

Introduction

La méta-analyse de King-Sears et ses collègues (2023) est l'une des plus récente ayant évalué l'impact de l'enseignement basé sur la *Conception Universelle de l'Apprentissage (UDL)* sur la réussite académique des apprenants, en la comparant à un enseignement traditionnel. Cette méta-analyse a sélectionné 20 études expérimentales couvrant des apprenants de la maternelle à l'âge adulte.

Sur les 20 études sélectionnées, j'ai réalisé une seconde analyse en conservant uniquement celles effectuées au primaire et au secondaire :

- Dont les interventions ont été réalisées sur un minimum de 12 semaines, tel que recommandé par Robert Slavin dans ses [*Best Evidence*](#) ;
- Impliquant une population régulière d'élèves ;
- Ayant un enseignement prodigué en présentiel.

Des 20 études retenues dans cette méta-analyse, seulement deux recherches respectent les critères mentionnés précédemment. Il s'agit de deux études réalisées au primaire. Voici une analyse critique de ces deux recherches.

1. Étude de Proctor et al. (2011)

Cette recherche a évalué sur 16 semaines l'impact du programme numérique ICON, basé sur la Conception Universelle de l'Apprentissage (UDL), auprès d'élèves de 5e année du primaire, incluant des apprenants bilingues. L'intervention, menée en grand groupe, ciblait la lecture, le vocabulaire et la compréhension.

Analyse critique

- Faiblesses méthodologiques et limites d'échantillonnage : Cette recherche repose sur un plan quasi-expérimental où l'assignation au groupe expérimental ou de contrôle s'est faite au niveau de la classe plutôt qu'à celui de l'élève. L'échantillon ne comptant que 12 classes, cela a limité la puissance statistique de l'étude et empêche de tirer des conclusions causales, gonflant potentiellement artificiellement l'effet mesuré de l'intervention. De plus, le groupe de contrôle a suivi un enseignement « habituel » très hétérogène (le temps consacré à la littérature variait de 55 à 135 minutes par jour selon les classes) ; les chercheurs admettent qu'un groupe de contrôle bénéficiant d'un environnement alternatif rigoureusement surveillé aurait été plus pertinent.
- Défauts de mise en œuvre (fidélité) : L'intervention prévoyait la lecture de 8 textes numériques, mais seulement 54 % des élèves du groupe expérimental ont achevé le programme complet (avec une moyenne de 6,96 textes lus), obligeant les auteurs à analyser

les résultats en fonction du « temps passé sur la tâche » plutôt que par une simple comparaison de groupes.

- **Instabilité des mesures et absence de transfert** : Bien que l'outil ait amélioré le vocabulaire, il n'a eu aucun effet significatif sur la compréhension écrite globale (évaluée par les tests *Gates-MacGinitie*), ce qui est pourtant l'objectif fondamental de toute intervention sur le vocabulaire. Pour justifier cette absence de transfert, les auteurs suggèrent que les tests standardisés requièrent des compétences secondaires (comme la gestion du temps ou l'élimination de mauvaises réponses) que leur programme n'enseignait pas. Enfin, les tests créés par les chercheurs eux-mêmes présentaient de sérieuses failles : le test d'étendue du vocabulaire (VBT) s'est révélé psychométriquement instable (manquant de variance pour détecter des effets), tandis que le test de profondeur (VDT) reposait sur une échelle très restreinte de 4 points, ce qui a pu accentuer de manière artificielle les sensibilités statistiques entre les différentes classes.

2. Étude de Yu et al. (2021)

Cette recherche s'est déroulée sur un an auprès d'élèves de 4^e année du primaire et visait à structurer les processus de résolution de problèmes en sciences. Elle a évalué un carnet scientifique numérique inclusif (SNUDLE) lors de sessions combinant apprentissage en grand groupe et travail autonome avec la technologie.

Analyse critique

- **Problèmes majeurs de mise en œuvre (fidélité et dosage)** : L'étude a gravement souffert d'un niveau d'implémentation insuffisant dans les classes. Sur l'ensemble du semestre, les enseignants ne se sont connectés en moyenne que 9,63 jours, et les élèves n'ont réalisé en moyenne que 6,53 investigations. Cette faible assiduité a été exacerbée par une catastrophe naturelle survenue lors de la première année (provoquant des fermetures d'écoles) et par un accès parfois inconstant aux tablettes tactiles.

- **Limites de la collecte des données individuelles** : L'évaluation de l'utilisation spécifique du logiciel par chaque élève a été biaisée par le fait que les élèves travaillaient souvent en binôme ou en groupe en se connectant sur le compte d'un seul élève. Il est donc impossible pour les chercheurs, dans les sources, de lier avec une certitude absolue l'utilisation précise des fonctionnalités individuelles aux résultats académiques individuels.

- **Absence d'effet global et manque de transfert sur les tests nationaux** : L'intervention n'a démontré aucun impact statistiquement significatif sur les performances académiques ou la motivation de la population globale des élèves. Bien que l'étude révèle des tailles d'effet très importantes et positives (de 0,82 à 1,01) ciblant spécifiquement les élèves en situation de handicap et des effets positifs chez les élèves allophones, un problème de transfert demeure. En effet, ces améliorations académiques ont été mesurées uniquement via les quiz liés au curriculum immédiat et les évaluations du district ; ces gains ne se sont pas reflétés sur le test scientifique national standardisé (MAP science). Cela suggère que l'outil aide à

accomplir les tâches immédiates, mais ne transforme pas nécessairement la maîtrise globale et standardisée de la discipline.

Conclusion

L'analyse de ces deux études met en lumière des résultats intéressants, mais trop limités pour soutenir de manière convaincante l'efficacité des interventions fondées sur l'UDL en milieu scolaire ordinaire. Dans l'étude de Proctor et al. (2011), les faiblesses méthodologiques — notamment le plan quasi-expérimental, la petite taille de l'échantillon et l'hétérogénéité du groupe contrôle — réduisent fortement la capacité de l'étude à établir un lien causal clair entre l'intervention numérique et les apprentissages. Les problèmes de fidélité d'implantation et l'instabilité des outils de mesure nuisent également à l'interprétation des résultats, notamment l'absence de transfert vers des tests standardisés de compréhension.

L'étude de Yu et al. (2021) présente un autre type de contraintes, principalement liées à la mise en œuvre. La faible utilisation du carnet numérique par les enseignants et les élèves, combinée à des perturbations externes importantes et à des limites dans la collecte de données individuelles, affaiblit la robustesse des conclusions. Les effets positifs observés pour certains sous-groupes, bien qu'encourageants, demeurent étroitement liés aux évaluations de proximité et ne se transfèrent pas vers des mesures plus globales comme les tests nationaux standardisés.

Ces deux recherches montrent que les interventions basées sur l'UDL n'ont pas encore démontré une efficacité solide, généralisable et durable dans les conditions réelles de classes ordinaires. Les limites méthodologiques, les problèmes d'implantation et l'absence de transfert vers des mesures standardisées convergent vers un même constat : **des études plus rigoureuses, mieux contrôlées et implantées avec constance sont nécessaires pour évaluer de manière fiable l'impact réel de l'UDL sur les apprentissages scolaires.**

Références

- King-Sears, M. E., Stefanidis, A., Evmenova, A. S., Rao, K., Mergen, R. L., Owen, L. S., & Strimel, M. M. (2023). Achievement of learners receiving UDL instruction: A meta-analysis. *Teaching and Teacher Education*, 122, Article 103956. <https://doi.org/10.1016/j.tate.2022.103956>.
- Proctor, C., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing*, 24(5), 517–544. <http://mutex.gmu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eih&AN=59928743&site=ehost-live>.
- Yu, J. W., Fikes, A. E., Ferguson, K., Wei, X., Tiruke, T., Hall, T. E., & Blackorby, J. (2019). Efficacy study of the science notebook in a Universal Design for Learning environment:

Preliminary findings. Grantee Submission. ERIC.
<http://mutex.gmu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED605911&site=ehost-live>.

