de normalité n'étant pas respectées (test de Shapiro-Wilk, $p < 0,0001$). Les résultats indiquent que les élèves, après entraînement, au niveau n , ont des performances systématiquement supérieures aux élèves non entraînés, au niveau $n+1$. Par exemple, les élèves de CE2 entraînés ont des performances moyennes supérieures aux élèves de CM1 non entraînés (test de Mann-Whitney-Wilcoxon, $p < 0,0001$). Globalement les tailles d'effet sont de faible à moyenne (0,29-0,60), mais du même ordre de grandeur que l'évolution naturelle des enfants sur une année sans TACIT (0,22-0,59). Ces résultats confirment la plus-value pédagogique de l'entraînement par la plateforme, et de manière intéressante, pour tous les niveaux académiques testés.

Tableau 3

Comparaisons inter-niveaux, entraînés niveau n versus non-entraînés niveau $n+1$.

Comparaison	<i>U de Mann-Whitney</i>	<i>Z</i>	<i>p</i>	<i>D de Cohen</i>
CE1_DE vs CE2_PE	110844,0	-5,15	$p < 0,0025$	0,29
CE2_DE vs CM1_PE	198636,0	-7,56	$p < 0,0025$	0,39
CM1_DE vs CM2_PE	260360,5	-7,48	$p < 0,0025$	0,36
CM2_DE vs SIX._PE	216298,0	-12,09	$p < 0,0025$	0,60

Note : PE= Première évaluation ; DE= Dernière évaluation ; SIX. = Sixième. Un seuil de décision corrigé $\alpha=0,0025$ a été appliqué.

Au-delà de ce résultat général, nous cherchons dans l'analyse qui suit à identifier les facteurs de la réussite de fin d'année.

4.2 Analyse des facteurs de réussite

L'objectif de l'analyse qui suit est d'identifier les variables individuelles et pédagogiques favorisant l'apprentissage avec l'utilisation de la plateforme, dans un modèle de régression linéaire gaussien.

Les variables incluses dans la modélisation sont : le niveau classe (du CE1 à la 6ème), le retard scolaire (en nombre d'années de retard de l'élève), la durée de l'entraînement (en nombre de jours écoulés entre la première et la dernière évaluation de l'année), le genre, et le nombre d'exercices réalisés entre pré-test et post-test. Pour des raisons de colinéarité des prédicteurs dans la régression, l'âge brut n'a pas été inclus dans l'équation à cause de sa forte redondance avec le niveau classe, et celui-ci est un repère important pour les conséquences pédagogiques du modèle et son interprétation.

Outre ces variables explicatives, la variable Première Évaluation (avant tout entraînement) a été contrôlée en l'incluant comme covariable. Cela permet d'interpréter les coefficients du modèle sur les autres variables comme des mesures de leur pouvoir explicatif sur la performance à niveau initial contrôlé.

Une séquence de modèles emboîtés, incluant ces variables, tour à tour et dans cet ordre, a été testée. Une comparaison des modèles par rapports de vraisemblances a montré que toutes ces variables avaient une contribution significative à l'explication de la variance.

« Insérer ici le tableau 4 »

Tableau 4. Coefficients des prédicteurs dans le modèle final

Variables	Coefficients	Erreur-type	T	p	η^2 cumulé
Première évaluation	0,645	0,02	40,77	$p < 0,0025$	0,447
Niveau CE2	0,252	0,04	5,71	$p < 0,0025$	
Niveau CM1	0,322	0,04	7,21	$p < 0,0025$	0,466
Niveau CM2	0,477	0,05	10,60	$p < 0,0025$	
Niveau Sixième	0,551	0,05	11,71	$p < 0,0025$	
Retard scolaire	0,289	0,04	7,61	$p < 0,0025$	0,475
Durée d'entraînement	0,001	0,00	6,35	$p < 0,0025$	0,484
Genre (=>Féminin)	0,169	0,02	7,37	$p < 0,0025$	0,492
Nombre d'exercices (Log)	0,063	0,02	3,78	$p < 0,0025$	0,494

Note : test de normalité de Shapiro-Wilk : $W=0,999$ ($p=0,057$) ; test d'homogénéité des variances de Levene : $F=1,80$ ($p=0,064$). Un seuil corrigé pour le nombre de comparaison a été utilisé ($\alpha=0,0025$).

Les résultats du tableau 4 montrent plusieurs points. D'abord, les performances finales sont logiquement d'autant plus élevées que le niveau académique des élèves est élevé, soulignant l'accroissement des compétences d'un niveau au suivant. Ensuite, on note que le retard scolaire est prédictif d'une moins bonne performance finale¹, malgré le redoublement. Par ailleurs, sur le plan pédagogique, il est intéressant de noter qu'une période d'entraînement plus étalée conduit à de meilleures performances finales. Une analyse plus fine serait nécessaire pour mesurer dans cet effet d'étalement des sessions dans l'année ce qui peut relever de l'avancée en âge. En termes d'effet de genre, nous observons que les filles ont des performances supérieures à celles des garçons, comme cela a déjà été observé dans des études internationales (ex. PISA, 2012 ; PIRLS, 2016).

Enfin, en lien avec l'objectif initial de validation de l'efficacité de la plateforme, nous notons qu'un nombre d'exercices réalisés plus important contribue à de meilleures performances.

5 - Conclusion et perspectives

En résumé, les résultats de la première analyse montrent, sans surprise, une progression d'un niveau académique au suivant, sans entraînement : cette progression est une mesure de l'effet conjugué de l'environnement scolaire et familial, et du développement cognitif, hors entraînement avec TACIT. Ce qui est plus intéressant est que, par rapport à cette ligne de base, le niveau de performance chez les individus entraînés se révèle supérieur à celui des individus

¹ La variable "Retard scolaire" prend une valeur négative quand il y a un retard. Le coefficient de régression positif $\beta=0,289$ correspond donc à une pénalité sur la performance finale, une fois multiplié par la valeur de prédicteur.

non-entraînés du niveau académique supérieur, pourtant plus âgés d'une demi-année en moyenne. Enfin, les résultats de l'analyse des facteurs de réussite montrent en particulier, au-delà de l'impact des co-variables invoquées (niveau académique, genre, retard scolaire...), que les niveaux de compétence atteints sont liés au nombre d'exercices d'entraînement passés sur l'interface.

Ces résultats, montrant des effets positifs de l'utilisation du module de Compréhension implicite de textes, sont cohérents sur tous les niveaux scolaires évalués. Ils corroborent nos précédentes études de validation (ex. Villessèche et al., 2019 ; Le Bohec et al., 2019).

Au-delà de ce module Implicite, TACIT comprend également un module Vocabulaire. Si la maîtrise du vocabulaire apparaît comme un préalable fondamental au développement de la compréhension, certains auteurs (Giasson, 2004) ont souligné que le seul entraînement à l'enrichissement du vocabulaire ne suffit pas à combler l'écart entre les élèves qui ont un riche vocabulaire et les autres. L'autonomie des élèves dans leur capacité à reconstruire le sens d'un mot inconnu en contexte, apparaît au moins aussi important. A cet égard, le module Vocabulaire de TACIT comprend 5000 exercices dont un grand nombre d'items spécifiquement dédiés à l'entraînement de l'inférence du sens des mots en contexte. Sur cette base, une analyse des liens entre les performances sur les deux modules est en projet. Il s'agirait plus précisément d'examiner les liens entre l'inférence du sens des mots inconnus en contexte et la compétence inférentielle nécessaire en compréhension de textes.

Plus précisément, ce lien entre les deux compétences pourrait varier selon le niveau académique des élèves. En effet, les composantes impliquées dans la compréhension de textes (le décodage, le vocabulaire, la morphologie, l'analyse de la syntaxe, l'analyse sémantique des phrases et la production des inférences) évoluent avec le développement de l'enfant au fur et à mesure de ses apprentissages (Bianco et al., 2014). Par exemple, le décodage et l'identification des mots sont des processus coûteux cognitivement, lors des apprentissages, mais ce coût diminue à mesure qu'il devient de plus en plus automatique, libérant des ressources cognitives pour l'exécution des autres processus ou composantes (Ecalte et Magnan, 2015a).

A un niveau plus général, Daussin et al. (2011) soulignaient que, si les mécanismes de base de la lecture restaient stables sur les différentes cohortes, les compétences langagières de base (orthographe, vocabulaire, et syntaxe) sont en baisse continue. Dans le module vocabulaire, les aspects liés à la morphologie dérivationnelle sont déjà travaillés (e.g. identification des radicaux, des suffixes, ...), mais la maîtrise de la morphologie flexionnelle (accords, temps verbaux, ...), l'orthographe grammaticale et la syntaxe ne sont pas, actuellement, directement entraînés dans TACIT. Ils sont néanmoins partiellement travaillés dans les items les plus grammaticaux du module Compréhension. Des développements supplémentaires concernant l'orthographe grammaticale sont en cours : un nouveau module centré sur les homonymes (homophones hétérographes ou non et homographes homophones ou non), notamment grammaticaux, est actuellement en phase de finalisation. Il pourrait à terme constituer une des composantes d'un module plus général d'Orthographe Grammatical.

6 - Références

Bataille, P., et Midelet, J. (2018). *L'école inclusive : un défi pour l'école* (2^e éd.). ESF : Sciences Humaines.

Bentolila, A. ROLL (Réseau des Observatoires Locaux de la Lecture) (s.d.). Pour une pédagogie de la compréhension et de la différenciation. Consulté le 15/06/2023. <https://www.roll-descartes.fr/>

Bianco, M. (2010). La compréhension de textes : Peut-on l'apprendre et l'enseigner ? In *Psychologie des apprentissages scolaires* (Crahay M. et Dutrévis M), 230-256. Bruxelles : De Boeck.

Cain, K., et Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, 11(5), 489-503.

Cain, K., Oakhill, J. V., Barnes, M. A., et Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory et Cognition*, 29(6), 850-859.

Cèbe, S., Goigoux, R. et Thomazet (2004). *Enseigner la compréhension : principes didactiques, exemples de tâches et d'activités*. <https://hal.science/hal-00922482>

Colmant, M., et Le Cam, M. (2017). *PIRLS 2016 : Évaluation internationale des élèves de CM1 en compréhension de l'écrit. Évolution des performances sur quinze ans*. Note d'information de la DEPP n° 17.24. <https://www.education.gouv.fr/media/13694/download>.

Daussin, J.-M., Keskpaik, S., et Rocher, T. (2011). L'évolution du nombre d'élèves en difficulté face à l'écrit depuis une dizaine d'années. *France, portrait social*, 137—152. <https://www.insee.fr/fr/statistiques/1373895>

Ecalte, J., Potocki, A., Jabouley, D., et Magnan, A. (2013). *LoCoTex : logiciel de compréhension de textes*. Adeprio Diffusion. <https://www.orthomalin.com/actualites/vient-de-sortir/locotex>

Fritschmann, N. S., Deshler, D. D., et Schumaker, J. B. (2007). The Effects of Instruction in an Inference Strategy on the Reading Comprehension Skills of Adolescents with Disabilities. *Learning Disability Quarterly*, 30(4), 245–262.

Giasson, J. (2004). État de la recherche sur l'intervention auprès des lecteurs en difficulté. *Formation et pratiques d'enseignement en questions*. n° 1, p. 27-35.

Goigoux, R. et Cèbe, S. (2009). *Lector et lectrix. Apprendre à comprendre des textes narratifs. Cycle 3*. Paris : Retz.

Goigoux, R. et Cèbe, S. (2013). *Lectorino et lectorinette. Apprendre à comprendre des textes narratifs. CE1-CE2*. Paris : Retz.

Gorzegno, A., Legrand, C., Virely, P., et Gallet, C. (2010). *Stratégies pour lire au quotidien : Apprendre à inférer de la GS au CM2 (édition revue et augmentée)*. Dijon: Canopé-CRDP.

Hall, C. S. (2016). Inference instruction for struggling readers: A synthesis of intervention research. *Educational Psychology Review*, 28(1), 1-22.

Hembise, C., Gelin, L. et Daniel, M. (2021). Lalilo: a reading assistant for children featuring speech recognition-based reading mistake detection. Annual Conference of the International Speech Communication Association (INTERSPEECH), Show et Tell contribution, Aug 2021, Brno, Czech Republic. fahal-03410194.

<https://www.usine-digitale.fr/editorial/edtech-la-start-up-francaise-lalilo-passe-dans-le-giron-de-l-americain-rennaissance.N1073729>

Groupe département prévention de l'illettrisme – 36 (s.d.). Je lis, je comprends. Consulté le 15/06/23.

http://www.ac-orleans-tours.fr/dsden36/circ_chateauroux/evaluations/cycle_3/#c25336

Kintsch, W., et Rawson, K. A. (2005). Comprehension. In M. J. Snowling et C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209–226). London : Blackwell Publishing.

Le Bohec, O., Deline, S., De La Haye, F., Noël, Y., Nogues, J., Lavandier, K., Pálková, B. et Quaireau, C. (2019). *Apprentissage de la compréhension de l'implicite des textes à l'aide d'un dispositif numérique adaptatif : une étude de validation*. 60ème congrès de la Société Française de Psychologie : Apprentissages, Vulnérabilités, Préventions, Poitiers, 4, 5 et 6 septembre.

Lefebvre, P., Bruneau, J., et Desmarais, C. (2012). Analyse conceptuelle de la compréhension inférentielle en petite enfance à partir d'une recension des modèles théoriques. *Revue des sciences de l'éducation*, 38(3), 533-553.

Oakhill, J., Cain, K., et Elbro, C. (2019). Reading Comprehension and Reading Comprehension Difficulties. In D. A. Kilpatrick, R. M. Joshi, et R. K. Wagner (Éds.), *Reading Development and Difficulties* (p. 83-115). New York : Springer International Publishing.

OECD (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing.

<https://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview-FR.pdf>

OECD (2019). *PISA 2018 Results (Volume I) : Where All Students Can Succeed*. OECD Publishing. <https://www.oecd.org/publications/pisa-2018-results-volume-i-5f07c754-en.htm>

Perea, M., Panadero, V., Moret-Tatay, C., et Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with young normal readers and developmental dyslexics. *Learning and Instruction*, 22(6), 420–430.

Perfetti, C. A. (2000). Comprehending written language: A blueprint of the reader. In *The neurocognition of language*. Vol. 167, p. 208. Oxford: Oxford University Press..

Potocki, A., Ecalle, J., et Magnan, A. (2015a). Computerized comprehension training in young readers: For whom and under which conditions is it efficient? *Journal of Computer Assisted Learning*, 31(2), 162-175.

Potocki, A., Magnan, A., et Ecalle, J. (2015b). Computer based training in four groups of struggling readers: Specific effects on word reading and comprehension. *Research in Developmental Disabilities*, 45-46, 83-92.

R Core Team. (2019). *R: a language and environment for statistical computing (version 3.5.3)*.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests. Studies in mathematical psychology. Danmarks Paedagogiske Institut*.

Snow C. E. (2002). *Reading for understanding: Toward a research and development program in reading comprehension*. Santa Monica, CA: Rand.

Villessèche, J., Le Bohec, O., Quaireau, C., Nogues, J., Besnard, A.-L., Oriez, S., De La Haye, F., Noel, Y., et Lavandier, K. (2019). Enhancing reading skills through adaptive e-learning. *Interactive Technology and Smart Education*, 16(1), 2-17.

Yuill N. et Joscelyne T. (1988). Effect of organizational cues and strategies on good and poor comprehenders' story understanding. *Journal of Educational Psychology*, 80 (2), 152-158.