Groupe de travail de l'ANRT
« Pédagogie par le jeu »

L'ÉVALUATION DES DISPOSITIFS PÉDAGOGIQUES INNOVANTS

Coordonné par Violette Nemessany, ANRT
en collaboration avec Ampiric,
pôle pilote de formation des enseignants et de recherche pour l'éducation

ANRT
ASSOCIATION NATIONALE
RECHERCHE TECHNOLOGIE

L'ÉVALUATION DES DISPOSITIFS PÉDAGOGIQUES INNOVANTS

Coordination :
Violette Nemessany, ANRT

Président du groupe de travail :
Franck Tarpin-Bernard, Humans Matter

Directrice de publication :
Clarisse Angelier, ANRT

En collaboration avec Ampriric, pôle pilote de formation des enseignants et de recherche pour l'éducation.

Soutenu par l'État dans le cadre de l'action « Territoires d'innovation pédagogique » du Programme d'investissements d'avenir, opéré par la Caisse des Dépôts.

Ces travaux sont soutenus financièrement par les souscripteurs FutuRIS

Cet ouvrage collectif rend compte des travaux menés, en 2022 et 2023, dans le cadre du groupe de travail « Pédagogie par le jeu » de l'ANRT, présidé par Franck Tarpin-Bernard, directeur général délégué de Humans Matter.

Son contenu n'engage que la responsabilité des auteurs, et non celle des institutions qui apportent leur soutien à l'ANRT.

L'ANRT et le président du groupe de travail remercient chaleureusement toutes les personnes qui, de près ou de loin, ont apporté leurs compétences, leur temps et leur soutien à l'élaboration de cet ouvrage.

SOMMAIRE

PRÉFACES 3
Elie Allouche 3
Jean-Philippe Lachaux 6

LISTE DES AUTEURS 9

INTRODUCTION 15
Franck Tarpin-Bernard

PARTIE 1 : Cadrage 18

Évaluer les dispositifs pédagogiques ludiques 19
Violette Nemessany

Développer, expérimenter et évaluer des dispositifs pédagogiques innovants : l'exemple du CréativLab d'Ampiric et la question du passage à l'échelle 37
Nicolas Mascret, Thierry Berthet, Pascale Brandt-Pomares, Magali Coupaud, Mathilde Favier, Jonathan Mirault, Johannes C. Ziegler

PARTIE 2 : Les outils 55

Les neurosciences cognitives entrent à l'école : évaluer la collaboration de recherche ? 56
Pascale Toscani

Sentiment d'efficacité personnelle : déterminant ou médiateur du lien entre stress perçu et agentivité 68
Stéphanie de Chalvron, Hanna Chainay, Franck Tarpin-Bernard, Alexandre Beaussier, Riadh Lebib, Valentine Facque, Florence Cousson-Gélie

CEPAJe, un outil pour évaluer une séance ludopédagogique 80
Julian Alvarez

Les biomarqueurs numériques sont-ils des outils puissants et agnostiques d'évaluation des dispositifs pédagogiques numériques ? 98
Théo Marchand, Laurence Vanin, Vanessa Douet-Vannucci

PARTIE 3 : Les dispositifs 112

Création et évaluation d'un programme pédagogique inclusif basé sur l'activite physique et le jeu sportif à destination d'enfants ayant un trouble du spectre de l'autisme 113
Olivia Collet, Geneviève Cabagno, Sandrine Le Sourn-Bissaoui

Utilisation d'un jeu en pédagogie universitaire ; évaluation de l'attention des étudiants 128
François Lecellier

La place des jeux au sein de la formation 140
Élodie Delahaye, Jean Marc Ferrandi, Lionel Muniglia

Sciado Tracker : un module pour le suivi des apprenants dans les simulateurs 155
Nesrine Rahmouni, Laurent Lavogiez

PRÉFACES

Elie Allouche

Chef de projet recherche appliquée et incubateur de projets numériques,
bureau du soutien à l'innovation numérique et à la recherche appliquée (DNE-TN2),
ministère de l'Éducation nationale et de la Jeunesse

En visant à fournir aux enseignants un ensemble d'outils, de ressources et de formations en soutien à leur liberté pédagogique et à leurs innovations, la stratégie du numérique pour l'éducation, publiée en janvier 2023 (MENJ, 2023a) a pour ambition de développer les compétences numériques des professeurs et de leurs élèves. Ces compétences, évaluées par la certification Pix, ont désormais un rôle déterminant dans la capacité d'innover et de faire face aux défis du monde contemporain. Pour les élèves, développées tout au long de leur parcours scolaire, ces compétences sont aussi centrales pour l'exercice d'une citoyenneté éclairée comme pour l'insertion dans la vie professionnelle.

Mais il faut rappeler ici que la mise en œuvre de dispositifs innovants et leur évaluation, dans le cadre de pratiques numériques disciplinaires et interdisciplinaires en constante évolution – voir par exemple la richesse des productions des travaux académiques mutualisés (MENJ, 2023b) – constituent d'abord une part essentielle des pratiques pédagogiques quotidiennes des enseignants. Cela leur permet d'être à l'écoute de leurs élèves, de les accompagner dans leurs apprentissages, et de contribuer à leur réussite personnelle et professionnelle.

Pour accompagner l'innovation des acteurs de terrain, la Direction du Numérique pour l'éducation a mis en place notamment les dispositifs des territoires numériques éducatifs et des challenges innovation, grâce au financement France 2030. Par ces actions, la DNE accompagne et soutient la dynamique d'innovation de l'ensemble des acteurs de l'écosystème, dont les collectivités territoriales et les acteurs de la filière industrielle. L'un des enjeux est l'orientation des élèves, en particulier des filles, vers les formations et les filières du numérique. Cette dynamique d'innovation inclut l'enseignement professionnel ou technologique, « qui nécessite des outils techniques avancés, et qui doit pouvoir tirer profit d'innovations technologiques, comme l'immersion virtuelle ou les jumeaux numériques ».

Les actions et projets de la DNE en matière d'innovation, entendue ici comme un soutien aux processus de socialisation d'inventions (Cerisier, 2014), vont de pair avec un soutien à la recherche appliquée via la coordination des groupes thématiques numériques – les #GTnum. À l'issue d'un

processus de sélection conduit avec la DGESCO, le Conseil scientifique de l'éducation nationale et l'inspection générale de l'Éducation, du Sport et de la Recherche, les #GTnum associent laboratoires et territoires académiques pour produire des connaissances et participer à leur diffusion, tout particulièrement sur éduscol (MENJ, 2023c) et le carnet Hypothèses « Éducation, numérique et recherche » (DNE-TN2, 2019).

L'innovation et la recherche appliquée à la DNE ont ainsi pour but d'améliorer la diffusion, la prise en compte et l'appropriation professionnelle des résultats de la recherche scientifique, pour mieux comprendre et accompagner la transformation numérique de l'éducation, enseigner, se former et apprendre. Accompagner la transformation numérique de l'éducation intègre essentiellement trois dimensions :

- l'impact des mutations socioculturelles sur le système éducatif et les acteurs de la communauté éducative ;

- la façon dont l'institution et les acteurs prennent part à ces mutations ;

- le déploiement d'infrastructures, d'équipements, de services, de ressources et de formations destinés à accompagner les équipes éducatives dans leur parcours d'appropriation, notamment par la transformation des outils en instruments pédagogiques (Allouche, 2022).

Cette transformation s'effectue avec l'impulsion et l'accompagnement de la DNE, qui définit et met en œuvre le service public du numérique éducatif dans le cadre d'une gouvernance qui associe les collectivités territoriales.

À ce titre, la recherche participative, « du laboratoire à la classe et de la classe au laboratoire », conduite dans le cadre des #GTnum, permet de construire une démarche ouverte et pluraliste, valorisant la diversité des perspectives, des données et des méthodes, associant les acteurs de terrain à divers degrés avec l'idée de faire de la recherche « avec » plutôt que « sur » (DNE-TN2, 2021) pour favoriser et réussir les passages à l'échelle.

Parmi les thématiques traitées par les #GTnum figurent, entre autres, celles qui sont traitées dans le présent ouvrage, à savoir l'apprentissage par le jeu, l'inclusion, l'hybridation (présentiel/distanciel) des enseignements et apprentissages, et l'intelligence artificielle (IA). Cette thématique de l'IA, déjà traitée en 2020-2022 par deux #GTnum, est certainement l'un des faits marquants de l'année 2023, avec la montée en puissance des IA génératives et des grands modèles de langage, qui constituent à la fois un défi de taille et un ensemble d'opportunités pour tous les acteurs de l'éducation et de la formation (DNE-TN2, 2023a).

Pour conclure cette préface, nous souhaitons saluer, d'une part, l'ANRT, à l'initiative de ce recueil et partenaire précieux de l'Éducation nationale en matière de recherche et d'innovation, en particulier par le soutien et la

valorisation des doctorants Cifre, ainsi que l'organisation du colloque du 9 mars 2023 sur la recherche partenariale en éducation (DNE-TN2, 2023b), et d'autre part, l'ensemble des collègues rédactrices et rédacteurs du présent ouvrage pour leurs apports et leurs réflexions sur l'évaluation des dispositifs pédagogiques innovants.

Bibliographie

Elie Allouche. (2022, décembre 5). Numérique éducatif et recherche appliquée : les actions de la DNE. Innovation pédagogique. [En ligne].

Jean-François Cerisier. (2014). « On demande toujours des inventeurs et l'on cherche encore les innovateurs », *Distances et médiations des savoirs.* Article 8. [En ligne].

DNE-TN2. (2019). Éducation, numérique et recherche – Veille et diffusion des travaux de recherche sur le numérique dans l'éducation. *Éducation, numérique et recherche.*

DNE-TN2. (2021, septembre 9). Les formes de partenariat recherche/éducation : références, repères et pistes de travail [Billet]. *Éducation, numérique et recherche.*

DNE-TN2. (2023a). Intelligence artificielle et éducation : apports de la recherche et enjeux pour les politiques publiques [Billet]. *Éducation, numérique et recherche.*

DNE-TN2. (2023b, janvier 11). Colloque « Recherche partenariale : éducation, enseignement et formation » (09/03/23) [Billet]. *Éducation, numérique et recherche.*

MENJ. (2023a, janvier). Stratégie du numérique pour l'éducation 2023-2027. *Ministère de l'Éducation nationale et de la Jeunesse.*

MENJ. (2023b, mai). Les travaux académiques mutualisés (TraAM) : des laboratoires de pratiques numériques. Eduscol. *Ministère de l'Éducation nationale et de la Jeunesse - Direction générale de l'Enseignement scolaire.*

MENJ. (2023c, mai). Enseigner et apprendre avec la recherche : les groupes thématiques numériques (GTnum). Eduscol. *Ministère de l'Éducation nationale et de la Jeunesse - Direction générale de l'Enseignement scolaire.*

Jean-Philippe Lachaux

Directeur de recherche à l'INSERM
Centre de recherche en neurosciences de Lyon

N'attendez pas de cette préface qu'elle vous éclaire davantage sur le sujet traité par cet ouvrage : ce serait bien difficile, tant ce livre est complet. Je peux juste m'honorer d'avoir été l'un de ses premiers lecteurs, avec, au fil de ma lecture, le sentiment récurrent que toutes les interrogations qui me venaient à l'esprit trouvaient très rapidement leur réponse. Je ne saurais trop le recommander à toute personne désireuse de développer ou de tester un dispositif pédagogique innovant, car, immanquablement, se posera à elle la question de son évaluation : une tâche ardue, qui demande à être abordée, et définie, avec soin.

Sous ce terme générique d'évaluation se cache une multitude d'objectifs différents, qui vous seront dévoilés dans toute leur diversité à travers ces pages. On peut évaluer dans l'idée d'aboutir à un jugement couperet et définitif : « ça marche », ou « ça ne marche pas ». Mais il est bien sûr bien plus utile d'envisager l'évaluation comme un diagnostic intermédiaire dans un processus d'amélioration d'un dispositif, à l'affût de tout ce qui ne fonctionne pas, ou pas assez bien, et dans un but d'amélioration. On retrouve alors le rapport à l'erreur que l'on cherche ardemment à développer chez l'élève : l'erreur comme information et comme guide pour progresser.

J'ai volontairement utilisé une formulation caricaturale – « ça marche » – pour susciter une réaction de la part du lecteur : qu'entend-on par ce terme ? Et qui décide que « ça marche », selon quels critères ? À nouveau, vous trouverez toutes les réponses dans ce livre. Il sera beaucoup question de la « pyramide des preuves », un concept emprunté à la recherche clinique et sur lequel j'aimerais revenir un instant. La pyramide des preuves établit une hiérarchie entre différents niveaux de preuve – concernant l'effet d'un dispositif ou d'un traitement –, depuis l'observation anecdotique isolée, jusqu'à la synthèse de plusieurs études conçues selon les standards de rigueur acceptés par la recherche. Toutefois, je voudrais ici suggérer qu'il est possible de renverser cette pyramide, et que c'est même utile, non pas dans un simple but de provocation et de contestation, mais dans l'idée d'une évaluation plus complète.

Je m'explique : quand les grandes compagnies pharmaceutiques travaillaient d'arrache-pied à développer un vaccin contre le COVID-19, il était évidemment absolument nécessaire d'employer un dispositif d'évaluation le plus rigoureux possible, pour convaincre, sans l'ombre d'un doute, que le vaccin avait un effet sur une population large, qui ne pouvait être expliqué par aucun autre facteur. Les chercheurs impliqués devaient adopter la posture extrême que l'on attend d'eux dans ce cas-là, la seule garante d'une évaluation impartiale : chercher par tous les moyens à montrer que leur vac-

cin ne marchait pas, à l'affût de la moindre faille dans la procédure de test, jusqu'à devoir admettre, de guerre lasse, que l'effet est réel. Un évaluateur rigoureux doit donc se comporter comme un ennemi de sa création, ce qui exige une bonne flexibilité psychologique, vous en conviendrez. Dans ce processus d'évaluation du vaccin, l'anecdote isolée n'a bien sûr pas sa place : que le docteur Lebon ait observé un effet intéressant chez madame Triffeuil à Amiens n'a évidemment aucune importance. Et pourtant… Ce qui intéresse ce monsieur Lebon, c'est d'immuniser sa patiente ! Que le vaccin ait un effet de protection de 80 % chez 90 % des patients lui fournit une indication intéressante, qui doit l'encourager à vacciner cette bonne dame, mais qui ne garantit pas à 100 % un effet (notez bien que j'ai pris cet exemple au fil de ma plume, en inventant des chiffres farfelus, et sans aucune volonté de réouvrir le débat sur le vaccin !).

Revenons à l'apprentissage, avec un parallèle évident : l'évaluation rigoureuse d'un dispositif innovant, avec le plus haut niveau de preuve, est évidemment une étape importante pour sa diffusion et son financement. On ne peut rien conclure de quelques commentaires enthousiastes d'enseignants épars, c'est entendu. Il n'en demeure pas moins que pour madame Tilleul, professeure d'histoire-géographie à Nantes, la démonstration rigoureuse que tel dispositif a augmenté le taux d'apprentissage de telle notion du programme de 70 %, chez 60 % des élèves, est une information intéressante, qui doit l'encourager à s'y intéresser ; mais qui ne garantit aucunement que l'application de ce dispositif par elle-même, avec ses élèves, va produire l'effet qu'elle attend, pour subvenir à ses « besoins ». La seule évaluation vraiment utile pour cette enseignante, c'est celle qu'elle peut faire elle-même en l'utilisant avec ses élèves. C'est en fonction de son ressenti – le niveau de preuve le plus « faible » de la pyramide – qu'elle décidera ou non d'investir le temps nécessaire pour l'utiliser une année de plus. Pour cette dame, les résultats des méta-analyses constituent finalement une preuve moins forte. C'est une donnée théorique qui doit être confrontée à sa propre pratique. Et c'est encore plus vrai dans le cadre de l'enseignement que dans celui de la clinique, car, si on peut concevoir qu'un traitement mette en jeu des mécanismes chimiques universels, ce n'est pas le cas d'un dispositif pédagogique, qui implique des processus psychologiques et sociaux beaucoup plus complexes et variables : l'influence des facteurs particuliers à la relation classe-enseignant est majeure. On pourra rétorquer que cette influence peut être réduite considérablement par des dispositifs numériques confrontant directement l'élève à une machine, sans l'influence « parasite » d'un médiateur humain, mais cela nous renvoie à d'autres débats.

En conclusion de cette courte préface, je souhaiterais vous relater une anecdote dont j'ai fait personnellement l'expérience, et que je trouve particulièrement éclairante sur les limites d'une certaine rigueur scientifique. Mon équipe et moi-même souhaitions évaluer l'impact d'un programme sur les capacités attentionnelles des élèves. Inspirés par nos pratiques de laboratoire, nous avions assez logiquement mis en place un protocole soigné, mêlant des tests neuropsychologiques classiques et une mesure informa-

tisée de l'attention, sur tablette. Mes collègues et moi assurions la passation des tests, après un accueil bienveillant des enseignants, cantonnés au rôle « d'ouvreurs de portes », tels Passe-Muraille dans le jeu Fort Boyard. Les résultats aux tests furent bluffants, tant dans les classes ayant suivi le programme que dans les classes contrôle, comme si aucun élève n'avait le moindre problème d'attention. Si tous les élèves sont parfaitement attentifs, pourquoi envisager une éducation de l'attention ? Bien sûr, nous avions fait deux erreurs majeures, que ne manquèrent pas de souligner les professeurs de ces classes : la première, c'était d'être venus nous-mêmes réaliser les tests, dans le but de contrôler aussi précisément que possible tous les paramètres de la passation. Certes, la précaution était louable, mais notre présence avait complètement modifié l'état attentionnel des élèves. Autant aller photographier les animaux de la forêt la nuit, en voiture, tous phares allumés ! La deuxième erreur, soulignée par les enseignants, était d'avoir utilisé des tests de laboratoire totalement déconnectés du travail quotidien des élèves. Nous avions voulu transformer la classe en un laboratoire d'un jour, dans un but de rigueur. Mais dans cette tentative d'éliminer tout facteur confondant, nous avions jeté le bébé avec l'eau du bain.

La leçon de cette mésaventure, c'est que l'évaluation d'un dispositif innovant – comme sa conception – doit être envisagée main dans la main avec les acteurs de terrain, dans le but d'atteindre des conclusions aussi générales que possible, mais sans oublier le cas particulier de madame Tilleul à Nantes.

LISTE DES AUTEURS

Elie Allouche est chef de projet recherche appliquée et incubateur de projets numériques au sein de la direction du Numérique pour l'éducation (bureau TN2 chargé du soutien à l'innovation numérique et à la recherche appliquée) au ministère de l'Éducation nationale et de la Jeunesse. Engagé dans le champ des humanités numériques, pour la production et la diffusion des savoirs en accès ouvert, Elie Allouche est notamment chargé du carnet Hypothèses « Éducation, numérique et recherche » https://edunumrech.hypotheses.org/ et des groupes thématiques numériques #GTnum. Auparavant, il a occupé plusieurs postes dans l'enseignement, la recherche et l'accompagnement de la transformation numérique de l'éducation, en tant que directeur de l'Atelier Canopé du Val-de-Marne, chargé de cours en technologies éducatives à l'université Paris Descartes, directeur du centre départemental de documentation pédagogique des Hauts-de-Seine, et professeur agrégé d'histoire-géographie dans l'enseignement secondaire.

Julian Alvarez est responsable de la R&D en ludopédagogie chez Immersive Factory. En parallèle, il est professeur associé aux Laboratoires LARSH (Université Polytechnique Hauts-de-France) et CRIStAL (université de Lille). Julian Alvarez est responsable pédagogique du diplôme d'université « Apprendre par le jeu » à l'Inspé Lille Nord-de-France. Depuis 1991, il a été impliqué dans plus de 150 réalisations de jeux et de serious games pour le compte d'acteurs privés et publics.

Alexandre Beaussier est ingénieur de formation à ENSTA, titulaire d'un mastère en sciences, « Space Science », de l'University College London, ainsi que d'un mastère spécialisé en management de projets technologiques de l'ESSEC. Il dirige le pôle « innovation » de Humans Matter, en explorant les nouveaux modes d'apprentissage, de décision et d'engagement à la lumière des sciences cognitives.

Thierry Berthet est politiste, directeur de recherche CNRS et directeur du laboratoire d'économie et de sociologie du travail (LEST). Ses travaux portent sur les politiques comparées de lutte contre le décrochage scolaire, l'action publique en matière d'orientation, la territorialisation des politiques publiques et les politiques de l'emploi. Ses travaux sur le décrochage scolaire portent notamment sur l'expérience des dispositifs de remédiation par les jeunes décrocheurs et les opérateurs de terrain. Thierry Berthet pilote le dispositif de suivi et d'évaluation (DSE) du pôle pilote Ampiric.

Pascale Brandt-Pomares est professeure en sciences de l'éducation et de la formation ; vice-présidente déléguée au centre d'innovation pédagogique et d'évaluation d'Aix-Marseille Université et directrice de l'Inspé d'Aix-Marseille. Didacticienne des enseignements technologiques, elle s'intéresse aux

processus d'enseignement-apprentissage. Ses travaux portent en particulier sur les rôles et fonctions du numérique dans l'activité enseignante. Elle est coresponsable du programme de recherche Artefacts de formation de l'UR 4671 ADEF, et pilote l'Action 1 du pôle pilote Ampiric sur la formation initiale et continue des enseignants.

Geneviève Cabagno est professeure en STAPS à l'université Rennes 2. Elle est affiliée au laboratoire Valeurs, innovations, politiques, socialisations et sports (VIPS^2, UR 4636). Ses recherches investiguent plus particulièrement les processus psychosociaux favorisant l'engagement dans la pratique physique ou résultant de celle-ci, notamment dans le cas de publics vulnérables.

Hanna Chainay est professeure en psychologie cognitive à l'université de Lyon 2. Elle codirige l'école doctorale de neurosciences et cognition (NSCo, ED 476) et est membre du comité d'éthique de l'université de Lyon. Ses recherches portent sur les processus normaux et pathologiques de la mémoire, et plus particulièrement, les effets des émotions sur la mémoire.

Stéphanie de Chalvron est docteure en psychologie, chercheuse associée au Laboratoire Paragraphe (EA349) de l'université Paris VIII, chargée de cours, et directrice scientifique de Humans Matter, centre international de design cognitif. Ses recherches explorent l'impact des processus psychosociaux dans le développement de la capacité d'agir dans les domaines de l'éducation et de la santé.

Olivia Collet est doctorante en STAPS, affiliée au laboratoire VIPS^2 et LP3C de l'université Rennes 2. Doctorante CIFRE au sein de l'ONG PLAY International, sa thèse porte sur « la création et l'évaluation d'impact d'un programme pédagogique inclusif basé sur l'activité physique et le jeu sportif pour des enfants ayant un trouble du spectre de l'autisme (TSA) », sous la direction de Geneviève Cabagno et de Sandrine Le Sourn-Bissaoui.

Magali Coupaud est ingénieure de recherche au sein du pôle pilote Ampiric en charge de l'ingénierie de formation, et membre du Laboratoire Apprentissage, Didactique, Évaluation, Formation (ADEF) d'Aix-Marseille Université. Elle s'intéresse à la didactique de la biologie, et en particulier aux enjeux didactiques de l'enseignement de l'évolution du vivant.

Florence Cousson-Gélie est professeure en psychologie de la santé et du développement de l'université Paul-Valéry-Montpellier 3. Elle est également directrice scientifique d'Epidaure, département de l'institut du cancer de Montpellier (ICM). Le cadre général de ses recherches est d'étudier le processus dynamique de la relation entre l'individu et son environnement, en particulier lorsque l'individu est confronté à un cancer.

Élodie Delahaye est ingénieure pédagogique innovation, chargée de mission sur le projet PIA3 Hybrid Innovative Learning Lab (HILL) et ingénieure agroalimentaire à ONIRIS.

Vanessa Douet-Vannucci est chercheuse clinicienne (CNU 69, ED85). Elle combine expérience universitaire et entrepreneuriale pour examiner les potentialités de l'épidémiologie numérique dans la compréhension et l'approche transdiagnostique des troubles du neurodéveloppement et de santé mentale associés chez les enfants et adolescents.

Valentine Facque est docteure en neurosciences cognitives et cheffe de projet en innovation chez Humans Matter. Elle contribue notamment à la création de parcours innovants et à la mesure objective de fonctions cognitives.

Mathilde Favier est cheffe de projet du pôle pilote Ampiric. Elle en coordonne le pilotage stratégique, opérationnel, et scientifique. Elle s'interroge sur la mise en œuvre de l'interdisciplinarité en éducation, et la facilitation de conditions propices à un dialogue constructif et apaisé entre les différentes traditions disciplinaires autour des questions d'éducation.

Jean-Marc Ferrandi est professeur de marketing à Oniris. Il participe au pilotage du PIA3 HILL. Cofondateur de l'Institut du marketing social, il est coresponsable de la chaire MARÉSON « Marketing responsable et bien-être ». Ses recherches, au sein du Laboratoire d'économie et de management Nantes-Atlantique (LEMNA), portent sur les fondements psychologiques et culturels du comportement du consommateur (distance psychologique, changement comportemental) et sur l'éco-innovation.

Jean-Philippe Lachaux est directeur de recherche à l'INSERM, et travaille au Centre de recherche en neurosciences de Lyon. Ses travaux concernent les mécanismes neuronaux de l'attention et de la concentration. En dehors de son travail de recherche, il mène une action sur le long terme pour promouvoir la maîtrise douce de l'attention comme une valeur dans notre société, notamment à travers ATOLE, un programme de découverte et de développement de l'attention à destination des élèves.

Laurent Lavogiez est directeur conception de Sciado Partenaires. Il est expert en formation et certification, avec plus de vingt ans d'expérience en gamification et simulation pédagogique. Il a créé plusieurs entreprises spécialisées dans le domaine de la pédagogie ludique, au sein desquelles il a occupé différentes fonctions telles que la modélisation, le développement d'outils et l'animation de dispositifs pédagogiques. Il a également travaillé avec des universités, des écoles d'ingénieurs et de management pour les aider à mettre en place des programmes de formation efficaces et innovants.

Sandrine Le Sourn-Bissaoui est maître de conférences en psychologie à l'université Rennes 2. Elle est affiliée au Laboratoire de Psychologie, Cognition, Comportement, Communication (UR1285, LP3C) et membre de la FHU EXcellence Center for Autism – Tours (EXac-t). Dans un contexte sociétal inclusif, ses recherches investiguent les trajectoires développementales des personnes avec un trouble du neurodéveloppement, et plus spécifiquement d'un trouble du spectre de l'autisme, à différents âges de la vie.

Riadh Lebib est docteur en neurosciences, chargé de cours à Sciences Po et l'EFAP. Designer cognitif chez Humans Matter, son travail consiste à imaginer et développer des solutions opérationnelles pour les professionnels de l'humain (santé, éducation, travail, société), qui s'appuient sur les connaissances théoriques issues des sciences cognitives, des sciences du comportement, des sciences de l'économie, ou encore de la sociologie.

François Lecellier est maître de conférences en traitement du signal et des images à l'université de Poitiers au sein de l'Institut Xlim (UMR CNRS 7252). Il enseigne l'informatique et la gestion de projet. Ses activités de recherche actuelles concernent la mesure et l'analyse de l'attention visuelle des utilisateurs en situation d'apprentissage. Il est également vice-représentant du groupe de travail « Éducation Innovation » au sein du groupe de Coimbra, et chargé de mission pour l'accompagnement renforcé des étudiants à l'Université.

Théo Marchand est titulaire d'un master en sciences cognitives. Il effectue une thèse CIFRE, intitulée « Combining innovative mobile electroencephalographic devices and gamified tasks to pre-diagnose neurodevelopmental disorders in ecological context », sous la codirection de Rod O'Connor du laboratoire de bioélectronique (ED SIS 488488) et Vanessa Douet-Vannucci du département R&D de O-Kidia.

Nicolas Mascret est professeur à l'Inspé d'Aix-Marseille Université. Il effectue ses recherches en psychologie sociale et en psychologie sociale expérimentale au sein de l'Institut des sciences du mouvement de Marseille (UMR7287, ISM, AMU-CNRS), dont il est directeur adjoint. Elles portent notamment sur la motivation, le stress, l'anxiété et l'acceptabilité des technologies, qui sont étudiés dans le domaine scolaire et dans le domaine du sport et de l'activité physique. Nicolas Mascret est le responsable scientifique et technique du pôle pilote Ampiric.

Jonathan Mirault est ingénieur de recherche en expérimentation et analyse de données au sein du pôle pilote Ampiric. Il conçoit et met en place des expériences scientifiques, collecte et analyse les données recueillies sur les projets déployés par le pôle pilote. Il se focalise particulièrement sur l'utilisation de méthodologies telles que l'oculométrie, l'électro-encéphalographie, la réalité virtuelle, la chronométrie mentale et la psychophysique.

Lionel Muniglia est maître de conférences en chimie alimentaire et innovation, au Laboratoire d'ingénierie des biomolécules, à l'ENSAIA-Université de Lorraine de Nancy. Il collabore au PIA3 HILL sur les questions de pédagogie innovante. Lionel Muniglia est aussi directeur scientifique de la start-up Bioli. Ses projets portent notamment sur l'éco-innovation en alimentaire et en cosmétique.

Violette Nemessany est docteure en sociologie et ethnologie. Elle occupe un poste de directrice de projets à l'ANRT. Auparavant, elle a occupé un

poste d'attachée temporaire d'enseignement et de recherche (ATER), puis d'ingénieure de recherche à l'université de Lorraine. Elle a ensuite travaillé dans une société d'études et de conseil en marketing. Depuis 2018, elle coordonne le groupe de travail « Pédagogie par le jeu » de l'ANRT et, depuis 2020, le CréativLab du pôle pilote Ampiric.

Nesrine Rahmouni est doctorante et ingénieure de recherche chez Sciado Partenaires. Ses travaux de recherche se concentrent principalement sur les environnements informatiques de l'apprentissage, en particulier sur la personnalisation de l'apprentissage en utilisant des méthodes automatiques et d'intelligence artificielle.

Franck Tarpin-Bernard est directeur général délégué de Humans Matter. Ingénieur de formation et passionné d'innovations, il s'est tout d'abord tourné vers le monde de la recherche dans les nouvelles technologies d'information et de communication, en embrassant une carrière universitaire d'enseignant-chercheur en informatique. Spécialiste de l'interaction homme-machine et des technologies interactives d'apprentissage humain, auteur d'une cinquantaine de publications, il s'est progressivement rapproché des sciences cognitives. En 2000, il a cofondé la société SBT Humans Matter, qui marie le digital et les sciences cognitives au sein d'une approche dite de design cognitif, afin d'accompagner la transformation des comportements dans les secteurs de la santé, de l'éducation et des organisations. Depuis 2021, Franck Tarpin-Bernard préside le groupe de travail « Pédagogie par le jeu » de l'ANRT.

Pascale Toscani est docteure en psychologie cognitive. Elle a été maître de conférences en psychologie cognitive à l'université catholique de l'Ouest (UCO) d'Angers pendant 12 ans. Elle est aujourd'hui chercheuse associée au Laboratoire du LIRDEF, à l'université de Montpellier. Elle a créé le Laboratoire collaboratif indépendant multidisciplinaire et multiprofessionnel du GRENE MONDE (Neurosciences cognitives, apprentissages et transitions éducatives).

Laurence Vanin est philosophe, essayiste et auteure de nombreux ouvrages. Docteure en philosophie politique, épistémologie, et habilitée à diriger les recherches en épistémologie et histoire des sciences (HDR, CNU72), elle enseigne à l'université Côte d'Azur, et dirige la chaire « Smart-City : philosophie et éthique », dans laquelle elle porte une réflexion philosophique appliquée aux technosciences. Elle a notamment axé ses travaux sur les vulnérabilités et le numérique en santé (éthique des relations aux objets connectés et aux robots.).

Johannes Ziegler est chercheur en sciences cognitives, directeur de recherche au CNRS, directeur du Laboratoire de psychologie cognitive, directeur de l'Institut Convergences « Langage, communication et cerveau » à Aix-Marseille Université. Il coordonne les actions de recherches fondamentales au sein du pôle pilote Ampiric. Ses recherches portent sur l'apprentis-

sage de la lecture, ses dysfonctionnements et ses bases neurales, ainsi que le développement et l'évaluation d'outils d'entraînement et de remédiation. Johannes Ziegler est membre du Conseil scientifique de l'éducation nationale (CSEN).

INTRODUCTION

Franck Tarpin-Bernard

Directeur général délégué, Humans Matter,
président du groupe de travail « Pédagogie par le jeu », ANRT

L'arrivée des technologies digitales, tout comme l'apparition de pédagogies alternatives au modèle traditionnel descendant ont ouvert un champ des possibles extrêmement riche. Les formateurs, qu'ils soient enseignants en formation initiale ou continue, se voient proposer une multitude de dispositifs ludopédagogiques présentant des caractéristiques très variées. Face à cette profusion, il nous a semblé, au sein du groupe de travail « Pédagogie par le jeu » de l'ANRT, que nous pouvions croiser les expériences et les connaissances de nos membres académiques et socio-économiques pour revisiter la question de l'évaluation de ces dispositifs.

Cet ouvrage rassemble dix articles proposés par certains de nos membres, ayant fait l'objet de relectures croisées exigeantes. Nous les avons rassemblés en trois parties, qui éclairent différentes problématiques liées à l'évaluation des dispositifs pédagogiques innovants.

La première partie pose le cadre de nos travaux. Violette Nemessany présente, dans son article, une synthèse des travaux du groupe de travail autour de cette question de l'évaluation. Tout d'abord, elle y décrit les enjeux et les motivations qui poussent les acteurs à évaluer les jeux pédagogiques. Puis, elle présente les différentes méthodes d'évaluation des dispositifs pédagogiques ludiques, les difficultés rencontrées par les acteurs, et les impacts de ces évaluations. En conclusion, les membres de l'ANRT ont identifié douze critères indispensables à l'évaluation d'un jeu pédagogique. Ils constituent une grille devant permettre à tout praticien ou décideur d'éclairer son choix de déployer tel ou tel dispositif.

Nicolas Mascret et ses coauteurs proposent ensuite d'explorer les questions du développement, de l'expérimentation et de l'évaluation de dispositifs innovants à travers l'exemple du CréativLab du pôle pilote Ampiric et le défi du passage à l'échelle. Ces questions s'inscrivent dans un processus complexe lorsque les évaluations sont réalisées selon une démarche scientifique rigoureuse, et lorsque l'ambition est d'envisager le déploiement à grande échelle des dispositifs. L'article décrit les procédures et modalités d'évaluation de ces dispositifs pédagogiques innovants, issus d'une collaboration entre entreprises, chercheurs et enseignants. Les enjeux du passage à l'échelle de ces dispositifs sont également examinés dans le contexte de l'éducation fondée sur les données probantes, tout en identifiant les limites liées à la pérennisation et la généralisation de ces dispositifs, ainsi que les limites liées à l'influence des prescriptions chez les enseignants.

Le cadre étant posé, les trois articles suivants, qui composent la deuxième partie de l'ouvrage, s'intéressent à différents outils innovants d'évaluation. Le premier article, écrit par Pascale Toscani, traite de la question difficile de l'évaluation de l'impact de l'introduction des neurosciences cognitives dans les pratiques pédagogiques. Pour intégrer les neurosciences cognitives dans leurs pratiques, les enseignants ont adopté des choix professionnels différents. Certains ont été en recherche d'outils, de techniques, de modèles types, à appliquer « clés en main » dans leurs classes. D'autres ont souhaité être accompagnés par des laboratoires de neurosciences pour appliquer, dans des protocoles, des recherches tout autant « clés en main ». Ce constat pose un problème dans les deux cas : les enseignants ne connaissent pas les neurosciences, et les chercheurs ne connaissent pas la classe. Dès lors, comment évaluer des dispositifs dont on est censé mesurer l'efficacité dans la classe, à travers des critères et indicateurs stables ? L'article propose un état des lieux des dispositifs d'intégration des neurosciences cognitives à l'école, et pose la problématique de l'évaluation complexe de ces dispositifs qui impliquent de puissants paradigmes d'apprentissage.

Stéphanie de Chalvron et ses collègues nous invitent ensuite à plonger dans les notions d'« agentivité » et de « sentiment d'efficacité personnelle » (SEP), et à étudier le rôle joué par le stress perçu dans la relation qui unit ces deux notions. En utilisant une méthodologie statistique poussée, les auteurs démontrent que les formations professionnelles visant à améliorer l'agentivité devraient en priorité s'attacher à développer le SEP des participants. À travers des outils de mesure faciles à déployer, ils proposent de les mobiliser systématiquement pour évaluer l'impact des dispositifs pédagogiques.

L'article rédigé par Julian Alvarez présente le modèle CEPAJe, un outil conçu pour évaluer une séance ludopédagogique. À partir d'exemples concrets d'emploi de jeux sérieux (serious games), l'auteur montre comment ce modèle permet de préparer, en amont, une séance ludopédagogique, puis, en aval, d'évaluer cette dernière afin de pouvoir améliorer l'activité pour la fois prochaine.

Pour clôturer cette deuxième partie dédiée aux outils, Théo Marchand et ses collègues analysent la notion de biomarqueur numérique. Avec l'avènement des technologies numériques, les biomarqueurs issus de la collecte massive de traces constituent un nouveau type de donnée pouvant traduire les comportements et les processus cognitifs. Depuis plusieurs décennies, l'industrie du jeu vidéo a développé des méthodologies d'évaluation pour maximiser l'expérience utilisateur. Aujourd'hui, une approche hybride d'évaluation combinant jeu vidéo et biomarqueurs numériques émerge, afin d'examiner l'usabilité des contenus pédagogiques, leur pertinence, et leur efficacité sur les performances de l'utilisateur.

La troisième partie de cet ouvrage met en lumière la problématique de l'évaluation des dispositifs pédagogiques innovants à travers quatre expériences extrêmement variées. L'article proposé par Olivia Collet et ses collègues se

concentre sur la création et l'évaluation d'un programme pédagogique inclusif basé sur l'activité physique et le jeu sportif à destination d'enfants ayant un trouble du spectre de l'autisme (TSA). En effet, si l'activité physique semble constituer un moyen efficace pour améliorer les compétences sociales des enfants présentant un TSA, les professionnels ne proposent pas toujours les activités physiques en lien avec cet objectif. Par ailleurs, l'évaluation de l'effet de telles activités est rarement étudiée. Ce texte vise à présenter la démarche de cocréation d'un programme pédagogique inclusif, basé sur l'activité physique et le jeu sportif, à destination d'enfants présentant un TSA. Il s'agit d'évaluer l'effet de ce programme aussi bien pour les enfants avec un TSA que pour les enfants neurotypiques, leurs parents, et les professionnels impliqués dans le dispositif.

François Lecellier décrit trois manières distinctes d'évaluer l'utilisation du jeu vidéo Minecraft au sein de l'université de Poitiers. D'abord, au moyen des notes obtenues par les étudiants traduisant leur acquisition de compétences. Puis, par le biais de questionnaires anonymes transmis aux étudiants. Et enfin, au travers de la mesure de l'attention visuelle des étudiants lors des séances. L'ensemble des trois évaluations montre une augmentation de l'attention et de l'engagement des étudiants pendant l'utilisation du jeu, par rapport à des séances plus traditionnelles.

Quant à Élodie Delahaye et ses collègues, ils explorent la place des jeux au sein de la formation ingénieur agroalimentaire. À partir d'une enquête de grande ampleur, ils caractérisent la place des dispositifs ludopédagogiques dans la formation, et identifient les motivations, attentes et craintes sous-jacentes différenciées des apprenants et des formateurs. L'étude souligne toute l'importance de proposer une évaluation adaptée des dispositifs selon l'acteur. Elle permet de poser les bases des indicateurs à intégrer pour faciliter l'adhésion aux dispositifs ludiques, et évaluer leur efficacité en termes d'apprentissage.

Pour clore cet ouvrage, Nesrine Rahmouni et Laurent Lavogiez s'intéressent aux simulateurs pédagogiques. Si ces simulateurs sont des environnements particulièrement utiles pour aider à l'apprentissage par la pratique, il s'avère nécessaire de travailler sur leur personnalisation pour répondre aux besoins individuels de chaque apprenant. L'article décrit alors un module pour permettre aux formateurs d'obtenir des informations sur le processus d'apprentissage de chaque apprenant, afin de lui offrir une expérience adaptée. L'évaluation d'un dispositif de personnalisation est une tâche difficile, puisque, par définition, chaque expérience individuelle est unique.

Voilà, nous espérons que ce voyage autour de la question de l'évaluation de dispositifs pédagogiques innovants vous ouvrira de nouvelles perspectives. En tout cas, nous avons pris beaucoup de plaisir à l'élaborer, en partageant nos diverses et enrichissantes expériences en tant que formateurs, chercheurs et industriels.

Bonne lecture.

PARTIE I

Cadrage

Évaluer les dispositifs pédagogiques ludiques

Violette Nemessany

Résumé

Comment mesurer l'efficacité des jeux pédagogiques dans le domaine de l'éducation, l'enseignement et la formation ? Quels sont les critères d'évaluation qui fournissent une mesure fiable de la performance atteinte, et des compétences socio-comportementales acquises par les apprenants ? Autant de questions auxquelles les membres du groupe de travail « Pédagogie par le jeu » de l'ANRT ont apporté des éléments de réponse à l'occasion des travaux menés en 2022. Cet article décrit, en premier lieu, les enjeux et les motivations qui poussent les acteurs à évaluer les jeux pédagogiques. En second lieu, il présente les différentes méthodes d'évaluation des dispositifs pédagogiques ludiques, les difficultés rencontrées par les acteurs, et les impacts de ces évaluations. Pour conclure, les membres de l'ANRT ont identifié douze critères indispensables à l'évaluation d'un jeu pédagogique.

INTRODUCTION

Depuis plusieurs années, on assiste, en France, à un développement des jeux pédagogiques dans le domaine de l'éducation, l'enseignement et la formation. Ils sont, le plus souvent, le fruit de l'initiative individuelle et isolée de pédagogues passionnés, soucieux d'accroître l'efficacité de leur action par des moyens originaux. Mais leur multiplication est aussi le résultat de collaborations, de plus en plus fréquentes, qui s'établissent d'une part entre des enseignants, des formateurs, et des chercheurs qui poursuivent des objectifs pédagogiques scientifiques, et d'autre part, des professionnels du jeu, et des responsables de structures socio-économiques à la recherche de produits innovants. Cette dynamique à la fois pédagogique, scientifique et économique fait surgir la question de l'évaluation des jeux pédagogiques. Comment mesurer précisément ces innovations, et distinguer, dans les nouveaux jeux mis au point, les plus susceptibles d'enrichir effectivement les pratiques d'enseignement, d'éducation et de formation ? Il s'agit d'une étape incontournable pour « (re)penser, concevoir, mettre en œuvre, opérationnaliser, améliorer »[1] et ainsi, légitimer les jeux pédagogiques qui méritent d'être diffusés à grande échelle.

Une chose est sûre : le seul jugement d'un des acteurs de l'innovation pédagogique — enseignant, chercheur ou développeur — ne suffit pas pour prouver la pertinence, ni la valeur d'usage, du jeu proposé. Cette mesure exige la prise en compte du point de vue de chacun des différents acteurs concernés, mis en situation d'éprouver pratiquement la justesse d'une innovation. Mais alors, comment faire ? Existe-t-il des critères d'évaluation, qui fournissent une mesure fiable de la performance atteinte et des compétences socio-comportementales acquises par les apprenants ? Quelles sont les conditions à respecter pour garantir la rigueur et l'impartialité des résultats ? Comment les évaluations sont-elles perçues et acceptées par les pédagogues, leur hiérarchie, les parents et les concepteurs des dispositifs ludiques ? Au-delà, contribuent-elles à faire évoluer les pratiques d'apprentissage ? Parviennent-elles à convaincre les responsables institutionnels et politiques de réformer le système éducatif ?

Autant de questions auxquelles les membres de l'ANRT s'efforcent d'apporter des éléments de réponse dans le cadre de leur groupe de travail sur la « Pédagogie par le jeu ». Après avoir décrit les enjeux et les motivations qui poussent les acteurs à évaluer les jeux pédagogiques, la première partie de cet article présentera les différentes méthodes d'évaluation, chacune reposant sur des règles quantitatives, qualitatives et éthiques qui lui sont propres. La deuxième partie éclairera les obstacles et les résistances auxquels les enseignants, les chercheurs et les responsables de structures so-

1. Anne Jorro, Nathalie Droyer (ss la dir.), (2019), *L'Évaluation, levier pour l'enseignement et la formation,* De Boeck Supérieur, p. 8.

cio-économiques sont confrontés dans le cadre des évaluations des dispositifs pédagogiques à caractère ludique. Elle soulignera également les efforts d'acculturation et d'ajustement nécessaires pour les amenuiser. Enfin, la troisième partie analysera comment les résultats des évaluations des jeux pédagogiques permettent d'accompagner les transformations pédagogiques, techniques, culturelles et politiques.

I. MULTIPLES STRATÉGIES D'ÉVALUATION

1. OBJECTIFS ET MOTIVATIONS

Évaluer l'efficacité des jeux pédagogiques est devenu une préoccupation majeure, tant pour leurs concepteurs que pour les utilisateurs et les institutions publiques. L'objectif est « d'analyser précisément quels effets produisent ces dispositifs ludiques sur l'ensemble de la communauté éducative afin d'identifier les initiatives qui mériteraient d'être généralisées »[2]. Pour y parvenir, « le processus d'évaluation consiste à recueillir un ensemble d'informations pertinentes à confronter, par une démarche adéquate, à un ensemble de critères valides dans l'opérationnalisation, et fiables dans l'utilisation. Les résultats de cette confrontation doivent permettre une prise de décision cohérente avec la fonction visée par l'évaluation »[3]. Cette dernière peut prendre des formes différentes selon les besoins et les enjeux auxquels sont confrontés les acteurs. Par exemple :

- Du point de vue des *pédagogues-maîtres du jeu,* l'évaluation répond à une stratégie de légitimation. Elle fournit les preuves susceptibles de faire reconnaître, auprès non seulement des apprenants et de leur famille, mais aussi des collègues et de la hiérarchie :

  1) leurs compétences en tant que concepteurs de jeux pédagogiques, qui, pour la plupart, ont été conçus pendant leur temps de loisir, en dehors de toute reconnaissance académique,

  2) et la valeur éducative de ces jeux dans le cadre de la pratique pédagogique.

- Du côté, des *entreprises de la EdTech,* l'évaluation s'inscrit dans une stratégie d'innovation. Elle vise à accompagner la transformation des dispositifs pédagogiques ludiques en cours de conception. Dans

2. Eric Sanchez, Margarida Romero (2020), *Mythes et réalités : apprendre en jouant,* Éditions Retz, p. 123.

3. Olivier Rey, Annie Freyfant (2014), « Évaluer pour (mieux) faire apprendre », *Dossier de veille de l'Institut français de l'éducation,* n° 94, septembre, p. 7.

cette perspective, les indicateurs d'évaluation doivent permettre de « donner des précisions sur la qualité pédagogique des outils, leur potentiel ludique ainsi que leur utilité dans un contexte de formation. Ils servent également à mesurer le respect du cahier des charges initial et notamment, le fait que toutes les compétences ciblées au début du projet sont bien intégrées dans le scénario »[4]. En bref, l'évaluation aide les entreprises à « apprécier le potentiel d'un concept innovant, mesurer une performance atteinte à un certain stade de maturité du projet »[5], « repérer les difficultés de réalisation et remédier aux problèmes au plus vite afin de réduire le temps et le coût de conception »[6].

- En ce qui concerne les *familles,* l'évaluation est au service d'une stratégie d'intéressement. Elle leur offre des connaissances capables d'orienter leur choix vers les dispositifs ludiques les plus appropriés pour les besoins de leurs enfants. Dans ce cas, l'évaluation répond à un souci de modernisation des méthodes d'apprentissage, des modes de transmission des savoirs, et de la formation aux métiers de demain.

- Enfin, pour les *politiques publiques,* l'évaluation s'intègre dans une stratégie de financement, de communication et de contrôle. Elle peut servir, à la fois, à justifier les investissements publics dans les entreprises du secteur de la EdTech, qui proposent des solutions numériques susceptibles d'améliorer l'expérience de l'apprentissage, à en mesurer les effets, et à éclairer les orientations à donner à l'action publique.

2. SE DONNER UN AVAL ACADÉMIQUE

Pour que les résultats d'une évaluation soient valides et fiables, il importe de suivre un protocole scientifique rigoureux. C'est pourquoi la collaboration avec des chercheurs issus des établissements publics de recherche est indispensable. Ceux-ci sont engagés à concevoir, coordonner, et valoriser des méthodes d'évaluation rigoureuses reposant sur des règles quantitatives, qualitatives et éthiques, indispensables pour mesurer l'efficacité des dispositifs pédagogiques innovants. En effet, dans le cadre de leur activité de recherche, les chercheurs s'engagent à respecter quatre principes fon-

---

4. Iza Marfisi-Schottman, Sébastien George, Franck Tarpin-Bernard, Patrick Prévot (2012), « Comment évaluer la qualité d'un learning game pendant sa conception ? », *Conférence Technologies de l'Information et de la Communication pour l'Enseignement,* TICE 2012, décembre, p. 4.

5. Axelle Charpentier (2020), « Comment évaluer les innovations pédagogiques », in Sciences humaines, « Bonheur d'apprendre et d'enseigner. Enquête sur ces profs qui changent l'école », *Les grands dossiers de sciences humaines,* n° 58, mars-avril-mai, p. 77.

6. Iza Marfisi-Schottman, Sébastien George, Franck Tarpin-Bernard, Patrick Prévot (2012), *Ibid.*, p. 4.

damentaux. Ces principes ont été énoncés dans la Charte européenne du chercheur, en 2005 :

- *« la fiabilité :* visant à garantir la qualité de la recherche dans la conception, la méthodologie, l'analyse et l'utilisation des ressources ;

- *l'honnêteté :* qui consiste à élaborer, entreprendre, évaluer, déclarer et faire connaître la recherche d'une manière transparente, juste, complète et objective ;

- *l'intégrité scientifique :* qui signifie le refus de se plier à des pressions financières, sociales ou politiques ;

- *et la responsabilité :* tout au long de l'activité de recherche, de l'idée à la publication, en passant par l'organisation et la gestion du projet »[7].

Plusieurs motivations justifient le rapprochement entre le monde éducatif, économique et académique. Pour les pédagogues, le partenariat avec les chercheurs « peut leur permettre de comprendre de l'intérieur les méthodes utilisées par la science, leurs atouts et leurs limites. L'interaction structurée entre chercheurs et éducateurs favorise une culture commune de l'éducation basée sur des preuves. Il ne faut pas oublier que beaucoup d'enseignants innovent au quotidien en testant de nouvelles pratiques dans leur classe. Ils sont les premiers observateurs des effets des innovations pédagogiques et techniques »[8].

Inversement, du point de vue des chercheurs, les collaborations avec des enseignants et des responsables des entreprises du secteur de la EdTech leur offrent l'avantage d'accéder à des connaissances, des compétences ainsi que des moyens techniques et financiers supplémentaires. « Dans certains cas, cela peut être une source de revenus récurrente pour l'établissement public de recherche ». Le partenariat multiplie aussi les occasions de renouveler des thématiques de recherche. En effet, les échanges avec les pédagogues et les entreprises permettent « d'identifier des sujets porteurs, d'orienter des stratégies de recherche futures, voire de faire pro-

---

7. « Ces principes ont été explicités dans la déclaration de Singapour sur l'intégrité en recherche en 2010, dont le préambule souligne que : « quelle que soit la manière dont la recherche est menée et organisée selon les disciplines et les pays, il existe des principes communs et des obligations professionnelles similaires qui constituent le fondement de l'intégrité en recherche où qu'elle soit menée. », in Lucienne Letellier (2014), « Pratiquer une recherche intègre et responsable », Guide, *Comité d'éthique du CNRS*, p. 4, www.cnrs.fr/comets ; ALLEA (2018), « Code de conduite européen pour l'intégrité en recherche », Édition révisée, p. 4, www.allea.org.

8. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), « La recherche translationnelle en éducation », Rapport, Conseil scientifique de l'éducation nationale, MENJS, Réseau Canopé, p. 12, www.education.gouv.fr.

gresser les connaissances sur des projets originaux, interdisciplinaires et à dimension internationale »[9].

Cette nécessité de « créer un lien étroit entre la recherche publique, l'éducation et les entreprises » n'est pas démentie par Jean-Marc Monteil, de la mission interministérielle Éducation et Numérique. Il ajoute que « de véritables travaux de recherche pluridisciplinaire doivent être entrepris pour mesurer précisément l'efficacité de ces outils ludiques ». Or, « à ce jour, aucune étude scientifique n'a réellement fait la preuve de leur impact ». Jean-Marc Monteil regrette que « les enquêtes existant sur ces nouvelles façons d'apprendre portent sur de trop petits échantillons. Elles ressemblent plus à des enquêtes de satisfaction ou d'opinion qu'à des études de performance »[10]. C'est la raison pour laquelle il est impossible « de convaincre les acteurs de l'innovation du caractère extrapolable des résultats de ces études dans d'autres environnements pédagogiques »[11].

3. VERS DES MÉTHODOLOGIES D'ÉVALUATION ADAPTABLES

Il existe aujourd'hui une grande variété de méthodes d'évaluation (observations, entretiens, comparaisons, questionnaires, études expérimentales, etc.). « La pratique s'est institutionnalisée avec le déploiement d'évaluations standardisées ». Toutefois, la plupart de ces méthodes reposent sur des normes qui « sanctionnent les «bons» résultats selon les critères définis à l'avance ». Elles ne permettent pas de questionner l'efficacité des jeux pédagogiques sur :

- le développement des compétences transverses, comme l'autonomie, la créativité, l'esprit critique, l'acceptation de l'échec, etc. ;
- le sentiment d'efficacité personnelle des apprenants ;
- et leur capacité d'agir (l'agentivité).

Il est par conséquent nécessaire de développer des instruments de mesure adaptés au cas par cas, capables « d'accompagner les transformations, d'identifier les conséquences inattendues ou de justifier les options prises au cours d'un processus d'innovation »[12].

---

9. Violette Nemessany (2016), « Regards croisés sur la recherche partenariale. Comment se construisent les relations entre les établissements publics de recherche et les entreprises », Rapport, ANRT, p. 23-24.

10. Mireille Broussous, Centre-Inffo, « L'apprentissage par le jeu est-il efficace ? », 18 février 2020, www.centre-inffo.fr/.

11. Axelle Charpentier (2020), *op. cit.,* p. 75.

12. Axelle Charpentier (2020), *op. cit.,* p. 77.

La recherche translationnelle

Dans le rapport « La recherche translationnelle en éducation » publié par le Conseil scientifique de l'éducation nationale[13], les auteurs soulignent l'importance de :

> « renforcer, en éducation, la place de la recherche dite «translationnelle». Ce terme trouve son origine en médecine, où il indique le mariage de la recherche fondamentale avec celle du terrain. Car la recherche translationnelle est à deux voies : elle permet à la fois d'identifier des méthodes efficaces d'intervention et d'augmenter les connaissances fondamentales, dans un aller-retour constant entre les deux. Mais son objectif fondamental reste celui de proposer et d'évaluer des solutions aux problèmes concrets pour qu'elles soient fondées sur les meilleures preuves disponibles »[14].

Ce constat est partagé par Johannes Ziegler, directeur de recherche au CNRS et du Laboratoire de psychologie cognitive à Aix-Marseille Université[15]. Dans le cadre d'une réunion du groupe de travail de l'ANRT[16], il va plus loin en proposant de se référer à la « pyramide des preuves ». Il s'agit :

> « d'une liste de méthodes de recherche permettant d'apporter des preuves. À la base de la pyramide se trouvent les méthodes qui donnent les résultats les plus incertains. En se déplaçant vers le haut de la pyramide, on trouve les méthodes qui permettent de contrôler en partie ou totalement les biais et donc d'arriver à un plus haut niveau degré de certitude »[17].

Johannes Ziegler poursuit en décrivant les quatre niveaux de la « pyramide des preuves » :

> 1) *« Les observations informelles :* qui peuvent prendre la forme d'un essai réalisé par un pédagogue dans sa classe, le recueil d'un témoignage, ou la passation d'un questionnaire. Elles sont utiles dans la première phase d'un projet, également dénommée "étude pilote". Mais du fait de leur caractère subjectif, elles possèdent une faible valeur de preuve.

---

13. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), *op. cit.,* p. 12.

14. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), *op. cit.,* p. 3-4.

15. Johannes Ziegler est aussi membre du Conseil scientifique de l'éducation nationale.

16. Réunion du groupe de travail de l'ANRT, sous la présidence de Franck Tarpin-Bernard, « Comment évaluer les impacts d'un dispositif pédagogique ? », 1er juin 2022.

17. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), *op. cit.,* p. 21.

2) *L'étude quasi expérimentale :* elle compare au moins deux groupes, avec des effectifs faibles et moyens. Le but est de recueillir des mesures objectives. Cette étude s'inscrit dans une expérimentation préliminaire.

3) *L'étude expérimentale ou "randomisée contrôlée" :* est à privilégier auprès d'un effectif suffisamment important pour obtenir un niveau de preuve élevé. Elle en tire des enseignements, des données statistiques, etc. Elle consiste à comparer deux groupes (expérimental d'un côté, et témoin de l'autre) constitués par tirage aléatoire. Ce type d'étude peut être long et coûteux en termes de contraintes administratives, de ressources humaines et financières. Pour être menée à bien, elle doit être soigneusement planifiée et la collaboration avec les chercheurs est indispensable.

4) *L'étude de passage à l'échelle :* doit être déployée dès lors que les résultats des études randomisées contrôlées se sont avérés concluants. Elle concerne l'expérimentation des dispositifs auprès de très grands effectifs, sur le plan régional ou national. Elle implique la formation de nombreux intervenants entièrement indépendants des chercheurs ».

Pour conclure, « la référence à cette pyramide des preuves permet de mobiliser une grande variété de méthodes. Loin d'être en concurrence les unes avec les autres, ces méthodes créent des synergies et soutiennent collectivement la découverte de moyens pour améliorer l'éducation à tous les niveaux. Il serait contre-productif d'affirmer que seuls une méthode ou un petit nombre de méthodes ont de la valeur dans le domaine de l'éducation fondée sur les preuves. Il est souvent utile de commencer par des études observationnelles, avant de mener des études randomisées contrôlées »[18].

II. LES OBSTACLES À L'ÉVALUATION DES DISPOSITIFS PÉDAGOGIQUES LUDIQUES

1. DIFFICULTÉS ET RÉSISTANCES

L'évaluation de l'efficacité des dispositifs ludiques en milieu scolaire, universitaire et professionnel n'est pas sans poser de nombreuses difficultés. Certaines concernent la définition des critères d'évaluation. Ces critères jouent un rôle normatif. Ils doivent être déterminés par la finalité de l'évaluation, adaptés au contexte, cohérents avec les objectifs, traduits en indicateurs

18. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), *op. cit.,* p. 23.

quantitatifs et qualitatifs[19]. Reste que dans de nombreux cas, « force est de constater que les critères ne font pas toujours sens pour ceux qui les suivent parce qu'ils ne répondent pas nécessairement à des préoccupations de terrain, et peut-être aussi, parce que ceux qui sont en charge de les mesurer n'ont généralement pas participé à leur définition et ne s'en approprient pas les résultats »[20].

ENCADRÉ 1 — Exemple d'une bonne pratique : construire des critères d'évaluation valides grâce à la « méthode de Churchill »[21]

Stéphanie de Chalvron, directrice scientifique chez Humans Matter, résume en quelques phrases une méthode de recueil de données développée en 2010 :

> « Il s'agit de demander aux personnes interrogées d'aider à la construction d'un questionnaire en suivant les recommandations de la méthode de Churchill. Le but est ici de récolter le plus d'avis et de verbatims possible. Le fichier est ensuite analysé selon la "méthode Q", qui permet de dégager les grands thèmes et de les catégoriser. Les assertions sont codées sur ces thèmes et une analyse factorielle permet d'extraire les saturations (poids factoriels) de chaque assertion (verbatim) et donc de les sélectionner, non pas sur la base de la fréquence, mais sur le pourcentage d'explication de variance de tel ou tel phénomène qu'une assertion proposée permet d'obtenir. Cette méthode, également dénommée méthode de "Churchill 2.0", se distingue de la méthode scientifique, validée par des publications mais elle s'en inspire et fait l'objet actuel de la rédaction d'un article scientifique. Elle a été expérimentée sur de très gros échantillons et sur divers sujets dont le "numérique à l'école" en collaboration avec l'université de Poitiers ».

D'autres difficultés résultent des contraintes méthodologiques. En effet, « la montée en puissance, depuis les années 1980, d'une demande générale pour que le système éducatif puisse "rendre des comptes" publiquement, régulièrement et à tous les niveaux, sur ses performances, a renforcé cette pression à la production d'indicateurs standardisés »[22]. Entre autres, cela a des conséquences sur les différentes méthodologies d'évaluation habituellement employées pour mesurer les effets des dispositifs pédagogiques innovants. Il est alors difficile de prendre en compte, non seulement les caractéristiques individuelles des apprenants et leurs besoins éducatifs parti-

19. OCDE (2020), « Des critères améliorés pour des évaluations meilleures », février, p. 1, www.oecd.org/fr.

20. Claudine Golka (2010), « L'évaluation, un levier du changement ? », *Bulletin des bibliothèques de France*, T. 55, n° 4, p. 12, https://bbf.enssib.fr.

21. Gary A. Churchill (1979), "A paradigm for developing better measures of marketing constructs", *Journal of Marketing Research,* 16, pp. 64–73.

22. Olivier Rey, Annie Freyfant (2014), *op. cit*., p. 4.

culiers, mais aussi le rôle du pédagogue-maître du jeu dans la réussite d'une séquence pédagogique ludique. À cet égard, plusieurs travaux scientifiques attestent que « l'usage d'un même jeu peut avoir des conséquences différentes selon la manière dont il est utilisé dans une situation pédagogique »[23]. En un mot, aucun dispositif innovant « n'est valable pour tous les publics ou toutes les disciplines indépendamment des contextes sociaux et institutionnels dans lesquels elle se réalise. […] On dispose aussi de preuves de l'existence d'un effet propre des pratiques enseignantes »[24]. De manière plus précise, il est observable que l'enthousiasme des pédagogues, leur créativité, leur habileté et l'intérêt qu'ils portent aux méthodes de pédagogie active exercent une influence, difficilement mesurable par les évaluations, sur le désir d'apprendre des individus[25]. À cela s'ajoutent les difficultés liées à l'évaluation des compétences socio-comportementales, telles que l'estime de soi, le sentiment d'efficacité personnelle, la capacité à agir, etc. Cette question est épineuse. En effet, ces savoir-être ne sont pas officiellement reconnus dans les programmes pédagogiques nationaux.

ENCADRÉ 2 — Exemple d'une bonne pratique : l'Observatoire dynamique de la capacité à agir pour mesurer le sentiment d'efficacité personnelle (SEP) et l'agentivité.

L'Observatoire dynamique de la capacité d'agir a été lancé par l'entreprise Humans Matter. Cet outil permet de mesurer des variables liées aux compétences psychologiques, cognitives et comportementales que les apprenants peuvent développer pendant la formation. L'objectif est d'améliorer le SEP et l'agentivité. Le concept du SEP a été développé par Albert Bandura dans les années 80. « Il correspond à la croyance des personnes en leur capacité personnelle à organiser et réaliser des ensembles d'actions requises pour atteindre des types de performances attendus ». Le SEP est le principal moteur de la capacité d'agir. « Il s'agit d'une compétence qui n'est pas innée, mais qui peut s'acquérir. Il est le moteur et le marqueur de l'agentivité, définie comme "la volonté de tout individu d'être auteur et acteur de son propre développement". L'agentivité désigne le fait pour un individu d'avoir conscience qu'il est lui-même l'auteur de ses actions. En d'autres termes, si l'agentivité d'une personne est élevée, c'est qu'elle est animée par la volonté d'agir par elle-même ».
Du point de vue de Stéphanie de Chalvron, directrice scientifique chez Humans Matter :

23. Eric Sanchez, Margarida Romero (2020), *op. cit.,* p. 123.

24.Pascal Bressoux (2012), *L'Influence des pratiques enseignantes sur les acquisitions scolaires des élèves,* La Découverte, « Regards croisés sur l'économie », 2012/2 n° 12, p. 210.

25. « Depuis quatre décennies, un ensemble de recherches menées dans le domaine de l'éducation a pu confirmer l'intuition de nombreux parents : les progrès de leurs enfants dépendent de manière significative du talent et des compétences de leurs professeurs ». C'est ce qu'on appelle « l'effet enseignant ». *Cf.* Pierre-Yves Cusset (2011), « Que disent les recherches sur "l'effet enseignant" ? », *La Note d'analyse,* n° 232, Centre d'analyse stratégique.

> « la formation professionnelle doit amener les individus à développer leur SEP, et en même temps, leur agentivité. C'est pourquoi l'entreprise Humans Matter s'est donné pour mission de créer l'Observatoire dynamique de la capacité d'agir. Celui-ci se propose de se focaliser sur trois axes :
>
> • *la prévention* : en détectant précocement des signaux faibles et des risques sur le facteur humain et la capacité d'agir ;
>
> • *l'impact* : en mesurant l'efficacité des dispositifs de formation et d'accompagnement ;
>
> • *le développement* : en orientant des stratégies de développement du facteur humain et d'évolution des comportements ».
>
> L'Observatoire a été testé auprès d'un grand nombre de personnes. Il a été développé autour de quatre temps de mesure : 1) avant ; 2) pendant ; 3) juste après ; 4) et au minimum trois ou quatre mois après la formation. Ceci, car il cherche à dégager non pas des mesures à des instants T, mais des trajectoires de développement selon une approche "lifespan" (collecte des données, quatre vagues de baromètre autour d'un même groupe d'individus sur un an, voire plus) ».

Dans un autre registre, « si du point de vue économétrique, les évaluations apparaissent comme un moyen privilégié pour fournir une mesure fiable des effets des outils pédagogiques innovants, elles ne sont pas toujours acceptées par les acteurs de terrain »[26]. Des résistances humaines se font jour, comme en témoignent les chercheurs du Conseil scientifique de l'éducation nationale :

> « Certains s'opposent à l'idée même d'expérimentation en éducation, pour des raisons faussement éthiques : les élèves ne sont pas des "cobayes". Mais cet argument ignore le fait que chaque jour, les enseignants testent de nouvelles pratiques sur les élèves, sans que personne n'y trouve à redire. Malheureusement, ces tentatives se font selon une méthodologie qui ne permet pas de tirer de conclusions rigoureuses, et ne permet donc pas de faire avancer la connaissance au bénéfice de tous. D'un point de vue éthique, il est donc nettement préférable de soumettre moins d'élèves à de vraies expérimentations, guère différentes de celles qui ont lieu tous les jours, mais qui en plus permettent d'accumuler de la connaissance »[27].

2. LES ENJEUX ÉTHIQUES

L'évaluation des dispositifs pédagogiques innovants amène à s'interroger sur les questions éthiques. « Quelle est la valeur des outils d'évaluation utilisés ? Comment être assuré que ce sont les meilleurs outils d'évalua-

26. Axelle Charpentier, *op. cit*., p. 75.

27. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), *op. cit.,* p. 27.

tion, compte tenu des contenus d'apprentissage et du niveau cognitif des apprenants ? Qu'est-ce qui doit être évalué ? Comment être certain que le jugement évaluatif est neutre et impartial ? »[28].

Mais au-delà, il y a un autre défi à relever. La multiplication des évaluations produit une masse de données personnelles sur le comportement des apprenants. Comme le soulignent les chercheurs en sciences de l'éducation, il est indispensable que « la collecte et l'analyse de ces datas suivent des règles éthiques strictes »[29]. Dans cette perspective, il convient de mentionner :

- d'une part, le « cadre éthique de l'expérimentation pédagogique en situation réelle » du Comité national consultatif d'éthique[30], qui « prend en compte les risques liés aux expérimentations éducatives chez les apprenants et y répond avec des recommandations ciblées »[31] ;

- et d'autre part, l'installation du comité d'éthique pour les données d'éducation.

« Ces avancées récentes dans le domaine de l'éducation pourraient contribuer à la mise en œuvre d'une régulation éthique des évaluations »[32].

III. LES TRANSFORMATIONS ENGENDRÉES PAR LES ÉVALUATIONS

Ainsi, comme on a pu le constater, l'évaluation des dispositifs pédagogiques ludiques n'est pas un exercice facile. Elle se heurte à des écueils et des résistances. Pour les dépasser, trois conditions doivent être réunies. Il est indispensable que l'évaluation soit :

1) *coconstruite* dans le cadre d'un partenariat entre pédagogues, chercheurs et responsables de structures socio-économiques ;

---

28. Denis Jeffrey (2013), *L'Éthique dans l'évaluation scolaire.* Québec, Presses de l'université Laval, p. 12.

29. Laurent Tardif, Jérôme Clerget, Antoine Lafontaine, Anas Knefati (2021), « Les Learning Analytics et la transformation pédagogique et numérique : un champ en construction », in Collectif DESIR, *Transformations pédagogique et numérique dans l'enseignement supérieur. Quatre années pour changer les pratiques,* Presses des Mines, p. 65.

30. Cadre éthique de l'expérimentation pédagogique en situation réelle (2019), *cf.* www.ccne-ethique.fr.

31. Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021), *op. cit.,* p. 27.

32. Axelle Charpentier, *op. cit.*, p. 76.

2) *comprise* par les pédagogues ;

3) et *bien menée*, sur la base « des critères de crédibilité, fiabilité, transférabilité et de reproductibilité »[33].

Dès lors, les résultats des évaluations des dispositifs ludiques conduisent les acteurs à modifier les pratiques pédagogiques. Il est possible de distinguer quatre transformations caractéristiques : technique, pédagogique, culturelle et politique.

1. LA TRANSFORMATION TECHNIQUE

Du point de vue des concepteurs, l'évaluation de l'expérience utilisateur d'un jeu pédagogique apparaît comme un moyen privilégié de rendre compte des disparités d'usages pouvant être à l'origine des résistances vis-à-vis du dispositif ludique. Du même coup, les conclusions des évaluations permettent d'identifier ce sur quoi des efforts sont nécessaires pour améliorer techniquement le jeu. En d'autres termes, elles donnent la possibilité aux concepteurs de faire évoluer leur outil en mettant en place des moyens pour l'adapter aux objectifs pédagogiques et aux besoins des apprenants.

2. LA TRANSFORMATION PÉDAGOGIQUE

Lorsque les indicateurs d'évaluation valident l'efficacité d'un dispositif ludique, ils constituent, pour les pédagogues, une preuve de la qualité d'un jeu pédagogique. Cette appréciation positive contribue à assurer la pérennité de l'outil innovant et « va lui permettre de trouver des applications plus larges »[34]. En effet, elle offre la possibilité aux pédagogues de faire reconnaître d'autres façons de faire, d'autres pratiques et d'autres compétences. Les données recueillies dans le cadre des évaluations jouent, par conséquent, un rôle de levier. Ils rassurent la communauté éducative sur le bien-fondé des jeux pédagogiques ; et ils encouragent les plus sceptiques à faire évoluer leurs pratiques d'apprentissage en intégrant ces jeux à l'ensemble des outils pédagogiques.

3. LA TRANSFORMATION CULTURELLE

Tant pour les pédagogues que pour les responsables des entreprises du

33. Document de travail de l'ANRT (2022), Pascale Toscani, « L'évaluation des dispositifs pédagogiques innovants, en neurosciences cognitives et apprentissages dans les systèmes éducatifs français », résumé, Groupe de travail, 9 août.

34. Elsa Chusseau, Charleyne Caroll (2022), « Les conditions de pérennisation et d'essaimage », in Collectif DESIR, *Transformations pédagogique et numérique dans l'enseignement supérieur. Quatre années pour changer les pratiques,* Presses des Mines, Collection Libres opinions, p. 54.

secteur de la EdTech, la collecte et le traitement des données issues des évaluations des jeux pédagogiques nécessitent la collaboration avec des chercheurs opérant dans des établissements publics de recherche. Cette évolution se traduit par une transformation culturelle. En effet, la collaboration entre pédagogues, chercheurs et responsables du monde socio-économique ne se décrète pas. Elle s'organise pour créer les conditions mutuellement bénéfiques. Cela suppose un effort d'ajustement de chaque côté, pour à la fois apprendre à parler le même langage et repenser les modes de fonctionnement, chacun pouvant se faire de l'autre une idée déformée.

4. LA TRANSFORMATION POLITIQUE

Les évaluations positives des jeux pédagogiques, fondées sur la science et les preuves, sont des outils indispensables à la décision politique. Elles donnent aux décideurs une bonne raison – comprendre une raison conforme à l'intérêt général – de mettre en place les réformes pour accompagner, accélérer et pérenniser la pédagogie par le jeu. Pour Stanislas Dehaene, professeur au Collège de France et titulaire de la chaire de psychologie cognitive expérimentale, « il y a urgence à agir »[35]. En effet, à ce jour, la transition pédagogique demeure embryonnaire, et sa trajectoire est encore instable. C'est pourquoi il importe d'apporter les preuves de la valeur éducative des dispositifs ludiques pour que la transformation des méthodes d'apprentissage et de formation prenne une place prioritaire dans les agendas politiques.

CONCLUSION

LES CRITÈRES D'ÉVALUATION D'UN DISPOSITIF PÉDAGOGIQUE LUDIQUE

Afin d'aider les acteurs à apporter les preuves de la justesse des outils pédagogiques innovants, les membres de l'ANRT, sous la présidence de Franck Tarpin-Bernard, directeur général délégué chez Humans Matter, ont identifié différents critères indispensables à l'évaluation d'un dispositif pédagogique ludique[36].

---

35. Stanislas Dehaene (2020), « Il y a urgence à développer le plaisir à l'école », in Sciences humaines, « Bonheur d'apprendre et d'enseigner. Enquête sur ces profs qui changent l'école », Les grands dossiers de Sciences humaines, n° 58, mars-avril-mai 2020, p. 10.

36. ANRT (2022), « Atelier d'intelligence collectif », réalisé par Franck Tarpin-Bernard, séance du groupe de travail « Pédagogie par le jeu », 27 janvier.

1. FICHE PRODUIT

Tout dispositif pédagogique ludique doit d'abord fournir un ensemble minimal d'informations avant d'être évalué. Il s'agit de :

- **la caractérisation de la cible :**
  - âge,
  - niveau prérequis ;
- **l'identification des connaissances apprises et/ou compétences développées ;**
- **le positionnement sur la finalité :**
  - *ouverture* : permet une entrée en matière ludique pour capter l'attention et l'intérêt ;
  - *complémentaire* : permet un renforcement d'une approche traditionnelle ;
  - *autoportant* : remplace une autre approche pédagogique traditionnelle ;
- **la caractérisation des modalités d'utilisation :**
  - individuelle vs collective – taille des groupes ;
  - durée d'une session, voire plusieurs ;
  - intégration dans le parcours avec des phases amont et/ou aval du type débriefing ;
  - matériel nécessaire (exemple, tablette).

2. LES CRITÈRES

Les membres de l'ANRT ont ensuite dressé une liste de facteurs à considérer pour évaluer un jeu pédagogique. Ces facteurs peuvent être regroupés en douze critères, répartis dans quatre familles.

Critères liés à la conception du dispositif

Solidité des bases scientifiques	Est-ce que les contenus, les principes pédagogiques, les mécanismes cognitifs utilisés dans le dispositif s'appuient sur des bases reconnues ?

Qualité du game design	Est-ce que le dispositif a été conçu en exploitant des principes reconnus de game design ? Certaines mécaniques proposées enfreignent-elles des principes fondamentaux ?
Qualité de réalisation	Les éléments matériels et/ou digitaux ont-ils un niveau qualitatif avéré (solidité du matériel, qualité des univers visuels/sonores, pertinence du thème, etc.) pour une réutilisation ?

Critères liés à l'expérience apprenant

Utilisabilité, accessibilité et engagement	L'utilisation est-elle suffisamment intuitive pour les apprenants joueurs, le dispositif est-il accessible au plus grand nombre (RGA), tolérant aux erreurs de manipulation et les limitations sont-elles explicitées, le dispositif génère-t-il un bon niveau d'engagement ?
Qualité du feed-back	Le dispositif explicite-t-il les compétences mises en œuvre (y compris les *soft skills*), permet-il de voir/mesurer des progrès ?
Transférabilité	Le dispositif intègre-t-il des éléments facilitant la transférabilité des apprentissages (débriefing, ancrage dans la réalité, travail métacognitif, etc.) ?

Critères liés à l'expérience formateur

Simplicité d'appropriation	Le dispositif s'accompagne-t-il de moyens appropriés pour son bon usage (guides de prise en main, formations, communautés d'utilisateurs, etc.) ?
Capacité de personnalisation	Le dispositif permet-il d'être adapté au contexte d'utilisation et aux envies du formateur (flexibilité d'usage, paramétrage, insertion de contenus personnalisés, etc.) ?
Coût et simplicité de mise en œuvre	Évaluation des coûts directs (achat matériel, licence d'utilisation, formation) et indirects (temps de prise en main, durée des sessions, mises à jour).

Critères liés à l'usage à long terme

Qualité des éléments de preuve	Nombre et validité (méthodologie, taille des groupes) des études scientifiques ou enquêtes menées sur le dispositif.
Accès aux données d'usage	Fourniture de données exploitables (par le formateur ou des équipes scientifiques tierces) permettant d'évaluer les apports du dispositif.
Niveau de maturité	Informations sur les volumes d'apprenants/formateurs ayant utilisé le dispositif, ancienneté de la mise sur le marché, et historique des mises à jour tenant compte des retours utilisateurs.

Enfin, sur la base de ces critères, il serait utile de concevoir une plateforme numérique, destinée à collecter et partager des évaluations de dispositifs pédagogiques innovants. Un label pourrait être attribué aux outils qui auront démontré leur efficacité en termes de capacité à faire progresser les apprenants. Ce gage de qualité, tout à la fois pédagogique, technique et éthique, garantirait la valeur éducative objectivement attribuable aux ressources évaluées. Il favoriserait leur référencement, améliorant ainsi leur visibilité auprès des utilisateurs tels que les pédagogues, les apprenants et les familles. Cette reconnaissance officielle justifierait notamment l'octroi de financements par les pouvoirs publics.

Bibliographie

Pascal Bressoux (2012). *L'Influence des pratiques enseignantes sur les acquisitions scolaires des élèves,* Paris, La Découverte, « Regards croisés sur l'économie », 2012/2 n° 12.

Axelle Charpentier (2020). « Comment évaluer les innovations pédagogiques », in Sciences humaines, « Bonheur d'apprendre et d'enseigner. Enquête sur ces profs qui changent l'école », *Les grands dossiers de Sciences humaines,* n° 58, mars-avril-mai.

Elsa Chusseau, Charleyne Caroll (2022). « Les conditions de pérennisation et d'essaimage », in Collectif DESIR, *Transformations pédagogique et numérique dans l'enseignement supérieur. Quatre années pour changer les pratiques,* Presses des Mines, Collection Libres opinions.

Pierre-Yves Cusset (2011). « Que disent les recherches sur "l'effet enseignant" ? », *La Note d'analyse,* Centre d'analyse stratégique, n° 232.

Stanislas Dehaene (2020). « Il y a urgence à développer le plaisir à l'école », in Sciences humaines, « Bonheur d'apprendre et d'enseigner. Enquête sur ces profs qui changent l'école », Les grands dossiers de Sciences humaines, n° 58, mars-avril-mai.

Stanislas Dehaene, Elena Pasquinelli, Franck Ramus, Marc Gurgand, Elisabeth Spelke (2021). « La recherche translationnelle en éducation », Rapport, Conseil scientifique de l'éducation nationale, MENJS, Réseau Canopé.

Claudine Golka (2010). « L'évaluation, un levier du changement ? », *Bulletin des bibliothèques de France,* T. 55, n° 4.

Denis Jeffrey (2013). *L'Éthique dans l'évaluation scolaire.* Québec, Presses de l'université Laval.

Anne Jorro, Nathalie Droyer, ss la dir. (2019). *L'Évaluation, levier pour l'enseignement et la formation,* Louvain-la-Neuve, De Boeck Supérieur.

Lucienne Letellier (2014). « Pratiquer une recherche intègre et responsable », Guide, Comité d'éthique du CNRS.

ALLEA (2018). « Code de conduite européen pour l'intégrité en recherche », Édition révisée.

Violette Nemessany (2016). « Regards croisés sur la recherche partenariale. Comment se construisent les relations entre les établissements publics de recherche et les entreprises », *Les Cahiers FutuRIS,* ANRT.

OCDE (2020). « Des critères améliorés pour des évaluations meilleures », février.

Olivier Rey, Annie Freyfant (2014). « Évaluer pour (mieux) faire apprendre », *Dossier de veille de l'Institut français de l'éducation,* n° 94, septembre.

Eric Sanchez, Margarida Romero (2020). *Mythes et réalités : apprendre en jouant,* Paris, Éditions Retz.

Laurent Tardif, Jérôme Clerget, Antoine Lafontaine, Anas Knefati (2021). « Les learning analytics et la transformation pédagogique et numérique : un champ en construction », in Collectif DESIR, *Transformations pédagogique et numérique dans l'enseignement supérieur. Quatre années pour changer les pratiques,* Paris, Presses des Mines.

Développer, expérimenter et évaluer des dispositifs pédagogiques innovants : l'exemple du CréativLab d'Ampiric et la question du passage à l'échelle

Nicolas Mascret, Thierry Berthet, Pascale Brandt-Pomares, Magali Coupaud, Mathilde Favier, Jonathan Mirault, Johannes C. Ziegler

Résumé

Le développement, l'expérimentation et l'évaluation des dispositifs innovants s'inscrivent dans un processus complexe, lorsqu'ils sont réalisés selon une démarche scientifique rigoureuse, et lorsque l'ambition est d'envisager leur passage à l'échelle. En appui sur le CréativLab du pôle pilote Ampiric, les procédures et modalités d'évaluation de dispositifs pédagogiques innovants, issus d'une collaboration entre entreprises, chercheurs et enseignants, seront présentées. Dans un second temps, les enjeux du passage à l'échelle de ces dispositifs seront examinés dans le contexte de l'éducation fondée sur les données probantes ; tout en identifiant les limites liées à la pérennisation et la généralisation de ces dispositifs, ainsi que les limites liées à l'influence des prescriptions chez les enseignants.

INTRODUCTION

L'efficacité d'un nouveau dispositif pédagogique ne va pas de soi. Ce n'est pas parce qu'il est dit « innovant » qu'il est par définition « plus efficace ». L'évaluation et l'expérimentation sont donc cruciales pour connaître la valeur d'un dispositif dans la vie réelle à l'échelle d'une classe, d'une école, d'une académie, ou même d'un pays tout entier.

1. POURQUOI EXPÉRIMENTER ET ÉVALUER DES DISPOSITIFS PÉDAGOGIQUES INNOVANTS ?

Le mot « expérimenter » est souvent utilisé dans un sens assez vague, celui d'essayer, sur un groupe restreint, quelque chose qui sort de l'ordinaire. Cette idée de l'expérimentation dans le quotidien est assez éloignée de l'expérimentation qui consiste à mettre à l'épreuve une hypothèse à l'aide de méthodes susceptibles de contrôler les biais. Expérimenter permet de tester la véracité d'une hypothèse, ou de déterminer l'efficacité d'une méthode pédagogique. Avant chaque expérimentation, les chercheurs, comme les enseignants, doivent veiller à la pertinence de l'intervention proposée, penser à son éventuel passage à l'échelle et à son adoptabilité, aussi bien qu'à son impact potentiel. La conception expérimentale d'une recherche fait partie intégrante des aspects éthiques, car il n'est jamais éthique de mener une recherche dont on sait à l'avance qu'elle ne permettra pas de tirer la moindre conclusion solide. Par exemple, parce qu'elle comprend des biais importants, qu'aucun groupe contrôle n'est prévu, ou qu'elle concerne trop peu d'enfants pour être significative. Quelle serait en effet l'alternative à l'expérimentation ? La prise de décision sur la base de simples intuitions, sans savoir si elles sont fondées ? Certains s'opposent à l'idée même d'expérimentation en éducation pour des raisons faussement éthiques. Mais les enseignants testent tous les jours de nouvelles pratiques sur leurs élèves, sans que personne ne s'y oppose. Malheureusement, ces tentatives se font selon une méthodologie qui ne permet pas de tirer de conclusions rigoureuses, de diffuser les avancées, et ne permet donc pas de faire avancer la connaissance au bénéfice de tous. Le pôle pilote Ampiric (**A**ix-**M**arseille – **P**ôle d'**I**nnovation, de **R**echerche, d'ense**I**gnement pour l'édu**C**ation) est né de la volonté de rapprocher les acteurs du domaine éducatif, qu'ils soient enseignants, scientifiques ou entrepreneurs, dans un seul et même but : améliorer l'éducation des enfants afin de maximiser leur bien-être, leurs connaissances et leurs compétences.

2. LE PÔLE PILOTE AMPIRIC

Financé par le Programme d'investissement d'avenir (PIA3) et s'inscrivant dans la stratégie France 2030, Ampiric est un pôle pilote porté par Aix-Mar-

seille Université en partenariat avec l'académie d'Aix-Marseille, Avignon Université, l'université Côte d'Azur, le CNRS, l'Association nationale de la recherche et de la technologie (ANRT), le Réseau Canopé, et la Région Sud. Le pôle pilote Ampiric (dénommé Ampiric dans la suite de cet article) vise à améliorer durablement l'apprentissage des savoirs fondamentaux des élèves tout au long de leur parcours scolaire, notamment de ceux qui sont en difficulté. Ampiric se fonde sur le développement de travaux de recherche innovants, pour mieux documenter l'évolution de la formation initiale et continue des enseignants, et développer, expérimenter, diffuser des approches pédagogiques innovantes. L'écosystème Ampiric se fonde ainsi sur une meilleure articulation entre la recherche, la formation et le terrain.

Ampiric poursuit cinq objectifs stratégiques :

- transformer la formation et l'accompagnement des enseignants pour faire évoluer leurs pratiques face aux défis liés aux apprentissages fondamentaux ;

- favoriser les coopérations recherche-terrain en soutenant les recherches collaboratives ;

- améliorer la compréhension des processus d'enseignement-apprentissage des savoirs fondamentaux en soutenant des recherches fondamentales, finalisées par les besoins de l'éducation et de la formation;

- coproduire des outils et des applications au service des apprentissages fondamentaux au sein d'un CréativLab ;

- diffuser et valoriser les ressources pédagogiques sur les apprentissages fondamentaux.

3. LE CRÉATIVLAB D'AMPIRIC

Dans le cadre de cet article, nous nous centrerons plus spécifiquement sur le quatrième objectif stratégique afin d'investiguer l'évaluation des dispositifs pédagogiques innovants dans le cadre du CréativLab d'Ampiric. Le CréativLab est un espace de coopération, de recherche et d'action dédié aux enseignants, chercheurs et entreprises. Coordonné par l'ANRT, et à travers elle, par son groupe de travail sur la « Pédagogie par le jeu », il bénéficie du soutien et de l'accompagnement de tous les partenaires d'Ampiric. En appui sur la recherche, l'objectif principal du CréativLab est de coproduire, d'expérimenter et de labelliser des ressources pédagogiques permettant d'améliorer les apprentissages fondamentaux.

Le fonctionnement du CréativLab est articulé autour des actions suivantes :

- **L'initiation et l'accompagnement de projets collaboratifs.** Chaque année, des projets portés conjointement par des responsables d'entreprises, des chercheurs et des enseignants sont sélectionnés, au service de l'amélioration de l'apprentissage des savoirs fondamentaux.

- **La mise en réseau des acteurs, entre terrain, formation, recherche et entrepreneuriat.** Diverses manifestations sont organisées pour opérationnaliser cette mise en réseau.

- **Une évaluation des ressources produites ou perfectionnées.** Les outils pédagogiques développés ont vocation à être évalués auprès de classes partenaires au sein du plateau expérimental « Jacques Ginestié » d'Ampiric, ou directement dans les établissements scolaires.

L'ambition de ce texte est donc double : d'une part, présenter les modalités d'évaluation des projets mises en œuvre dans le CréativLab d'Ampiric, et d'autre part, identifier dans quelle mesure cette évaluation spécifique interroge de façon plus transversale toutes les modalités d'évaluation des dispositifs pédagogiques innovants.

I. L'ÉVALUATION DES PROJETS DANS LE CRÉATIVLAB D'AMPIRIC

L'évaluation des projets au sein du CréativLab répond à un double enjeu : accompagner des projets de recherche partenariale, co-portés par des chercheurs et des acteurs du monde socio-économique en lien avec l'éducation (entreprises, associations, parents d'élèves, enseignants), et évaluer les ressources produites ou développées afin d'envisager un éventuel passage à l'échelle. Deux niveaux d'évaluation des projets sont identifiés : le premier correspond à la procédure de sélection des projets qui intègrent le CréativLab, et le second correspond au suivi et à l'évaluation des projets retenus.

1. LA PROCÉDURE DE SÉLECTION DES PROJETS

Le premier niveau d'évaluation, visant à initier les collaborations, suit une procédure de sélection en deux temps : tout d'abord, les acteurs répondent à un appel à manifestation d'intérêt (AMI) pour soumettre leur projet. Cet appel ne fait pas l'objet d'une sélection particulière, bien que les projets doivent répondre aux objectifs du CréativLab, et plus largement à ceux d'Ampiric, à savoir l'amélioration de l'apprentissage des savoirs fondamen-

taux des élèves. L'ensemble des projets déposés dans le cadre de cet AMI sont présentés lors d'un « speed-dating » organisé annuellement. L'objectif est de permettre aux porteurs de rencontrer les acteurs du monde socio-économique et de la recherche afin de réfléchir ensemble à une collaboration potentielle. Cette collaboration se concrétise par le dépôt d'un projet commun dans le second temps de l'appel. Les projets déposés font l'objet d'une double évaluation conduite par des chercheurs du périmètre d'Ampiric en fonction des disciplines et thématiques concernées. Cette évaluation porte sur les éléments suivants :

- la qualité scientifique du projet ;
- les perspectives de succès de la recherche partenariale ;
- l'implication de l'entreprise et la pertinence du projet dans son contexte entrepreneurial ;
- l'adéquation avec le laboratoire partenaire et sa pertinence dans le projet ;
- la participation des enseignants et la pertinence du projet au service de l'amélioration de l'apprentissage des savoirs fondamentaux.

Au terme de cette évaluation, les rapporteurs doivent émettre un avis sur l'opportunité de recevoir les porteurs de projets pour une audition. Cette dernière permet de clarifier certains aspects liés au projet, qui n'auraient pas été précisés dans le dossier écrit. Elle constitue la dernière phase de ce processus de sélection. Les projets doivent recevoir deux avis favorables pour être reçus en audition, lors de laquelle la commission de sélection émet un avis final sur l'opportunité de soutenir et d'accompagner le projet dans le cadre du CréativLab.

Ce processus de sélection très rigoureux a été élaboré pour s'assurer de la mise en place de conditions favorables au bon déploiement des projets, en termes de déroulement et d'atteinte des objectifs. Financés par de l'argent public dans le cadre du Plan France 2030, les projets soutenus doivent en effet répondre aux exigences fixées par les objectifs du pôle Ampiric relatifs à l'évolution de la professionnalité des enseignants et de l'amélioration de l'apprentissage des élèves. Pour ce faire, la problématique de recherche doit être pertinente et explicitée, et les conditions propices à une collaboration fructueuse entre les chercheurs et leurs partenaires doivent être réunies. Des zones d'ombre sur les modalités de collaboration ou sur l'implication et la motivation d'un partenaire à s'engager dans un projet, ou des incertitudes concernant des questions de méthodologie, de validation des résultats et des impacts de l'outil sur les apprentissages sont autant de facteurs qui ne permettront pas aux projets d'être retenus. À l'inverse, une méthodologie expérimentale rigoureuse et clairement explicitée, contrôlant un maximum de paramètres et en gardant une situation écologique (en salle de classe par exemple), l'articulation avec les recherches actuelles, la plus-value de tel ou

tel laboratoire en termes de compétences et d'approche scientifique et une collaboration active et pertinente entre la recherche et son partenaire sont autant d'éléments qui permettront aux projets de devenir lauréats.

2. LE SUIVI ET L'ÉVALUATION DES PROJETS

Le suivi et l'évaluation des projets retenus forment le deuxième niveau du processus mis en œuvre par le CréativLab. Ainsi, il est demandé aux projets lauréats de compléter un bilan annuel intermédiaire et de présenter leurs avancées lors des « speed-datings » des années suivantes. Ce bilan intermédiaire permet de mettre à jour les acteurs impliqués, de rendre compte des avancées par rapport aux objectifs initiaux, de signaler toute réorientation éventuelle, de présenter les premiers résultats potentiels, et de partager les difficultés rencontrées dans l'organisation de la collaboration, ou tout autre aspect du déploiement du projet. Le suivi des projets met aussi en évidence que les critères d'évaluation peuvent varier au fil des différentes étapes du développement du projet, comme l'illustre le projet de recherche basé sur le jeu Darwinium (encadré 1). De plus, les bilans intermédiaires offrent la possibilité au comité de suivi des projets lauréats de juger du respect des engagements initiaux par les parties prenantes, et d'examiner d'éventuels problèmes. Si les projets ne correspondent pas aux objectifs énoncés, le comité se réserve le droit de cesser son soutien. Cette situation s'est produite à plusieurs reprises. Certains projets ont été interrompus en raison du départ de collaborateurs (chercheurs ou équipes pédagogiques) ou de difficultés techniques/technologiques dans le déploiement des outils proposés.

Ce processus présente des avantages. Il permet tout d'abord de garantir la bonne utilisation et/ou l'optimisation des moyens mis à disposition, qu'ils soient financiers, humains ou matériels. Il assure par ailleurs l'adhésion collective de tous les acteurs et le bon déroulement de la collaboration, en garantissant la possibilité de se retirer d'une collaboration qui ne donnerait pas satisfaction. Cela est particulièrement important dans le développement de dispositifs pédagogiques innovants pour deux raisons : la première concerne le temps extrêmement précieux des équipes pédagogiques de terrain, qui doivent pouvoir trouver une utilité aux dispositifs mis en œuvre en relation avec leur pratique professionnelle. La seconde répond au devoir de vigilance que l'on se doit d'avoir à l'égard des élèves impliqués dans le déploiement des dispositifs, afin que ceux-ci ne soient pas délétères pour leur expérience éducative. Enfin, le suivi rigoureux des projets contribue au développement d'une culture commune au regard de critères de réussite constitutifs des projets soutenus. À mesure du déploiement du second niveau de suivi et d'évaluation, ces critères pourront venir constituer la base d'un référentiel visant à permettre la labellisation de certaines ressources et/ou certains outils dans le cadre du développement d'une filière française spécialisée dans le développement de dispositifs pédagogiques ludiques innovants.

ENCADRÉ 1 — Développement, expérimentation et évaluation du jeu Darwinium

Le projet de recherche présenté ici repose sur la conception du jeu Darwinium intégré à des ressources pédagogiques destinées aux enseignants de collège, afin d'aborder les concepts de hasard et d'évolution. Environné par Ampiric, il s'inscrit dans le cadre du projet ANR EGRESS (Educational Game about Randomness and Evolution for Students in Science). Ce jeu place des élèves de collège dans la situation de chercheurs observant l'évolution de populations chimériques placées dans un dôme expérimental, et devant rendre compte de cette évolution sous forme verbale narrative (explication écrite et orale de l'évolution) et graphique (représentation de la population). Ce projet s'inscrit dans la continuité de travaux qui ont souligné les difficultés d'appréhension de l'évolution du vivant et des idées de hasard qui y sont liées (Fielder et al., 2019). L'objectif de cet encadré est de mettre en évidence la diversité des choix méthodologiques retenus pour guider les étapes de conception du jeu et mesurer son efficacité.

La conception du jeu Darwinium s'inscrit dans un projet plus large de méta-conception du jeu intégré à une intervention didactique en classe (Ardito et al., 2015). Le principe général de ce processus de méta-conception est d'intégrer différents acteurs, comme des experts (didacticiens, spécialistes de l'évolution, ou joueurs experts), des utilisateurs (comme des enseignants, des élèves), dans le processus de conception même. L'activité de conception implique des cycles itératifs de développement avec réalisation de différents prototypes (éléments et mécaniques du jeu différents) et feed-back (Bonnardel et Didier, 2016). À différentes phases de test du jeu, plusieurs données sont collectées en contexte d'activité de classe ou non : verbatims, récits et dessins d'utilisateurs. La méthodologie est basée sur des analyses sémio-cognitives avec une attention particulière accordée aux spécificités d'une activité créative, aux interactions entre différents experts dans le processus de conception, à la place donnée aux utilisateurs finaux (enseignants et élèves) dans le processus de conception, et au retour d'expérience sur l'effet du jeu par rapport aux différents objectifs.

La deuxième partie du projet vise à mesurer l'impact du dispositif conçu sur le raisonnement évolutionniste de collégiens. L'étude en cours sera menée en situation écologique de classe de troisième, lors d'un enseignement des sciences de la vie et de la Terre (SVT), sur un large échantillon. Les mesures seront réalisées selon une méthode de pré-test / post-test, à partir d'un questionnaire fermé (Coupaud et al., 2019). Au sein d'une vingtaine de classes, un groupe contrôle et un groupe test seront constitués. La méthodologie est basée sur des analyses statistiques multivariées permettant une approche comparative des résultats.

Les procédures de sélection des projets dans le CréativLab et de suivi des projets retenus sont des illustrations, au sein d'Ampiric, du processus de développement, d'expérimentation et d'évaluation de dispositifs pédagogiques innovants. Toutefois, bien que spécifiques, ces procédures interrogent au-delà du CréativLab, et interpellent la question du passage à

l'échelle, transversale à toute évaluation de dispositifs pédagogiques innovants. En effet, la validation des outils ainsi développés doit faire l'objet de différents niveaux de preuve afin de permettre leur passage à une échelle plus large, et à terme aboutir à un transfert éventuel vers le marché. Au-delà du CréativLab d'Ampiric, les enjeux et les limites du passage à l'échelle des dispositifs pédagogiques innovants en contexte d'éducation se doivent alors d'être examinés de façon transversale.

II. LE PASSAGE À L'ÉCHELLE : ENJEUX ET LIMITES

1. L'ÉDUCATION FONDÉE SUR LES DONNÉES PROBANTES

Se développant au début du XX^{e} siècle, les bonnes pratiques scientifiques basées sur les niveaux de preuves voient le jour petit à petit dans le monde de la recherche. De nos jours, c'est un ensemble de pratiques qui tend à se systématiser pour optimiser l'impact de la recherche dans nos sociétés. Le rapport de la Haute Autorité de santé (HAS) publié en 2013 a défini le niveau de preuve d'une étude comme étant la capacité de celle-ci à répondre à une question précise, en appui sur les quatre critères suivants : 1) adéquation du protocole à la question, 2) existence de biais dans la réalisation, 3) adaptation de l'analyse statistique aux objectifs et 4) puissance statistique et taille de l'échantillon.

L'ensemble de ces bonnes pratiques basées sur les niveaux de preuves (*evidence-based practices* en anglais) est conceptualisé de manière globale, mais elles se spécifient en fonction du domaine de prédilection de l'étude en question. L'approche « éducation fondée sur des données probantes » ou *evidence-based education* apparaît dans les années 90 (Barrett, 1991). Elle réside dans des pratiques et des politiques éducatives préconisant le fondement des études sur les données les plus probantes produites par la recherche (Saussez et Lessard, 2009) afin de limiter l'influence des modes, des croyances et des idéologies (Davies, 1999 ; Slavin, 2002 ; 2008) dans le monde de l'éducation. Cette démarche, qui se veut un outil d'aide à la décision scientifique et rationnelle pour les praticiens et les politiques, est fondée sur quatre étapes : 1) formuler une question précise, 2) effectuer une revue systématique de recherche, 3) évaluer la qualité des preuves recueillies en se basant sur un système de production de la preuve, l'essai contrôlé randomisé et 4) formuler une réponse claire à la question posée, établir si la solution fonctionne et rendre ces preuves disponibles (Davies, 1999 ; Gough, 2007 ; Oakley, 2004).

L'ambition du passage à l'échelle s'inscrit donc dans un mouvement plus large qui accorde un rôle important à l'utilisation des méthodes scientifiques

pour faire progresser les pratiques pédagogiques. Ce mouvement est par exemple mis en avant en France par le Conseil scientifique de l'éducation nationale (CSEN), ou en Angleterre par l'*Education Endowment Foundation* (encadré 2).

ENCADRÉ 2 — Education Endowment Foundation (EEF)

Le secrétaire d'État britannique à l'éducation, Michael Gove, a annoncé fin 2010 son intention de créer un programme d'aide pour améliorer les conditions dans les écoles difficiles en Angleterre, et c'est en 2011 que l'EEF est fondée par une association caritative, The Sutton Trust, en partenariat avec Impetus. Le programme EEF est alors doté d'une subvention du ministère de l'Éducation de 125 millions de livres sterling. Il emboîte le pas du mouvement « What Works » aux États-Unis et au Royaume-Uni (White, 2019).

Ce fonds lance chaque année de nombreuses expérimentations en éducation, ayant pour but l'identification des pratiques et des aménagements efficaces, et répondant le mieux possible aux objectifs définis par les acteurs de l'éducation. Avant chaque expérimentation, l'EEF réalise d'abord une revue de la littérature scientifique, afin de savoir si une intervention est justifiée par des connaissances fondamentales ou empiriques. Puis, l'intervention est testée en situation réelle, en utilisant principalement des paradigmes en essais contrôlés randomisés (ECR) avec randomisation des groupes témoin et expérimentaux. L'expérimentation est caractérisée par plusieurs étapes :

1. Dans un premier temps, les enseignants et les chercheurs coconstruisent une intervention qui est à la fois faisable et souhaitable, du point de vue des pratiques éducatives des enseignants et de celui de la recherche.

2. Dans un deuxième temps, les enseignants et les chercheurs mènent une étude pilote sur un petit échantillon, afin de savoir si les conditions sont réunies pour que l'intervention puisse faire l'objet d'un ECR.

3. Ensuite, une première expérimentation est conduite sur un échantillon étendu de classes, dans des conditions idéales puisque l'intervention est réalisée par les chercheurs et les enseignants qui ont élaboré cette action.

4. Si cette première expérimentation produit des résultats positifs, l'intervention est alors testée à grande échelle par des enseignants et des expérimentateurs totalement indépendants de l'équipe initiale, après une formation du personnel censé la mettre en place dans la réalité.

Lorsque les résultats sont positifs, afin de multiplier les chances d'adoption des bonnes pratiques, l'EEF produit des boîtes à outils d'indications pratiques pour les enseignants, propose des formations, un accompagnement, et la création de réseaux pour la dissémination des résultats et l'échange des pratiques. L'EEF a démontré qu'il est aujourd'hui possible de mener de nombreuses expérimentations en éducation, en utilisant principalement des paradigmes en ECR.

Il est à noter que les premières tentatives de l'EEF ont souffert d'un taux d'acceptabilité relativement bas, avec une moyenne de 24 % d'élèves qui ont abandonné entre 2011 et 2012. Toutefois, depuis, l'EEF a introduit de nouvelles stratégies pour recruter et retenir les écoles. Par exemple, avec un développement des communications sur les avantages des ECR dans les écoles, par le biais d'événements de recrutement ; et par la diffusion de documents expliquant clairement la conception et l'évaluation des études, mais aussi le rôle central des écoles (Dawson et al., 2017).

En termes d'évaluation, par le biais d'un appel d'offres, l'EEF a nommé un panel d'organismes de recherche spécialisés dans l'évaluation de l'éducation, qui se mettraient en concurrence pour évaluer les programmes et seraient ensuite associés aux développeurs. Un point d'honneur a été mis en avant : publier toutes les conclusions, quel que soit le résultat (Nevill, 2016). Certains des premiers essais de l'EEF ont souffert de l'absence d'évaluation de qualité. Ce qui signifie qu'ils n'étaient pas en mesure d'expliquer les processus causaux qui sous-tendaient les résultats, ou de décrire leur mise en œuvre (Morris et al., 2016). Début 2012, l'EEF a publié des directives strictes sur la sélection des tests dans ses évaluations, les principaux critères étant qu'ils devaient avoir une large validité externe, et être fortement corrélés aux performances dans les évaluations nationales à fort enjeu (The EEF, 2012). Grâce aux efforts réalisés, ce programme a engendré un grand nombre de preuves qui aident à identifier ce qui fonctionne, et ce qui ne fonctionne pas, dans les écoles anglaises (pour une revue plus complète, voir Edovald et Nevill, 2021).

Le passage à l'échelle, comme envisagé par les méthodologies qui se basent sur une éducation fondée sur des données probantes, a pour ambition de s'appuyer sur une collaboration inédite entre les enseignants et les chercheurs, dans le but de faire émerger une véritable « recherche translationnelle » en éducation (Dehaene et Pasquinelli, 2021). Ce terme trouve son origine en médecine. Son objectif premier est de proposer et d'évaluer des solutions aux problèmes concrets que pose la santé, afin que ces solutions soient fondées sur les meilleures preuves disponibles. Cette approche s'applique également au domaine de l'éducation, qui constitue un enjeu sociétal aussi important que la santé. Tout comme la recherche médicale, la recherche en éducation doit aider à identifier des pratiques, des aménagements et des interventions efficaces, qui répondent le mieux possible aux objectifs définis par les acteurs de l'éducation. La recherche translationnelle doit aider à évaluer ce qui marche dans la vie réelle, à l'échelle d'une classe, d'une école, d'une académie, ou même d'un pays tout entier. Le terme fait donc référence à l'ensemble des étapes qui permettent de passer d'une connaissance scientifique fondamentale à sa mise en application pratique. Ce processus est progressif : du laboratoire à la salle de classe, en passant par une série de recherches sur des cohortes de plus en plus grandes d'élèves, jusqu'à la mise en pratique à grande échelle.

Comme le soulignent Stanislas Dehaene et Elena Pasquinelli (2021) dans la synthèse du CSEN, « en éducation comme en médecine, il n'est raisonnable de tester une nouvelle intervention que lorsque des recherches préalables indiquent qu'elle a de bonnes chances d'être plus efficace que les meilleures pratiques éducatives actuellement disponibles (présomption d'efficacité). La pré-identification des interventions potentiellement efficaces est indispensable pour éviter les coûts inutiles, aussi bien en termes de temps et d'argent que d'effets indésirables ». Le speed-dating d'Ampiric est un bon exemple de moment où les idées des partenaires économiques se confrontent aux regards des scientifiques.

Avant le passage à l'échelle d'un dispositif pédagogique innovant, il existe un continuum de méthodes, de plus en plus sophistiquées, fournissant une hiérarchie de niveaux de preuves (encadré 3).

ENCADRÉ 3 — La pyramide des preuves

L'approche basée sur les preuves a produit une liste de méthodes de recherche permettant d'apporter des preuves en faveur ou à l'encontre d'une intervention, ordonnées selon leur susceptibilité à des biais de différente nature. Dans son rapport sur la « recherche translationnelle », le CSEN (Dehaene et Pasquinelli, 2021) fait la proposition suivante :

Preuves de niveau 0 : Témoignages, questionnaires, interactions entre enseignants, parents, apprenants, chercheurs.

Ce riche tissu d'échanges permet de définir les buts des interventions de façon collégiale, mais aussi de recueillir la sagesse collective des personnes impliquées dans le processus éducatif. Même si les résultats de ces échanges ne fournissent pas une preuve de « ce qui marche », ils sont très importants à prendre en compte. En effet, ils sont la principale source d'hypothèses sur « ce qui pourrait marcher ».

Preuves de niveau 1 : Connaissances d'ordre fondamental obtenues en laboratoire.

Par ordre croissant de fiabilité pour la recherche translationnelle, on distingue :

- les connaissances issues de recherches de laboratoire éloignées des conditions réelles, ou avec des échantillons limités ;

- les connaissances issues de recherches de laboratoire avec conditions proches du réel, ou réalisées sur de nombreux sujets ;

- les résultats convergents de différentes lignes de recherche, avec réplication et confirmation des résultats, sur de vastes échantillons.

Preuves de niveau 2 : Études observationnelles dans les classes.

Sans modification des conditions initiales, ces études examinent la corrélation entre les pratiques et les progrès des élèves. La qualité de ces études varie selon l'utilisation de supports objectifs pour l'observation (grilles, questionnaires, tests), et d'outils statistiques pour contrôler les autres variables qui pourraient expliquer les résultats.

Preuves de niveau 3 : Études expérimentales avec modification contrôlée des conditions d'apprentissage et observation des conséquences.

Par ordre de fiabilité, on distingue :

– les études qui comparent un pré-test et un post-test (un design expérimental souvent très insuffisant, car tous les élèves s'améliorent avec le temps) ;

– les études qui comparent un groupe d'intervention et un groupe témoin en pré- et post-test, mais sans les randomiser ;

– les études randomisées contrôlées, dans lesquelles les élèves ou les classes sont assignés aléatoirement à la condition d'intervention ou de contrôle, et sont comparés en pré- et post-tests.

Preuves de niveau 4 : Synthèses et méta-analyses.

Résultats répliqués et confirmés par plusieurs groupes de recherche indépendants, idéalement réunis dans des analyses systématiques de la littérature (méta-analyses).

Il est important de souligner que toutes ces méthodes différentes ont un rôle à jouer dans la recherche en éducation. Elles ne sont pas en concurrence, mais créent des synergies, et soutiennent collectivement la découverte de moyens pour améliorer l'éducation à tous les niveaux. Il serait contre-productif d'affirmer que seuls une méthode ou un petit nombre de méthodes ont de la valeur dans le domaine de l'éducation fondée sur les preuves. Par exemple, il est souvent utile de commencer par des études observationnelles, avant de mener des études plus définitives (par études randomisées contrôlées) quant à l'efficacité causale des interventions éducatives.

Très souvent, les idées viennent d'abord des préoccupations du terrain, par exemple d'un enseignant qui souhaite tester une nouvelle idée avec ses élèves. Il est toutefois important d'être conscient que ces observations informelles ne fourniront pas de preuve. Si ces observations informelles suggèrent que l'idée est prometteuse, il est important de mener une véritable expérimentation – il ne s'agit pas encore d'un véritable essai randomisé contrôlé à grande échelle, mais d'une expérience à petite échelle (quelques classes), avec des mesures objectives des compétences cibles avant et

après intervention, et autant que possible un groupe contrôle (voir la méthodologie proposée par EEF). Si cette expérimentation produit des effets espérés, il peut être judicieux de chercher à obtenir le niveau de preuve le plus élevé, en menant un essai randomisé contrôlé à une échelle qui dépasse celle d'un seul établissement. Pour autant, le passage à très grande échelle, à savoir le déploiement d'une nouvelle pratique pédagogique auprès de nombreux enseignants, pose d'autres problèmes.

2. LE PASSAGE À L'ÉCHELLE : LA QUESTION DE LA PÉRENNISATION ET DE LA GÉNÉRALISATION

L'évaluation d'une expérimentation pose de redoutables questions théoriques et pratiques, notamment quand il s'agit de décider de l'opportunité de pérenniser et éventuellement généraliser le dispositif expérimenté. Le développement de logiques d'expérimentation de nouveaux dispositifs, à petite échelle, en vue de leur possible généralisation, tend à s'imposer comme une nouvelle technique de *policy making*. L'évaluation de ces innovations expérimentales a donc une responsabilité particulière, puisqu'il s'agit au fond de décider *ex-ante* du bien-fondé d'un dispositif ; et elle soulève des questions inhabituelles en termes évaluatifs, dont la plus prégnante est sans nul doute celle de sa pérennisation/généralisation. Cette évaluation suppose donc de prendre en compte un certain nombre de points de vigilance.

Premier de ces points, les expérimentations s'inscrivent le plus souvent dans l'écosystème local qui les porte ; et elles visent à répondre à des problématiques elles aussi localisées, dont il est souvent difficile de mesurer la validité externe pour d'autres écosystèmes. Il est ainsi plutôt rare que les expérimentations soient construites en intégrant dans leur design la perspective de leur généralisation. Il importe donc, dans l'évaluation, de bien mesurer le degré d'endogénéité de l'expérimentation en question, pour analyser sa transférabilité. C'est là souvent un point critique des évaluations randomisées, puisqu'il s'agit, dans le protocole expérimental d'évaluation, de neutraliser tous les biais endogènes possibles lors de la constitution des échantillons (Conseil d'État, 2019).

Second point critique, la représentativité de l'échantillon – ou des échantillons test et témoin en cas d'évaluation randomisée –, qui supposerait que celui-ci soit représentatif de la population cible d'un processus de généralisation. Évaluer la possibilité de généraliser les expérimentations conduites dans le cadre d'un pôle pilote, comme Ampiric, supposerait d'avoir la capacité de garantir que les populations d'élèves et d'enseignants soient représentatives de la population cible régionale ou nationale visée.

La conscience, pour les parties prenantes, de participer à une expérimentation peut également porter en soi un biais ; cela nécessite de s'interroger sur l'effet de la disparition de ce phénomène potentiel en phase de généralisation. Le caractère volontaire généralement en œuvre pour la participation

à une expérimentation dans le cadre des projets Ampiric suppose de s'interroger sur l'impact de la disparition de ce caractère en phase de généralisation. Il n'est ainsi pas dit qu'une expérimentation produisant un résultat positif auprès d'enseignants volontaires produise les mêmes effets dans le cadre d'une politique académique. Ce que rappelle également judicieusement Gordon Stobart à propos des conditions de généralisation, à toutes les écoles britanniques, d'une expérimentation d'évaluation par les apprentissages d'abord localisée, lorsqu'il écrit : « la manière dont les politiques centralisées sont parfois mises en œuvre empêche que la généralisation de l'expérimentation atteigne ceux pour lesquels le dispositif est pourtant une clé essentielle de leur réussite, à savoir les enseignants de terrain » (2011, p. 47).

Comme le rappellent Bérard et Valdenaire (2013, p. 5), « le changement d'échelle peut induire des effets dits d'équilibre général qui n'affectent pas un dispositif expérimental ». Dit autrement, rien ne garantit que la généralisation d'un dispositif expérimental, par la montée en généralité qu'elle produit, se traduise par la réplication à l'identique des effets constatés sur une population réduite. Ce qui rejoint la question de la représentativité.

Si l'évaluation d'une expérimentation permet de contrôler des effets de court terme – ceux mesurables dans le temps de la phase expérimentation/évaluation –, on peut s'interroger sur l'émergence d'effets de moyen terme, car on fait comme si « la généralisation se déroulerait dans les conditions de l'expérimentation sans évolution du contexte, ni des normes, et représentations sociales » (Beaupère, Bosse et Lemistre, 2014). Ces évolutions échappent de fait à une observation temporellement située. Ce questionnement est particulièrement prégnant au regard de la préconisation de pérenniser, ou pas, le dispositif au moment de son évaluation.

Pour finir, on peut rajouter que l'évaluation d'un dispositif tel qu'Ampiric, et en son sein le CréativLab, parmi d'autres actions et projets, engendre un bouquet de difficultés supplémentaires liées à sa complexité. En effet, le caractère emboîté et multiniveau des projets et des expérimentations dans l'expérimentation suppose d'adopter une démarche évaluative elle aussi multiniveau, touchant à des domaines éducatifs divers, des populations scolarisées hétérogènes (élèves, étudiants, enseignants), tant en formation initiale que continue, des territoires différenciés et des problématiques nombreuses. En évaluer la possibilité de pérennisation, et plus encore de généralisation, représente un challenge évaluatif de premier ordre.

Si la question de la pérennisation et de la généralisation des dispositifs évalués se pose pour les raisons évoquées plus tôt, elle s'accompagne de celle des limites des prescriptions lors de la confrontation au terrain.

3. LE PASSAGE À L'ÉCHELLE : LES LIMITES DES PRESCRIPTIONS

L'innovation ne se décrète pas. Pas plus en pédagogie qu'ailleurs. Pour advenir, elle suppose de faire différemment dans un but d'amélioration des méthodes d'enseignement, des ressources pédagogiques, des apprentissages… La question de la généralisation de l'innovation dans la classe, du changement là où les résultats de la recherche les plus probants doivent arriver, concerne *in fine* celle ou celui qui en a la responsabilité : l'enseignante ou l'enseignant. Or, il peut s'avérer qu'arrivée en bout de chaîne, la verticalité de prescriptions trouve parfois des obstacles dans la confrontation au réel. En France, les programmes d'enseignement et leurs documents d'accompagnement ne manquent pas de détails. Pour autant, certains aspects du programme peuvent être plus ou moins traités ; tout comme certaines pratiques, qu'ils ne préconisent pas particulièrement, peuvent largement se répandre. Cette distinction entre prescrit et réel s'inscrit dans l'analyse de l'activité enseignante, qui intéresse particulièrement la formation professionnelle des enseignants. À un niveau de master universitaire, celle-ci s'inscrit en faux avec la diffusion de méthodes pédagogiques, voire de recettes, peu étayées sur le plan scientifique, qui renverraient l'exercice de la profession d'enseignant à un métier d'exécution, loin d'être à la hauteur des enjeux de société en matière d'éducation. L'adossement de cette formation aux savoirs produits par la recherche est donc un enjeu majeur du travail de l'enseignant, pour lequel les processus d'enseignement-apprentissage occupent une place centrale. L'ancrage relatif aux savoirs à transmettre et à leur didactique disciplinaire conditionne l'organisation de l'enseignement, et induit des processus de transmission-appropriation de savoirs. La discipline joue un rôle majeur qui situe l'enseignement dans un champ épistémologique et dans un format scolaire précis (programmes, niveaux, enseignants…).

Toutefois, bien que lié au travail et à la professionnalité enseignante, cet ancrage disciplinaire ne suffit pas totalement à caractériser le travail de l'enseignant, ni les processus d'enseignement-apprentissage pour lesquels l'activité de l'enseignant, à l'origine de celle de l'élève, est déterminante. Les ressources, les matériels et les dispositifs pédagogiques sont des médiateurs instruments (Vérillon et Rabardel, 1995 ; Engeström et Sannino, 2010) de l'activité du professeur, activité dont l'efficacité se mesure au fait que les élèves ont appris. En d'autres termes, c'est ce que l'enseignant demande de réaliser qui est porteur de l'activité de l'élève. L'élaboration et le choix des tâches sont donc capitaux (Brandt-Pomares, 2008). Aussi, réunir les conditions pour que les enseignants développent leur activité pour favoriser l'apprentissage des élèves interroge la question de l'articulation entre le monde de la recherche en éducation et celui des pratiques enseignantes. Elle bute aujourd'hui sur la difficulté à répartir le travail d'ingénierie, de coconstruction entre chercheurs et enseignants. Le CréativLab apporte une solution en amont de ce qui peut devenir une innovation pédagogique efficace, par un travail d'articulation entre le monde de la R&D et celui de son application.

Sans exclure que, pour leur généralisation, la mise en œuvre des ingénieries développées dans le cadre du CréativLab fasse l'objet d'analyse en contexte écologique, le processus proposé par le CréativLab tient compte de la professionnalité des enseignants et de la complexité de leur tâche. Ce type d'élaboration qui intègre cette dimension professionnelle dès la conception des dispositifs favorise le processus d'appropriation des innovations par les enseignants, et leur pilotage de l'enseignement par les apprentissages.

CONCLUSION

Développer, expérimenter et évaluer des dispositifs pédagogiques innovants est un processus long et complexe. Le CréativLab d'Ampiric en est un exemple particulièrement illustratif, et des précautions d'évaluation sont prises pour donner aux projets retenus les meilleures chances de succès, en appui sur une collaboration étroite entre entreprises, chercheurs et enseignants. Mais le développement, l'expérimentation et l'évaluation des projets dans le cadre du CréativLab posent aussi la question du passage à l'échelle, qui est transversale à tous les dispositifs innovants. Ce passage à l'échelle est un enjeu, notamment dans le contexte de l'éducation, mais certaines précautions doivent être prises, tout comme les limites doivent être connues. Lors de l'évaluation de dispositifs innovants, il s'agit d'être vigilant pour placer le curseur de l'évaluation au bon niveau, permettant de trouver un équilibre entre rigueur de la méthodologie scientifique, échelle envisagée et finalité du dispositif.

Bibliographie

C. Ardito, P. Buono, M. F. Costabile, R. Lanzilotti, A. Piccinno, L. Zhu (2015). « On the transferability of a meta-design model supporting end-user development », in *Universal Access in the Information Society*, 14(2), p.169-186.

B. H. Barrett, R. Beck, C. Binder, D. A Cook, S. Engelmann, R. D. Greer, C. L. Watkins (1991). The right to effective education, in *The Behavior Analyst,* 14, p.79.

N. Beaupère, N. Bosse & P. Lemistre (2014). « Expérimenter pour généraliser le portefeuille d'expériences et de compétences à l'université : le sens de l'évaluation », in *Formation Emploi,* 126, p. 99-117.

J. Berard & M. Valdenaire (2013). « L'expérimentation pour renouveler les politiques publiques ? », in laviedesidées.fr.

N. Bonnardel & J. Didier (2016). « Enhancing creativity in the educational design context:

An exploration of the effects of design project-oriented methods on students' evocation processes and creative output », in *Journal of Cognitive Education and Psychology*, 15(1), p.80-101.

P. Brandt-Pomares (2008). Searching for information on the internet about the link between task and activity. In J. Ginestié (Ed.), *The cultural transmission of artefacts, skills and knowledge: Eleven studies in technology education,* Rotterdam: Sense Publishers, p.173-192.

Conseil d'État (2019). « Les expérimentations : comment innover dans la conduite des politiques publiques ? », Collection « Les études du Conseil d'État ».

M. Coupaud, J. Castéra, H. Armand, P. Brandt-Pomares & A. Delserieys (2019). « Validité et fiabilité d'un questionnaire visant à situer les conceptions de l'évolution du vivant de collégiens », in *Recherche en didactique des sciences et des technologies,* 20, p.27-59.

P. Davies (1999). « What is evidence-based education? », in *British Journal of Educational Studies,* 47, p. 108-121.

P. Davies (2004). Systematic reviews and the Campbell collaboration. *Evidence-based practice in education,* Corwall: Open university Press.

A. Dawson, E. Yeomans & E. R. Brown (2018). « Methodological challenges in education RCTs: reflections from England's Education Endowment Foundation », in *Educational Research, 60*(3), p.292-310.

S. Dehaene, E. Pasquinelli (2021). *La recherche translationnelle en éducation. Synthèse de la recherche et recommandations.* Conseil scientifique de l'éducation nationale, MENJS, Réseau Canopé.

T. Edovald & C. Nevill (2021). « Working out what works: The case of the Education Endowment Foundation in England », in *ECNU Review of Education,* 4(1), p.46-64.

Y. Engeström & A. Sannino (2010). « Studies of expansive learning: Foundations, findings and future challenges », in *Educational Research Review,* 5(1), 1-24.

D. Fiedler, G. Sbeglia, R. Nehm & U. Harms (2019). « How strongly does statistical reasoning influence knowledge and acceptance of evolution? », in *Journal of Research in Science Teaching,* p.1-24.

D. Gough (2004). *Systematic research synthesis. Evidence-based practice in education,* Corwall: Open university Press.

D. Gough (2007). « Weight of evidence: a framework for the appraisal of the quality and relevance of evidence », in *Research Papers in Education,* 22, p. 213-228.

Haute Autorité de santé (2013). *Niveau de preuve et gradation des recommandations de bonne pratique. État des lieux,* Paris: HAS.

S. Morris, T. Edovald, C. Lloyd & Z. Kiss (2016). « The importance of specifying and studying causal mechanisms in school-based randomised controlled trials: Lessons from two studies of cross-age peer tutoring », in *Education Research and Evaluation,* 22(7-8), p.422–439.

C. Nevill (2016). EEF blog : Do EEF trials meet the new gold standard? [Blog post]. https://educationendowmentfoundation.org.uk.

A. Oakley (2004). *Who's afraid of the randomized controlled trial? Social research methods: a reader,* Londres: Routledge, p. 261-268.

F. Saussez & C. Lessard (2009). « Entre orthodoxie et pluralisme, les enjeux de l'éducation basée sur la preuve », in *Revue française de pédagogie. Recherches en éducation,* 168, p.111-136.

R. Slavin (2002). « Evidence-based educational policies: transforming educational practice and research », in *Educational Researcher,* 31, pp.15-21.

R. Slavin (2008). « What works? Issues in synthesizing educational programs evaluations », in *Educational Researcher,* 37, p.5-15.

G. Stobart (2011). « L'évaluation par les apprentissages : d'une expérimentation locale à une politique nationale », in *Revue française de pédagogie,* 174, p.41-48.

The Education Endowment Foundation (2012). *EEF guidance on choosing and delivering attainment tests.*

P. Vérillon & P. Rabardel (1995). « Cognition and artifacts - a contribution to the study of though in relation to instrumented activity », in *European journal of psychology of education,* 10(1), p.77-101.

PARTIE II

Les outils

Les neurosciences cognitives entrent à l'école : évaluer la collaboration de recherche ?

Pascale Toscani

Résumé

Au début du 21e siècle, l'école a connu une rupture épistémologique de taille, par l'arrivée des neurosciences cognitives, qui n'avaient pas franchi jusque-là le seuil de la classe. Pour intégrer les neurosciences cognitives dans leurs pratiques, les enseignants ont adopté des choix professionnels différents. Certains ont cherché des outils, des techniques et des modèles types prêts à l'emploi pour les implémenter dans leurs classes. D'autres ont préféré bénéficier de l'accompagnement de laboratoires de neurosciences pour appliquer, des recherches tout aussi « clés-en-mains » au sein de protocoles. Pour autant, cela pose un problème dans les deux cas : les enseignants ne connaissent pas les neurosciences, et les chercheurs ne connaissent pas la classe. Dès lors, comment évaluer des dispositifs dont on est censé mesurer l'efficacité dans la classe, à travers des critères et indicateurs stables ? Il s'agira, dans cet article, de faire un état des lieux des dispositifs d'intégration des neurosciences cognitives à l'école, et de poser la problématique de l'évaluation complexe de ces dispositifs qui impliquent de puissants paradigmes d'apprentissage.

INTRODUCTION

En cherchant à comprendre le fonctionnement du cerveau, grâce aux progrès de l'imagerie cérébrale, les neuroscientifiques ont été conduits à s'intéresser de près aux mécanismes de l'apprentissage. Les applications les plus spectaculaires de ces recherches se sont tournées vers les pathologies neurologiques et vers la création de toutes les formes d'intelligence artificielle. Dans un premier temps, l'école s'est tenue à distance de ce qui lui apparaissait alors comme relevant du monde médical et technologique. Dans un deuxième temps, lorsque les recherches se sont intéressées au développement de l'enfant, aux méthodes de lecture ou à l'apprentissage des mathématiques, les enseignants se sont quelque peu méfiés de la mainmise des chercheurs sur leur terrain professionnel, l'apprentissage.

Mais au fur et à mesure que les études internationales faisaient douter de la performance de notre système scolaire français, pointaient les inégalités récurrentes ou le faible niveau en mathématiques, les enseignants se sont davantage intéressés à l'ensemble des recherches en neurosciences cognitives. Aujourd'hui, beaucoup attendent effectivement que les neurosciences apportent des réponses aux questions que leur posent les situations d'apprentissage. L'idée qu'une approche scientifique de l'apprentissage pourrait interroger des pratiques anciennes, bousculées par l'attitude actuelle des élèves, a fait son chemin.

Les neurosciences sont alors entrées dans la classe, sous des formes diversifiées. Tantôt de façon individuelle, tantôt de façon collective. Certains pédagogues ont fait le choix, de façon « autonome », d'utiliser des outils trouvés principalement sur le Net, ou dans des ouvrages de vulgarisation scientifique. Il s'agit d'une recherche d'efficacité la plus rapide possible, qui implique un « prêt-à-porter » pédagogique sans en connaître forcément les théories sous-jacentes. D'autres ont rencontré les neurosciences lors de journées pédagogiques « one shot ». Moments intéressants comme moyen d'information, mais qui ne laissent aucunement le temps de construire une réflexion sur le sujet.

Lorsque les enseignants mettent en pratique les connaissances acquises, et déploient des dispositifs appelés « expérimentations » dans leur enseignement, l'évaluation de leurs résultats auprès des élèves n'est quasiment jamais réalisée, il est donc impossible d'en mesurer l'efficacité. Pourtant, l'évaluation de ces dispositifs expérimentaux, tout comme celle de la recherche, devrait répondre à des critères précis tels que la crédibilité, la transférabilité, la fiabilité, et la reproductibilité. Pour le dire autrement, d'après les termes de Stanislas Dehaene et Elena Pasquinelli, il est nécessaire qu'une expérimentation soit basée sur la preuve. Ces deux auteurs décrivent, en référence à la recherche médicale, les niveaux possibles d'une « pyramide des preuves » pour les expérimentations réalisées à l'école :

- **« Niveau 0** : témoignages, questionnaires, interactions entre enseignants, parents, apprenants, chercheurs. Ce niveau est source d'hypothèse.
- **Niveau 1** : connaissances d'ordre fondamental obtenues en laboratoires.
- **Niveau 2** : études observationnelles dans les classes. Il s'agit de partir d'un phénomène dysfonctionnant afin de repérer les causes et les améliorations efficaces possibles.
- **Niveau 3** : études expérimentales avec modification contrôlée des conditions d'apprentissage et observation des conséquences.
- **Niveau 4** : synthèse et méta-analyse »[37].

Les enseignants ne sont pas formés à s'engager dans des recherches co-construites avec les laboratoires, qui positionneraient leur recherche aux niveaux 3 et 4 de la pyramide. Il semble important de commencer par accompagner les enseignants à intégrer dans leurs pratiques les niveaux 0, 1 et 2.

Dans le système éducatif, la progression des élèves est très dépendante de l'évolution des paradigmes professionnels des enseignants, ainsi que de leurs biais cognitifs sur l'apprentissage et son évaluation. L'évaluation sommative, dominante et surtout déterminante dans le monde scolaire, repose sur une norme, qui n'est pas applicable à tous les élèves au même âge, sur les mêmes temps scolaires, selon les mêmes méthodes. Entre liberté pédagogique, obligation de réalisation des programmes, volonté de changement, il y a peu de place pour l'évaluation des dispositifs d'intégration des neurosciences dans la classe et de leurs impacts, alors qu'elles constituent un enjeu majeur pour l'école de demain.

I. LA COLLABORATION DE RECHERCHE ENTRE L'ÉCOLE ET LE LABORATOIRE

1. UNE VOIE DE DÉVELOPPEMENT PROFESSIONNEL POUR TOUS LES ACTEURS IMPLIQUÉS DANS L'ÉDUCATION.

Dans une tribune du journal *Le Monde*, Eirick Prairat proposait de cerner la notion de liberté pédagogique dans sa dimension collective :

37. S. Dehaene, E. Pasquinelli (2021). « La recherche translationnelle en éducation », in *L'École éclairée par la science,* Odile Jacob, p. 44.

> « Un professeur est libre de ses options didactiques et de ses choix pédagogiques, sous réserve qu'ils correspondent à l'état actuel des connaissances et au niveau intellectuel des élèves dont il a la charge. Ce n'est pas une liberté restreinte mais raisonnée, car il n'y a de liberté qu'éclairée par la connaissance et le sens de la responsabilité »[38].

Cette définition fait référence à ce qu'il appelle « l'état actuel des connaissances ». Ce concept ne saurait se limiter à celui des sciences de l'éducation. Se profilent en arrière-plan la formation initiale et continue des enseignants en tant que moyen, mais surtout leur développement professionnel en tant qu'objectif. Il ne s'agit pas seulement d'imposer des connaissances de l'extérieur, mais plutôt de les sélectionner, de les étudier, de les adapter aux contextes, et de les faire expérimenter par les enseignants eux-mêmes, en collaboration avec ceux qui les ont fait émerger et les ont diffusées.

Au-delà de cette déclaration d'intention, chacun sait bien qu'il ne suffit pas de distribuer les partitions pour qu'un orchestre soit en mesure de donner un concert. La collaboration entre un laboratoire de recherche et une école suppose de définir des étapes qui prennent en compte les résistances, les habitudes et les routines de chacun des partenaires.

Ainsi, est-il important de pointer trois repères afin de baliser le cheminement et de clarifier la démarche :

- déterminer ce qui caractérise la recherche ;
- préciser ce qui caractérise la recherche-action et plus spécifiquement la recherche-action collaborative ;
- expliquer ce qui caractérise la particularité de cette recherche-action quand elle s'adresse à l'école, dont il est difficile aujourd'hui de cerner tous les enjeux, tant la variété de ses acteurs et de ses missions est complexe, et parfois même buissonneuse.

Pour ce qui est de la recherche, je m'en tiendrai au singulier, alors que nous savons bien qu'elle se décline de façon très différente selon les disciplines et les postures épistémologiques des chercheurs. Toutefois, ce qui la caractérise, c'est d'être en mesure de répondre à quatre critères : la crédibilité, la transférabilité, la fiabilité, et sa reproductibilité, c'est-à-dire la capacité à conduire aux mêmes résultats chaque fois que les expérimentations sont reconduites dans des conditions identiques. Afin de respecter ces critères, toute recherche s'impose une méthodologie rigoureuse composée de six étapes, avec des variantes selon les disciplines : l'analyse du contexte, la question de recherche, la recension des théories convoquées par le sujet, la problématique, l'expérimentation, et l'analyse des résultats.

---

38. Tribune du journal *Le Monde* parue le 25 mars 2019.

À partir de cette première approche de la recherche, comment peut-on définir la recherche-action collaborative ? Les objectifs de la recherche collaborative visent à produire un changement dans un milieu professionnel donné, et à améliorer les pratiques qui y sont liées. Elle implique les acteurs, qui deviendront co-chercheurs. C'est une investigation conjointe, et à ce titre, elle suppose le respect des valeurs et des principes démocratiques, comme l'engagement, l'équité, le partage des ressources et le respect de la diversité humaine. Tous les acteurs d'une recherche-action s'engagent à respecter la méthodologie propre à ce type de recherche, mais sont toujours invités, par ailleurs, à s'exprimer librement en sachant que leur point de vue sera entendu. Tous prennent part aux décisions qui balisent les étapes du processus de recherche.

Par sa nature, la recherche-action collaborative ne peut pas répondre à tous les critères de la recherche académique, notamment celui de la reproductibilité, dans la mesure où elle produit des résultats contextualisés, qui ne prétendent pas forcément à leur généralisation. D'un laboratoire à un autre, il est possible et même indispensable de recréer les mêmes conditions d'expérimentation pour confirmer la véracité des conclusions. Mais d'une école à une autre, d'une classe à une autre, d'une entreprise à une autre, d'une situation d'apprentissage à une autre, il serait illusoire de vouloir dégager des conclusions valables pour toutes les autres structures qui semblent similaires. Néanmoins, toute recherche-action éclaire des situations qui se révèlent souvent analogues, et peut, parfois, favoriser des transferts pour des problématiques équivalentes. Par ailleurs, la particularité de la contextualisation ne dédouane pas la recherche-action des autres contraintes de la recherche académique. Elle n'échappe ni à l'éthique des procédures, ni à la planification d'un projet qui tient compte des contraintes du milieu, ni à la viabilité du changement ou des solutions qui découlent des résultats.

Au-delà de la clarification des termes, nous devons aller plus loin dans l'étude comparative entre la recherche académique et la recherche-action collaborative, parce que c'est dans leur mise en œuvre réciproque que se jouent de multiples nuances. Ces nuances transparaissent derrière trois mots : collaboration, coopération, et terrain. Sylvain Connac (2021)[39] décrit :

- la collaboration comme une situation dans laquelle les acteurs travaillent individuellement sur un même projet : « je fais quelque chose avec l'autre parce que j'y trouve un avantage pour moi-même »
- et la coopération comme une situation dans laquelle les acteurs sont complètement interdépendants : « je fais quelque chose avec l'autre et pour l'autre parce que j'éprouve une satisfaction à l'idée de la satisfaction de l'autre ». Dans ce cas, le groupe ne peut atteindre le résultat attendu qu'à condition que chaque membre contribue d'une façon ou d'une autre à la production de l'ensemble.

39. Sylvain Connac (2017). *La Coopération entre les élèves,* Canopé Éditions, Collection Éclairer, p. 21.

Pour faire référence à la contextualisation évoquée plus haut, cela suppose que la recherche prenne en compte la particularité du lieu où elle se déroule, qu'il s'agisse d'une école, d'un lieu de formation ou d'une entreprise. La recherche académique ou la recherche fondamentale se confrontent aux savoirs que partage l'humanité, c'est-à-dire aux connaissances et aux théories qui fondent ces savoirs. Dans le cas de la recherche-action, nous parlons de terrain non seulement en tant que lieu, mais bien davantage en tant que source même des observations qui sont menées. C'est un peu comme si nous avions affaire à une science du terrain. Ce que nous désignons généralement par le terme de savoirs expérientiels. Il convient de noter que la recherche-action ne se dispense pas de faire référence aux théories qui président aux choix des pédagogues ou des formateurs. En somme, elle se donne une double mission, à savoir les exigences de la recherche académique, et la quête de savoirs expérientiels.

La deuxième fonction du terrain, c'est d'être le lieu du questionnement. Toute recherche est fondée sur une question, qu'on appelle précisément question de recherche. Or, dans le cas d'une recherche-action, cette question émane du terrain lui-même. Elle est l'expression de problématiques qui émergent des pratiques ou des situations auxquelles doivent répondre les pratiques professionnelles. C'est ce qui la rend passionnante, contextualisée. Mais ce qui la rend aussi fragile, parce que le terrain se dérobe souvent, étant donné qu'il est composé d'un nombre impressionnant de sédiments qui appartiennent à des disciplines différentes : de la pédagogie à la sociologie, en passant par la psychologie, la didactique, les sciences de l'éducation ou les sciences cognitives.

Mais pour autant, dans toute recherche, les chercheurs sont confrontés à la validité et la pertinence de la question de départ, à savoir : « cette question est-elle une bonne question ? ». Il s'agit de déterminer ce que l'on appelle une « question scientifique ». Et de ce fait, d'être clairvoyant pour limiter le risque de se leurrer, lorsque l'on est confronté à un regard sur sa propre pratique, pour ne pas tomber dans le piège de ses propres croyances, ou pour confirmer ses propres biais cognitifs. La recherche comme la recherche-action collaborative s'insèrent généralement dans un paradigme interprétatif dont il faut parfois se méfier. Daniel Favre précise que :

> « La peur de se tromper développe alors chez les individus une appétence exagérée pour les connaissances sûres et immuables, celles sur lesquelles on peut s'appuyer sans crainte. Il s'ensuit un besoin de certitudes peu compatibles avec le développement de l'esprit critique, la critique n'étant pas perçue comme ce qui fait avancer, mais comme ce qui mortifie »[40].

40. Daniel Favre (2021). *Reconnecter l'école avec le vivant. 10 pratiques pédagogiques à changer pour un nouveau paradigme,* Dunod, p. 125.

La science ne doit pas devenir une machine à fournir des vérités réclamées par ceux qui ont besoin de certitudes. Au contraire, elle doit avancer en assumant et en travaillant sur ses paradigmes, ses croyances et ses erreurs. Comme le résume François Jacob, « l'important dans la recherche, c'est l'imprévisible ». Dans chacun de ces dispositifs ancrés dans une configuration systémique, il y a des angles morts.

Les neurosciences n'échappent pas à cet enjeu de la complexité. Si elles viennent bousculer la philosophie, la sociologie, la psychanalyse ou la pédagogie, elles sont aussi interrogées par ces mêmes disciplines. Se former à des contenus neuroscientifiques ne peut se concevoir sans une analyse réflexive profonde sur sa propre vision de l'homme, que cette vision soit cognitive, psychique ou sociale. Il ne s'agit pas, ici, de remettre en cause les pratiques scientifiques de la recherche. Il s'agit plutôt de questionner son application stricte dans le domaine des apprentissages. Toutes les conditions de la recherche ne peuvent pas être réunies dans une école ou une entreprise comme elles le sont dans un laboratoire, et en particulier, la condition de disposer de structures de contrôle au cours des expérimentations.

Pour répondre au mieux aux exigences de rigueur, il est indispensable, dès le départ, de clarifier les sources d'informations dont disposent les enseignants pour fonder leur questionnement. Edouard Gentaz (2022)[41] définit quatre types de ressources d'informations en sciences cognitives :

- les connaissances scientifiques des revues scientifiques avec comité de lecture ;
- les connaissances scientifiques des revues scientifiques pour les professionnels, qui permettent de décoder ces mêmes sources scientifiques ;
- les connaissances scientifiques destinées à une large diffusion (revue de kiosque) ;
- les connaissances scientifiques présentées dans des ouvrages de synthèse.

Dans une situation idéale, chaque enseignant engagé dans un travail de recherche devrait avoir accès aux trois premiers types de ressources d'information. De plus, bien que cela ne fasse pas partie de sa formation, ni de ses activités quotidiennes, il devrait adopter une méthodologie de recherche rigoureuse. Dès lors, peut-on vraiment appeler « recherche collaborative » des méthodologies qui se limitent à des exigences unilatérales ? Cela ne va pas sans rappeler les écrits de Josse Clichtove, professeur à la Sorbonne en 1520, qui distingue les « arts libéraux », nobles, des « arts mécaniques » ou « serviles ». Il insiste ainsi sur le privilège de la dignité de l'esprit. L'Université se pose comme toute puissante dans ce domaine, et fait globalement fi des

41. Edouard Gentaz (2022). *Les Neurosciences à l'école : leur véritable apport,* Odile Jacob, p. 64.

savoirs expérientiels, dans le domaine de l'éducation. D'ailleurs, le modèle universitaire ressemble au modèle scolaire, en ce qu'ils sont prioritairement descendants. Les sachants enseignent aux non-sachants sans prendre suffisamment en compte les écarts entre les acquis et les compétences nécessaires à de nouvelles connaissances. Afin qu'enseignants et chercheurs puissent coopérer en interdépendance, il serait équitable que les chercheurs se donnent les mêmes règles que celles qu'ils attendent des enseignants. Stanislas Dehaene et Elena Pasquinelli (2021)[42] précisent que :

> « La recherche translationnelle est une démarche qui vise à évaluer ce qui marche dans la vie réelle à l'échelle d'une classe, d'une école, d'une académie, ou même d'un pays tout entier, et pas seulement en théorie, dans des expériences de laboratoires ou à petite échelle ».

Connaître les grandes étapes de la méthodologie de la recherche pour les uns, connaître l'univers scolaire et ses contraintes pour les autres. Cette bilatéralité des informations est d'autant plus importante que beaucoup d'enseignants pensent qu'ils ne sont pas à la hauteur d'une activité de recherche, même collaborative. Pourtant, dans les écoles et les entreprises, nombreuses sont les demandes d'accompagnement de recherche en neurosciences cognitives pour mieux comprendre l'humain. Mais le désir s'affronte souvent à diverses formes d'inquiétude, notamment celle engendrée par le complexe de l'imposteur. Les enseignants craignent de ne pas être à la hauteur, de voir leur hypothèse invalidée, ce qui laisserait supposer que rien n'a été appris. Ils doutent de la pertinence de leur question de recherche, de leur légitimité. En somme, ils se regardent comme des amateurs. Pourtant, au-delà de ces craintes, s'ouvrent les possibilités de comprendre l'incidence de ce type de travail sur la pratique, à long terme, des uns et des autres, et de favoriser des changements de postures professionnelles.

Les dispositifs d'intégration des neurosciences cognitives à l'école sont très variés. Parfois, les protocoles sont imposés à l'école, qui devient alors un terrain d'expérimentation permettant d'enrichir des bases de données utiles pour les laboratoires. Dans ce cas, la méthodologie correspond aux critères de la recherche. Mais l'école est alors peu participative. Parfois, l'école demande à un laboratoire de l'accompagner pour mener une recherche sur un thème propre à la problématique d'apprentissage. La question de recherche, dans ce cas, correspond souvent à un dysfonctionnement ou à un besoin qui appelle une solution. Trouver une réponse à « qu'est-ce que la mémoire », pour un chercheur, n'est pas de même nature que de répondre à la question : « comment améliorer la mémoire des élèves », pour les enseignants. Savoir ce qu'est la mémoire n'implique pas forcément la possibilité de sa mise en œuvre en neurodidactique et en pédagogie.

---

42. Stanislas Dehaene, Elena Pasquinelli (2021). *Op. cit.*

Dans le contexte de la recherche collaborative en milieu scolaire, le partenariat s'étend sur trois ans, durée minimale pour mettre en place un dispositif de recherche. Au cours de cette démarche, les différents acteurs s'engagent dans une investigation conjointe. Tous « se mettent en recherche », chacun à sa manière, avec une expertise différente mais complémentaire. L'objectif commun implique de travailler pour un codéveloppement d'expertises pour l'avenir. Dans cette perspective, il s'agit d'instaurer une dynamique de changement, en remodelant la pratique et le développement professionnel des enseignants, les méthodes d'apprentissage des élèves, et bien sûr, les paradigmes éducatifs. Cette démarche implique à la fois l'équipe éducative et les chercheurs, deux composantes qui font partie intégrante de ce dispositif.

L'aide proposée par les laboratoires à l'équipe d'enseignants ressemble à un dispositif de recherche, à savoir :

- partir d'un constat qui émane du terrain scolaire ;
- travailler sur une question de recherche, qui fait suite à l'analyse des constats ;
- réaliser une recension de références scientifiques pouvant répondre à la question de recherche ;
- formuler une problématique qui engagera une expérimentation ;
- expérimenter dans la classe ;
- évaluer et critiquer des résultats ;
- communiquer par l'écriture commune d'un article, et une présentation éventuelle à un colloque.

Néanmoins, cette forme de collaboration ne peut se réaliser que dans les établissements capables de libérer du temps à leur équipe d'enseignants, soit en moyenne trois à neuf jours de disponibilité par an (totalisant entre dix et trente jours sur trois ans). Ces périodes correspondent à des moments où les enseignants ne seront pas face aux élèves. Ajoutons à cela la lecture personnelle et/ou collective d'articles et d'ouvrages scientifiques. Enfin, pour que ce dispositif puisse se réaliser, il est nécessaire que le chef d'établissement soit à la fois leader et porteur du projet, puisqu'il est responsable des ressources humaines de son équipe, de l'évolution de leurs compétences, et de l'évolution de son projet d'établissement. Un projet d'action collaborative engage tout l'établissement, sous peine de rester… lettre morte.

Ce partage de regard et de compétences non hiérarchisées est fondamental. Car, si les uns sont dans la recherche fondamentale, et les autres sont en recherche de solutions ou d'applications possibles dans l'apprentissage, une alliance méthodologique sera complexe. Gentaz (2021)[43] engage les

43. Edouard Gentaz (2022), *op. cit.*

enseignants à se fier aux méta-analyses diffusées par les chercheurs pour trouver des résultats à leur problématique concernant les apprentissages. C'est nécessaire et important, sans aucun doute, mais encore faudrait-il que les enseignants soient en mesure de passer de la méta-analyse à des situations contextualisées, ce qui est l'un des objectifs de la recherche collaborative. Quel serait l'élément qui pourrait venir débloquer l'engrenage engagé depuis si longtemps, pour qu'une autre logique de collaboration entre un laboratoire et une école puisse émerger ? Peut-être en explorant les représentations des savoirs, qui animent encore les pratiques d'enseignement scolaire et universitaire pour faire évoluer les paradigmes liés à l'apprentissage. « Apprendre » est encore synonyme de « recevoir ». L'école comme l'Université peinent à s'adapter aux évolutions du monde, tant dans les changements sociétaux, environnementaux, que dans l'extension des diffusions du savoir par l'intelligence artificielle. Par exemple, « l'intelligence artificielle pourrait faire évoluer les pratiques pédagogiques, en mettant en place des dispositifs d'apprentissage adaptatif, pour guider les apprenants dans leurs apprentissages »[44].

Enfin, l'organisation de l'espace et du temps, immuable depuis des décennies, ne correspond plus à ce que les neurosciences cognitives nous apprennent actuellement sur les conditions optimales des apprentissages. Pourtant, à partir des recherches en neurosciences cognitives, il est envisageable que l'école et l'université s'engagent dans une réflexion collective sur l'apprentissage, pour entreprendre des actions favorisant un changement profond des pratiques. Le cerveau ne fonctionne pas à partir d'un empilage d'informations ou de blocs agglomérés les uns aux autres. Il répond à une algorithmicité cérébrale déclinée en fonctionnalités.

En bref, le cerveau est :

- projectif, donc anticipateur et générateur d'hypothèses ;
- narrateur et interprétateur ;
- simulateur d'actions et prédicteur des conséquences des actions ;
- émulateur ou transformateur du monde. En quelque sorte, il est humanisateur ou créateur de sens.

Aussi l'apprentissage doit-il être envisagé à partir de ces puissantes compétences neurologiques, et non selon un schéma de simple transmission.

---

44. UNESCO (2022), « L'apprentissage adaptatif : réflexions de la Fondation l'IA pour l'école », *Institut de France,* BIE, Bureau international d'éducation.

CONCLUSION

Isabelle Stengers, philosophe des sciences, nous invite à être à la hauteur du monde tel qu'il se fait, et ainsi, à vivre l'aventure du savoir à travers ce monde tel qu'il évolue ! En ce quart de siècle bien engagé, la recherche collaborative pourrait répondre à d'autres critères, et se développer à partir de quelques principes clairement définis. Les laboratoires de neurosciences cognitives s'intéressent à l'école ; mais, pour rejoindre l'école, peut-être faudrait-il en effet qu'ils acceptent les différents niveaux de la pyramide des preuves, afin d'amener peu à peu les enseignants à des protocoles plus rigoureux. Cette ascension vers la rigueur scientifique impose des règles éthiques d'égalité de paroles et de pensées. De leur côté, les enseignants sont prêts à adopter une approche scientifique de l'apprentissage. Ce qui implique de remettre en question certains paradigmes, hérités de leur passé d'élève ou de leur formation initiale.

Le point de convergence entre tous les acteurs d'une recherche collaborative serait de situer leur travail dans une logique de développement professionnel partagé. Concrètement, cela signifie d'encourager le doute, le scepticisme, l'écoute, permettant la remise en question des certitudes établies. Cette approche favorise le débat, la polémique et la créativité, contribuant à un échange constructif de savoirs en neurosciences de l'éducation. Chacun détient une partie de la réponse, mais cette réponse n'a de sens que dans une réflexion commune, engendrant des ruptures avec d'anciens schémas, et créant de nouvelles perspectives pour le système scolaire dans son ensemble. À ce jour, il n'existe pas de méthodologie clairement définie qui engagerait cette collaboration entre un laboratoire et une école. Pour la simple raison que ses contours n'ont pas encore été tracés de manière collaborative. Le chemin est entrepris, puisque le sujet est controversé. Il suscite des débats, voire des conflits, et peut même donner lieu à une crise. Cette situation est illustrée de manière intéressante par deux idéogrammes de la langue chinoise : l'un symbolise le danger, et l'autre représente l'opportunité.

Bibliographie

Bruno Bourassa, Mehdi Boudjaou (2012). *Des recherches collaboratives en sciences humaines et sociales. Enjeux, modalités et limites,* Laval, Presses de l'université Laval.

Edouard Gentaz (2022). *Les Neurosciences à l'école : leur véritable apport,* Paris, Odile Jacob, pp. 64-66.

Edouard Gentaz (2017). *« Co-construction de recherches en sciences cognitives interventionnelles : un outil pour développer la formation initiale et continue des enseignants »*, éditorial, ANAE, n° 146.

Elena Pasquinelli, Mathieu Farina (2020). *L'Art de faire confiance : pour un nouveau contrat entre la science et les citoyens,* Paris, Odile Jacob.

Eric Tardif, Pierre-André Doudin [dir] (2016). *Neurosciences et cognition : perspectives pour les sciences de l'éducation,* Bruxelles, De Boeck Supérieur, Collection Pédagogie en développement.

François Taddei (2022). *Et si nous ? Comment relever ensemble les défis du XXIème siècle,* Calmann Levy.

Guy Le Bouedec, Titoune Lavenier, Luc Pasquier (2016). *Les Postures éducatives : de la relation interpersonnelle à la communauté apprenante.* L'Harmattan.

Isabelle Stengers (2013). Une autre science est possible : manifeste pour un ralentissement des sciences, Paris, La Découverte, collection Les empêcheurs de penser en rond.

Isabelle Stengers (2017). *Another Science is Possible: A Manifesto for Slow Science.*

Jacques Attali (2022). *Histoires et avenirs de l'éducation,* Flammarion.

Les chercheurs ignorants (2015). *Les Recherches-actions collaboratives : une révolution de la connaissance,* Rennes, Presses de l'EHESP, collection Politiques et interventions sociales.

Marie-Anne Hugo, Claude Siebel (2018). *Recherches impliquées - recherche-action : le cas de l'éducation,* Bruxelles, De Boeck Université.

Michel Serres (1992). *Le Tiers-Instruit.* Folio, Essais.

Stanislas Dehaene (2018). *Apprendre ! Les talents du cerveau, le défi des machines,* Paris, Odile Jacob.

Stanislas Dehaene [dir] (2021). *L'École éclairée par la science : travaux du Conseil scientifique de l'éducation nationale,* Paris, Odile Jacob.

Stuart Firestein (2021). *Les Continents de l'ignorance,* Paris, Odile Jacob, collection sciences.

Sylvain Connac 2017). *La Coopération entre les élèves,* Poitiers, Canopé Éditions, Collection Éclairer.

Catherine Becchetti-Bizot, Guillaume Houzel et François Taddei (2017). *Vers une société apprenante : rapport sur la recherche et développement de l'éducation tout au long de la vie.* Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche. Mars.

Sentiment d'efficacité personnelle : déterminant ou médiateur du lien entre stress perçu et agentivité

Stéphanie de Chalvron, Hanna Chainay, Franck Tarpin-Bernard, Alexandre Beaussier, Riadh Lebib, Valentine Facque, Florence Cousson-Gélie

Résumé

L'agentivité, concept clé dans le contexte professionnel, repose sur le sentiment d'efficacité personnelle (SEP, Bandura, 2000). Toutefois, son impact positif sur l'agentivité serait altéré par le stress perçu (Myers et al., 2012). Ou bien, est-ce peut-être le SEP qui permet de diminuer l'effet négatif du stress sur l'agentivité ? Pour le savoir, nous avons comparé le rôle du stress perçu dans la relation qui unit le SEP à l'agentivité (modèle 1), et celui du SEP dans la relation qui unit le stress perçu à l'agentivité (modèle 2). De manière générale, nos résultats indiquent que le SEP atténue l'effet du stress, et que ce dernier a finalement peu d'effet sur le SEP. En conséquence, les formations professionnelles visant à améliorer l'agentivité devraient se concentrer sur le développement du SEP des participants. Nos outils de mesure pourraient être utilisés systématiquement pour évaluer l'impact des dispositifs pédagogiques.

INTRODUCTION

L'agentivité, définie par Bandura (2000) comme la capacité d'un individu à exercer un contrôle sur son propre fonctionnement et sur les phénomènes environnementaux, est un concept clé dans l'univers professionnel, et notamment pour ce qui concerne la formation. Il ne s'agit plus uniquement de considérer l'apprentissage comme ayant une valeur économique, mais comme un moyen d'autonomisation et de réalisation personnelle et professionnelle en tenant compte d'un environnement en évolution constante (Alheit & Dausien, 2005). Par ailleurs, les dispositifs pédagogiques ludiques portent en eux des caractéristiques qui semblent favorables au développement de l'agentivité des apprenants. En proposant des situations pédagogiques peu stressantes, mais au contraire très engageantes, ces dispositifs peuvent en effet agir sur l'agentivité et le sentiment d'efficacité personnelle (SEP). Dès lors, il nous semble essentiel d'être en mesure d'évaluer l'effet d'un dispositif pédagogique ludique sur l'ensemble des éléments constitutifs de la capacité d'agir des apprenants et des formateurs.

Dans cet article, à travers l'exploration des liens entre stress perçu et agentivité, nous construisons les bases de ce qui peut constituer un outil de mesure de l'usage d'un dispositif pédagogique afin d'évaluer son impact sur la capacité d'agir. L'agentivité s'enracine dans la théorie sociale cognitive de Bandura (1986), selon laquelle les conduites humaines sont régulées à la fois par l'environnement dans lequel un individu se trouve, et par ses caractéristiques personnelles, les deux étant en interaction permanente. Selon cette théorie, l'agentivité se produit quand l'individu exerce volontairement, et en toute conscience, son influence sur ses propres actions, sur autrui, et sur l'ensemble des systèmes qui forment son environnement (Bandura, 2006). En d'autres termes, quand il s'autorégule en mobilisant l'ensemble de ses ressources (cognitives, émotionnelles, physiques et environnementales). En ce sens, l'agentivité est aussi reliée à l'autodétermination. En effet, elle suppose que l'individu se fixe un objectif, qu'il soit en mesure de faire des choix, de s'engager, et de persévérer pour atteindre l'objectif fixé. L'autorégulation et l'autodétermination sont donc le fondement de l'agentivité, et elles sont sous-tendues par une ressource individuelle fondamentale, à savoir le sentiment d'efficacité personnelle (SEP) (Bandura, 2001).

Le SEP s'appuie sur la croyance en nos motivations et en nos ressources cognitives (raisonnement, mémoire, attention…), et surtout sur la croyance qu'on est capable – ou non – de mettre en place les comportements nécessaires pour réaliser une tâche particulière. Ainsi, pour Bandura & Cervone (1983), le SEP serait le principal mécanisme régulateur des comportements. D'ailleurs, les résultats de la méta-analyse de Sadri & Robertson (1993) démontrent que le SEP est relié positivement à la performance et au choix du comportement. Le SEP n'est donc pas une mesure des aptitudes réelles ou objectives d'une personne, mais reflète une croyance relative à ce qu'elle

se sent capable de faire dans diverses situations, quelles que soient ses aptitudes, difficultés ou facilités, à agir. Le SEP participe non seulement à la prise de décision, mais aussi au fait de s'investir, ou non – et de poursuivre cet investissement –, dans divers projets. Enfin, il régule les réactions émotionnelles face aux obstacles ou défis rencontrés. Il a été démontré que le niveau de SEP se révèle un bon prédicteur d'une pluralité d'issues, que ce soit dans les domaines de la santé mentale, des apprentissages, des choix et développements de carrière professionnelle, de l'employabilité… (Lecomte, 2004). Un SEP élevé influence positivement la performance d'un individu directement, en lui permettant de mobiliser et d'organiser ses compétences, et/ou indirectement, en influençant le choix de ses objectifs et de ses actions. Par ailleurs, un SEP élevé réduit le stress, car les personnes pensent avoir un certain contrôle sur les situations qu'elles rencontrent, et il diminue la vulnérabilité à la dépression.

Lazarus & Folkman (1984) définissent le stress comme relevant d'une transaction entre, d'une part, un individu ; et, d'autre part, la situation à laquelle il est confronté, et plus globalement, son environnement au sens large, dont il perçoit l'issue comme « débordant ses ressources et pouvant mettre en danger son bien-être ». Cette perspective transactionnelle du stress est proche de celle de Bandura, puisqu'elle insiste sur l'aller-retour permanent entre l'individu et son environnement, dans lequel l'individu n'est pas le seul à se trouver modifié (par la situation, mais aussi par ses propres cognitions et conduites). En effet, l'environnement est également modifié, et ce, en fonction de la manière dont l'individu réagit. Tout comme dans la théorie de Bandura, qui insiste sur la bi-directionnalité des influences, c'est-à-dire le fait que les individus sont à la fois produits et producteurs de leur environnement, la théorie transactionnelle du stress (Lazarus & Folkman, 1984) considère l'individu comme l'agent principal dans l'interaction, entre des dimensions biologiques, psychologiques et sociales qui lui sont propres, et les caractéristiques de la situation à laquelle il fait face. Il va tout d'abord évaluer cette situation (représente-t-elle une menace, un défi, a-t-il quelque chose à y gagner ou à y perdre ?), et ensuite il va s'évaluer lui-même en termes de ressources (cognitives, physiques, sociales…). Cette évaluation va donner lieu à une perception du stress (la situation est-elle finalement stressante ou pas ?).

Si l'effet direct du SEP sur l'agentivité n'est plus à démontrer (Bandura & Lecomte, 2019; Corner et al., 2015; Tapal et al., 2017), la nature des effets du stress perçu dans cette relation est un peu moins claire. En effet, Bandura affirme que le SEP permet, de manière générale, de réduire le stress, et ceci chez différentes populations, des adolescents (Faurie et al., 2016), des adultes au travail (Pérez-Fuentes et al., 2018; Pérez-Fuentes et al., 2019) ou encore des personnes âgées (Tovel et al., 2019; Trouillet et al., 2009). Ainsi, le SEP aurait un effet direct sur le stress perçu. Parallèlement, quelques recherches se sont penchées sur le rôle du stress perçu sur l'agentivité. Elles indiquent qu'il diminue l'agentivité notamment à prendre soin de soi (Myers et al., 2012). Dans cette étude, les auteurs démontrent que chez des étu-

diants, le stress limite les comportements de santé sains, comme s'alimenter correctement, faire du sport, respecter son sommeil. On peut donc supposer que le stress perçu diminue de manière générale l'effet bénéfique du SEP sur l'agentivité (modèle 1). L'étude de Schwarzer & Hallum (2008) démontre aussi cet effet du stress perçu. Le SEP y apparaît bien comme une ressource protectrice contre le stress au travail. Mais le stress perçu se traduit directement en épuisement professionnel. De même, l'étude de O'Leary (1992) montre que le SEP favorise l'adoption de bons comportements de santé, mais que le stress perçu diminue l'effet du SEP sur cette relation. On peut supposer que le stress perçu agit de la même manière pour ce qui concerne l'effet du SEP sur l'agentivité au travail. Un SEP élevé augmente l'agentivité des employés. Toutefois, si le stress perçu est élevé, alors il est possible que l'effet positif du SEP sur l'agentivité s'en trouve diminué. Toutes ces études mentionnent donc que le SEP est un déterminant (ou un précurseur) de l'agentivité, mais que cette relation est ce qu'on appelle médiatisée par le stress perçu, c'est-à-dire qu'elle dépend du niveau de stress perçu. Pour autant, certaines études démontrent que c'est plutôt le SEP qui est un médiateur, et le stress perçu, un déterminant (modèle 2). Par exemple, l'étude de Yakimova et al. (2021) montre que le SEP médiatise la relation entre les traits de personnalité et la qualité de vie (QdV). Celle de Lee et al. (2016) montre qu'il médiatise le lien entre le stress perçu et la satisfaction de vie. Par ailleurs, dans l'étude longitudinale de Maciejewski et al. (2000), il a été montré que le SEP est un médiateur des événements stressants, à la fois dans un groupe de personnes avec des antécédents de dépression, et dans un groupe de personnes sans antécédent. Plus précisément, dans le premier groupe, les personnes ayant un SEP élevé font moins de récidives (ou les symptômes sont moins graves). Dans le second, un SEP élevé protège contre la dépression.

L'absence de consensus dans la littérature, quant aux rôles de déterminant ou de médiateur attribués au SEP et au stress perçu, nous amène à nous demander quel est le meilleur modèle. D'un côté, le rôle bien établi du médiateur du stress perçu dans le modèle de Lazarus & Folkman (1984), de même que celui de déterminant de l'agentivité pour le SEP dans le modèle de Bandura (2000), suggèrent l'existence d'un modèle (modèle 1) où le SEP est déterminant (x), le stress perçu médiateur (m) et l'agentivité le critère (y). D'un autre côté, plusieurs travaux (Lee et al., 2016 ; Maciejewski et al., 2000 ; Yakimova et al., 2021) éclairent l'existence d'un autre modèle (modèle 2), où le stress perçu est le déterminant (x), le SEP le médiateur (m) et l'agentivité le critère (y). C'est une question importante, car, selon le cas, les formations destinées à augmenter l'agentivité des personnes devront soit plus se centrer sur l'amélioration du SEP, dans le cas où il serait médiateur ; soit se centrer sur la gestion des émotions, dans le cas où le stress perçu serait médiateur. En effet, si le SEP médiatise la relation entre le stress perçu et l'agentivité, cela signifie qu'il absorbe une part ou le total de l'effet du stress, et donc qu'il représente le facteur le plus important pour développer l'agentivité. En revanche, si le stress est médiateur, l'inverse se produit, c'est-à-dire qu'il absorbe une partie ou la totalité de l'effet du SEP ;

et donc il serait nécessaire de d'abord permettre aux personnes de bien gérer leur ajustement au stress, avant de chercher à développer le SEP. C'est une question importante dans le cadre de la mesure de l'efficacité des dispositifs pédagogiques innovants, notamment pour ceux qui portent en eux des caractéristiques favorables au développement de l'agentivité. Selon que ces dispositifs proposent des situations pédagogiques plus ou moins stressantes, comment agissent-ils sur l'agentivité et sur la gestion du stress ? *In fine*, comment doivent-ils également organiser les parcours pour mieux aider à la gestion du stress, et donc au développement du SEP et de l'agentivité ? S'il semble essentiel d'être en mesure d'évaluer l'effet d'un dispositif pédagogique ludique sur la capacité d'agir des apprenants et des formateurs, comment s'organisent réellement les variables constitutives de cette capacité d'agir ?

I. MÉTHODOLOGIE

1. DESCRIPTION DE L'ÉCHANTILLON

Au cours de l'année 2021, une série de questionnaires a été proposée en deux vagues (à deux mois d'intervalle) dans le cadre de l'observatoire de la Fabrique Spinoza®, à des personnes en activité professionnelle via l'institut Think®.

La première vague (étude 1) a rassemblé 1 005 répondants, dont 487 hommes âgés en moyenne de 44,24 ans (ET = 9,28) et 518 femmes âgées en moyenne de 41,28 ans (ET = 10,46). La seconde vague (étude 2) a rassemblé 1 004 individus, dont 483 hommes âgés en moyenne de 44,81 ans (ET = 10,08) et 521 femmes âgées en moyenne de 42,94 ans (ET = 10,34). Au total, les âges vont de 18 à 64 ans.

2. OUTILS DE MESURE

SEP

Nous avons utilisé quatorze items de l'échelle de Sherer et al. (1982) traduite en français par Chambon et al. (1992). Elle comporte à l'origine vingt et un items mesurés sur une échelle de Likert en cinq points de : 1 = pas du tout d'accord à 5 = tout à fait d'accord. Ces points sont répartis en deux dimensions : le sentiment d'efficacité générale et le sentiment d'efficacité personnelle sociale. Nous avons choisi de ne conserver que les quatorze items mesurant le sentiment d'efficacité personnelle. La consistance interne est bonne dans les deux échantillons ($\alpha = 0{,}85$ et $\alpha = 0{,}84$).

Agentivité

Il n'existe pas d'échelle courte déjà validée en français. Nous avons donc choisi de traduire et de valider l'échelle de Yamaguchi et al. (2020). Il s'agit d'une échelle de cinq items mesurés sur une échelle de Likert en cinq points : de 1 = fortement en désaccord à 5 = fortement d'accord. La consistance interne est bonne dans les deux échantillons (α = 0,81 et α = 0,83).

Stress

Nous avons utilisé la PSS-4 de Warttig et al. (2013), qui est une version courte de l'échelle de stress perçu de Cohen et al. (1983). Il s'agit donc d'une échelle de quatre items mesurés sur une échelle de Likert en cinq points allant : de 0 = jamais à 4 = souvent. La consistance interne est bonne dans nos deux échantillons (α = 0,78 et α = 0,80).

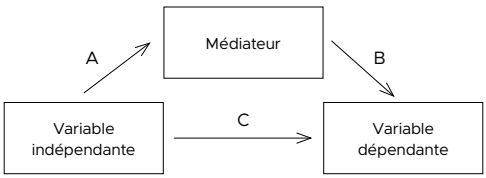
3. PROCÉDURE ET TESTS STATISTIQUES

Nous avons suivi la méthode indiquée par Baron et Kenny (1986). Mais nous avons utilisé des équations structurelles, car, selon El Akremi (2005), elles se prêtent mieux au traitement de variables latentes, en réduisant les biais d'estimation dus aux erreurs de mesure, à la multicolinéarité et aux échelles ordinales (Kenny et al., 1998). En outre, elles permettent de tester simultanément les effets directs et indirects (MacKinnon et al., 2002).

Dans le cadre d'une analyse de médiation, trois variables sont en jeu : la variable indépendante (X), la variable dépendante (Y) et le médiateur (M_e). La variable indépendante doit avoir un effet significatif sur la variable dépendante (effet direct ou effet C'), mais également sur le médiateur (effet A). La variable médiatrice doit avoir un effet sur la variable dépendante quand la variable indépendante est contrôlée (effet B). L'effet indirect (AB ou C-C') doit être significatif ; et l'effet total (effet C) doit non seulement être significatif, mais être plus élevé que l'effet direct (médiation partielle), voire l'annuler (médiation parfaite).

L'ensemble des analyses a été réalisé à l'aide de STATA 17® et nous avons utilisé le package medsem (Mehmetoglu, 2018) pour les analyses de médiation modérée.

FIGURE 1 — Modèle d'analyse de médiation



II. RÉSULTATS

Tout d'abord, nous avons examiné les effets médiateurs du stress perçu (modèle 1 : SEP en X, Stress perçu en M et agentivité en Y) et du SEP (modèle 2 : Stress perçu en X, SEP en M et agentivité en Y) dans chacune des deux études. Les indices d'ajustement des modèles sont bons (tableau 1).

Dans l'étude 1, tous les effets indirects sont significatifs, dans le modèle 1 $\beta = 0{,}10$, $p < 0{,}005$ comme dans le modèle 2, $\beta = -0{,}28$, $p < 0{,}005$. L'effet total est significatif dans les deux cas, $\beta_{\text{modèle1}} = 0{,}61$, $p < 0{,}001$ et $\beta_{\text{modèle2}} = -0{,}45$, $p < 0{,}001$, mais n'annule pas l'effet direct, $\beta_{\text{modèle1}} = 0{,}52$, $p < 0{,}001$ et $\beta_{\text{modèle2}} = -0{,}17$, $p < 0{,}001$.

On retrouve sensiblement les mêmes coefficients, quoiqu'un peu plus faibles dans l'étude 2, puisque là aussi les effets indirects sont significatifs, $\beta_{\text{modèle1}} = 0{,}06$, $p < 0{,}001$ et $\beta_{\text{modèle2}} = -0{,}27$, $p < 0{,}001$, de même que l'effet total, $\beta_{\text{modèle1}} = 0{,}58$, $p < 0{,}001$ et $\beta_{\text{modèle2}} = -0{,}38$, $p < 0{,}001$ qui n'annule pourtant pas l'effet direct, $\beta_{\text{modèle1}} = 0{,}53$, $p < 0{,}001$ et $\beta_{\text{modèle2}} = -0{,}12$, $p < 0{,}001$. Le stress perçu reste également un bon indicateur. En effet, dans tous les cas, il diminue fortement l'agentivité, de même pour le SEP qui l'augmente fortement. Ces résultats sont résumés dans le tableau 1.

TABLEAU 1 — Effets médiateurs du stress perçu (modèle 1) et du SEP (modèle 2)

	β_{direct}	$\beta_{indirect}$	β_{total}	R^2_{adj}	*Prop. Effet total médiatisé*	*Ratio indirect/ direct*
Modèle 1 – Étude 1				0,48****	0,16	0,18
Agentivité						
Stress perçu	-0,17****					
SEP	0,52****	0,10***	0,61****			
Stress perçu						
SEP	-0,56****					
Modèle 1 bis – Étude 2				0,43****	0,10	0,11
Agentivité						
Stress perçu	-0,12****					
SEP	0,53****	0,06****	0,58****			
Stress perçu						
SEP	-0,50****					
Modèle 2 – Étude 1				0,34****	0,63	1,73
Agentivité						
SEP	0,52****					
Stress perçu	-0,17****	-0,28***	-0,45****			
SEP						
Stress perçu	-0,56****					
Modèle 2 bis – Étude 2				0,26****	0,69	2,29
Agentivité						
SEP	0,53****					
Stress perçu	-0,12****	-0,27****	-0,38****			
SEP						
Stress perçu	-0,50****					

***** p < 0,001 ; *** p < 0,005*

Le SEP comme le stress perçu sont donc des médiateurs partiels. Le modèle 1 explique cependant plus de variance, respectivement 48 % et 43 % dans les deux études, que le modèle 2, respectivement 34 % et 26 %. Pour autant, dans les deux études, le stress perçu ne médiatise que 16 % et 10 % de l'effet total, alors que le SEP en médiatise 63 % et 69 %. Les résultats des analyses des bootstraps confirment cependant les effets de chaque médiateur (tableau 2).

TABLEAU 2 — Résultats du bootstrap des analyses de médiation

	Effet conditionnel indirect	Biais	SE	IC 95 %
Modèle 1 étude 1	0,10	-0,0001	0,02	[0,6 – 0,13]
Modèle 1 étude 2	0,06	0,0002	0,02	[0,3 – 0,09]
Modèle 2 étude 1	-0,29	-0,0005	0,02	[-0,34 – -0,25]
Modèle 2 étude 2	-0,27	0,0002	0,02	[-0,31 – 0,22]

Note : réplications 5000 – méthode par percentiles – IC = Intervalle de confiance -

III. DISCUSSION

Cette étude a pour objectif de comprendre le rôle joué par le stress perçu dans la relation qui unit le SEP et l'agentivité. Le SEP est le déterminant principal de l'agentivité (Bandura, 2000). Si on connaît les leviers qui permettent de développer le SEP, à savoir l'apprentissage social, la persuasion par autrui, la gestion des émotions, la maîtrise de soi (Lecomte, 2004), on ne connaît pas précisément le rôle joué par le stress perçu sur l'action du SEP sur l'agentivité. Puisque le stress est un médiateur de la relation entre SEP et burn-out, ou encore un médiateur entre SEP et adoption de bons comportements de santé (O'Leary, 1992b), nous nous sommes demandé s'il médiatise aussi la relation entre SEP et agentivité ; ou si, au contraire, comme l'affirment d'autres auteurs (Maciejewski et al., 2000; Yakimova et al., 2021), c'est le SEP qui médiatise la relation entre stress perçu et agentivité. La question reste ouverte. En effet, nos résultats indiquent que les deux sont bien des médiateurs partiels. Cela dit, nos résultats indiquent aussi que le SEP est un bien meilleur médiateur de l'effet du stress perçu sur l'agentivi-

té (modèle 2), que l'est le stress perçu sur l'effet du SEP sur l'agentivité (modèle 1). Par ailleurs, même si les effets directs nous montrent que le stress perçu fait baisser le SEP et l'agentivité, le poids du SEP sur l'agentivité est très important, peu importe qu'il soit placé en déterminant ou en médiateur. Le modèle où le SEP médiatise la relation entre stress perçu et agentivité est donc le meilleur modèle. Ce qui signifie que le SEP, comme Bandura (2000) l'affirme, permet de mieux s'ajuster aux situations stressantes, et donc que d'aider à son développement notamment via l'ensemble de ses sources (maîtrise personnelle, apprentissage social, persuasion par autrui, régulation des émotions). Le SEP contribuerait à augmenter l'agentivité et à mieux se prémunir des effets délétères et pathologiques (dépression, anxiété, burn-out) du stress.

De manière générale, nos résultats confirment l'ensemble des recherches de Bandura. Le SEP est un puissant prédicteur de l'agentivité humaine. En effet, il peut en être à la fois un déterminant et un médiateur. Le modèle ainsi que les outils développés dans cette étude permettront également d'évaluer des dispositifs pédagogiques non plus sur leur finalité primaire (apprendre telle ou telle connaissance) ; mais sur une finalité plus large, en posant l'hypothèse globale qu'un dispositif innovant dit « efficace » doit permettre de développer à long terme l'agentivité des individus et groupes d'individus.

Limites

L'une des limites de cette étude est sa validité écologique, liée au mode de recueil des données. Toutefois, il s'agit d'une étude exploratoire. Les modèles doivent donc être confirmés par des recherches ultérieures en milieu professionnel direct (dans les entreprises). Par ailleurs, nous n'avons pas contrôlé l'âge des personnes, alors qu'il est parfois indiqué que le SEP est plus faible chez les personnes plus jeunes que chez les personnes plus âgées. Enfin, nous n'avons pas tenu compte des revenus perçus, dans un contexte social délicat où les revenus pourraient avoir une influence sur l'agentivité, notamment via l'engagement ou le désengagement.

CONCLUSION

Notre étude démontre l'importance du SEP dans le développement de l'agentivité, ainsi que sa capacité à diminuer le stress perçu. Les formations d'adultes, qui visent à développer l'agentivité devraient donc d'abord chercher à améliorer le SEP, avant de chercher à aider les individus à mieux réguler leurs émotions. Nous disposons désormais de premiers outils fiables capables de mesurer les différentes composantes de la capacité d'agir, et ainsi, capables d'évaluer l'impact d'un dispositif pédagogique ludique sur la capacité d'agir des apprenants.

Bibliographie

P. Alheit & B. Dausien (2005). Processus de formation et apprentissage tout au long de la vie. L'Orientation scolaire et professionnelle, 34/1, 57-83.

A. Bandura (1986). Social Foundation of Thought and Action : A Social Cognitive Theory. Prentice-Hall.

A. Bandura (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. Educational Psychologist, 28(2), 117-148.

A. Bandura (2000). Self-efficacy : The foundation of agency. In Control of human behavior, mental processes, and consciousness : Essays in honor of the 60th birthday of August Flammer. (p. 17-33). Lawrence Erlbaum Associates Publishers.

A. Bandura & D. Cervone (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. Journal of Personality and Social Psychology, 45(5), 1017-1028.

A. Bandura & J. Lecomte (2019). *Auto-efficacité : comment le sentiment d'efficacité personnelle influence notre qualité de vie* (3e éd.). De Boeck Supérieur.

S. Cohen, T. Kamarck & R. Mermelstein (1983). A Global Measure of Perceived Stress. Journal of Health and Social Behavior, 24(4), 385. https://doi.org/10.2307/2136404

A. Corner, O. Roberts, S. Chiari, S. Völler, E. S. Mayrhuber, S. Mandl & K. Monson (2015). How do young people engage with climate change? The role of knowledge, values, message framing, and trusted communicators. WIREs Climate Change, 6(5), 523-534.

A. El Akremi (2005). Chapitre 12. Analyse des variables modératrices et médiatrices par les méthodes d'équations structurelles : in *Méthodes & Recherches* (p. 325-348). De Boeck Supérieur.

Émilie Morin, G. Therriault & B. Bader (2019). La pertinence du concept de sentiment de pouvoir agir, dans la recherche en éducation, pour faire face aux changements climatiques.

I. Faurie, C. Thouin & C. Sauvezon (2016). Étude longitudinale du stress perçu chez les étudiant·e·s : Effets modérateurs de l'estime de soi et du sentiment d'efficacité personnelle. L'Orientation scolaire et professionnelle, 45/1.

D. A. Kenny, D. Kashy & N. Bolger (1998). Data Analysis in Social Psychology. In S. Gilbert, S. Fiske, & G. Lindsey (Éds.), *Handbook of social psychology* (p. 233-265). McGraw-Hill.

R. S. Lazarus & S. Folkman (1984). *Stress, coping and adaptation.* Springer.

J. Lecomte (2004). Les applications du sentiment d'efficacité personnelle. *Savoirs*, Hors série(5), 59.

J. Lee, K. EunYoung & A. vWachholtz, (2016). The effect of perceived stress on life satisfaction. *Korean Journal of Youth Studies,* 23(10), 29.

P. K. Maciejewski, H. G. Prigerson & C. M. Mazure (2000). Self-efficacy as a mediator between stressful life events and depressive symptoms : Differences based on history of prior depression. *British Journal of Psychiatry*, 176(4), 373-378.

D. P. MacKinnon, C. M. Lockwood, J. M. Hoffman, S. G. West & V. Sheets (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods,* 7(1), 83-104.

M. Mehmetoglu (2018). medsem : A Stata package for statistical mediation analysis. *International Journal of Computational Economics and Econometrics*, 8(1), 63.

S. B. Myers, A. C. Sweeney, V. Popick, K. Wesley, A. Bordfeld & R. Fingerhut (2012). Self-care practices and perceived stress levels among psychology graduate students. *Training and Education in Professional Psychology*, 6(1), 55-66.

A. O'Leary (1992). Self-efficacy and health : Behavioral and stress-physiological mediation. Cognitive Therapy and Research, 16(2), 229-245. https://doi.org/10.1007/BF01173490

G. Sadri & I. T. Robertson (1993). Self-efficacy and Work-related Behaviour : A Review and Meta-analysis. *Applied Psychology,* 42(2), 139-152.

M. Sherer, J. E. Maddux, B. Mercandante, S. Prentice-Dunn, B. Jacobs, R. W. Rogers (1982). The Self-Efficacy Scale : Construction and Validation. *Psychological Reports,* 51(2), 663-671.

A. Tapal, E. Oren, R. Dar & B. Eitam (2017). The Sense of Agency Scale : A Measure of Consciously Perceived Control over One's Mind, Body, and the Immediate Environment. *Frontiers in Psychology,* 8, 1552.

H. Tovel, S. Carmel & V. H. Raveis (2019). Relationships Among Self-perception of Aging, Physical Functioning, and Self-efficacy in Late Life. *The Journals of Gerontology:* Series B, 74(2), 212-221.

R. Trouillet, K. Gana, M. Lourel & I. Fort (2009). Predictive value of age for coping : The role of self-efficacy, social support satisfaction and perceived stress. *Aging & Mental Health,* 13(3), 357-366.

S. L. Warttig, M. J. Forshaw, J. South & A. K. White (2013). New, normative, English-sample data for the Short Form Perceived Stress Scale (PSS-4). *Journal of Health Psychology,* 18(12), 1617-1628.

S. Yakimova, C. Pocnet, A. Congard & D. Jopp (2021). Rôle médiateur du sentiment d'efficacité personnel et des dimensions d'ouverture émotionnelle sur les liens entre personnalité et qualité de vie.

S. Yamaguchi, T. Shiozawa, A. Matsunaga, P. Bernick, U. Sawada, A. Taneda, T. Osumi & C. Fujii (2020). Development and psychometric properties of a new brief scale for subjective personal agency (SPA-5) in people with schizophrenia. *Epidemiology and Psychiatric Sciences,* 29, e111.

CEPAJe, un outil pour évaluer une séance ludopédagogique

Julian Alvarez

Résumé

Cet article décrit le modèle CEPAJe, qui permet de concevoir et d'évaluer une séquence ludopédagogique. De manière concrète, il s'agira de présenter le modèle CEPAJe, et d'illustrer par des exemples d'emploi de jeux sérieux (serious games), comment il permet de préparer, en amont, une séance ludopédagogique ; puis, en aval, d'évaluer cette dernière afin de pouvoir améliorer l'activité pour la fois prochaine.

INTRODUCTION

Quelles évaluations opérer dans le cadre de la ludopédagogie ? Est-ce pertinent de se cantonner à évaluer uniquement les apprenants, ou bien doit-on prendre en considération d'autres dimensions encore ? Par exemple, le jeu que l'on choisit de mobiliser joue un rôle prépondérant à son niveau. En effet, la capacité du jeu à engager, ou bien encore la nature et la qualité des feed-back proposés vont nécessairement influer sur les apprentissages des utilisateurs. Mais d'autres facteurs peuvent aussi être pris en compte, comme le contexte : une connexion internet capricieuse, se trouver à proximité d'un chantier bruyant, etc., peuvent nuire à l'expérience proposée. Quant au professeur lui-même, il peut jouer un rôle significatif sur l'engagement des apprenants.

Ces exemples nous invitent à reconsidérer le périmètre de l'approche évaluative. Doit-il se focaliser sur l'apprenant, ou convient-il de l'élargir à l'ensemble du dispositif sociotechnique mobilisé ? À la nature des feed-back dispensés ? Au contexte d'utilisation ? À la manière d'animer l'activité ? etc. Outre ces questionnements non exhaustifs, il faut également prendre en considération que chaque session d'une même activité ludopédagogique est unique et non reproductible (Laudati et Leleu-Merviel, 2018 : 26-27). Ainsi, la nature des feed-back offerts aux participants varie nécessairement d'une session à une autre. En fonction des événements, des prises de décisions, des références culturelles, de l'humeur du moment, les organisateurs de l'activité et les participants verront leurs performances améliorées ou altérées.

Dès lors, avec une telle approche systémique, on comprend rapidement que nous sommes très loin de pouvoir nous contenter d'une évaluation sommative, à l'instar d'un simple QCM visant à mesurer un stock de connaissances pour un sujet donné à un instant donné. Dans le cadre de cet article, nous proposons ainsi de présenter le méta-modèle évaluatif CEPAJe pour « Contexte, Enseignant, Pédagogie, Apprenant, Jeu » (Alvarez et Chaumette, 2017), qui correspond à la prise en compte de cinq dimensions distinctes, et ce, au regard des différentes phases de l'activité : son introduction, son animation et son débriefing. Si CEPAJe se destine à évaluer une séquence ludopédagogique, il est nécessaire d'illustrer concrètement comment faire usage d'un tel modèle au regard d'un exemple de jeu sérieux. Ceci, tant pour préparer une séance ludopédagogique en amont que pour en évaluer la mise en œuvre en aval.

I. PRÉSENTATION DU MODÈLE CEPAJE

L'un des premiers modèles s'attachant à évaluer une activité de jeu en situation formelle est le « Four-dimensional framework » (modèle à quatre dimensions) élaboré par De Freitas et Martin Oliver (De Freitas et Oliver, 2006). Si leur modèle ouvre des perspectives, il lui manque une représentation temporelle, à l'instar de ce que proposent les trois temps pédagogiques de Tremblay (2007) et la prise en compte de la médiation opérée par le professeur. C'est pourquoi il a été proposé de faire évoluer ce modèle pour concevoir CEPAJe. Ce dernier se présente sous la forme d'un tableau à double entrée, croisant les trois temps pédagogiques de Tremblay (Introduire l'activité, Animer l'activité, Débriefer l'activité) avec cinq dimensions qui composent les principaux items entrant en jeu dans une séquence ludo-pédagogique : le contexte, l'enseignant ou formateur, le scénario pédagogique, les apprenants et le jeu mobilisé.

TABLEAU 1 — Exemple de modèle CEPAJe

Critères évaluatifs/ dimensions	Introduction de l'activité (briefing)	Animation de l'activité	Débriefing de l'activité
Contexte	Le contexte est-il propice à introduire l'activité ?	Le contexte perturbe-t-il l'activité ? Y a-t-il des éléments de contexte qui viennent favoriser l'activité ?	Le contexte est-il propice à débriefer l'activité ?
Enseignant (Tuteur, médiateur pédagogique…)	Les objectifs de l'activité sont-ils clairement présentés tant sur le plan du jeu que sur le plan utilitaire ? Est-ce présenté de manière engageante ?	Habileté à animer le jeu et à accompagner les apprenants durant l'activité de jeu (aide à la lecture et à l'utilisation du jeu…) Les stratégies pour tenter de motiver les apprenants sont-elles adaptées à leurs profils ?	Le débriefing est-il bien mené pour amener les apprenants à faire le lien entre jeu et apprentissages ?
Pédagogie (Scénario de l'activité prenant place dans le scénario pédagogique)	Le scénario propose-t-il d'associer de manière cohérente et équilibrée les aspects jeu et visées utilitaires ?	Le scénario d'utilisation propose-t-il des rôles et objectifs clairs pour le compte des apprenants et des enseignants ?	Le scénario prévoit-il d'exploiter le jeu proposé pour permettre aux apprenants d'atteindre les objectifs utilitaires visés ?
Apprenant (Participants, utilisateurs, élèves, étudiants…)	Les apprenants ont-ils envie de s'engager dans l'activité de jeu sérieux proposée ? Les objectifs sont-ils clairs pour les apprenants ? Recense-t-on des apprenants qui se mettent en retrait par rapport à l'activité ?	Les apprenants montrent-ils de la motivation plutôt intrinsèque ou extrinsèque ? Recense-t-on de l'entraide ou de la rivalité entre pairs ? Recense-t-on des apprenants qui souhaitent modifier ou faire péricliter l'activité ?	Les apprenants manifestent-ils de la motivation à comprendre les aspects utilitaires de l'activité ? Les apprenants sont-ils à même de pouvoir aider les pairs à faire part de leurs ressentis, messages perçus, apprentissages, des suites concrètes à donner ou de pistes d'améliorations de l'activité ?
Jeu (Artefact : jeu, jouet, simulateur, serious game ou serious toy)	Le jeu propose-t-il des leviers motivationnels donnant envie de s'engager ?	Le jeu propose-t-il des systèmes d'aide pour lire ou utiliser le jeu ? (Tutoriels, moyens de débloquer le joueur, ressources pédagogiques complémentaires…) L'accessibilité est-elle prévue ? Le jeu propose-t-il de faire des liens entre le monde réel et sa diégèse (effet-V) ?	Un bilan est-il proposé au joueur tant sur les aspects jeu que sur les aspects utilitaires ?

Nous obtenons un ensemble de cases qui peuvent accueillir des questions à prendre en considération (tableau 1). Bien entendu, les questions peuvent être complétées et personnalisées selon les besoins de la discipline (mathématiques, français, technologie, informatique…) ou du domaine d'application (santé, éducation, formation, recherche…). De ce fait, des séances ludopédagogiques menées dans le cadre d'un enseignement scolaire, d'une formation professionnelle ou bien dans un contexte d'éducation thérapeutique (Alvarez et al. 2016) pourront présenter des questions différentes.

II. APPROPRIATIONS DU MODÈLE CEPAJe

Le modèle CEPAJe fait l'objet d'appropriations. Certains acteurs proposent de le simplifier, à l'instar de Mélanie Fenaert (2021), pour l'adapter à l'évaluation d'escape games pédagogiques (figure 1). Plus axé sur l'artefact, le modèle revisité retient les dimensions « contexte », « pédagogie » et « jeu » au regard de critères d'évaluation qui évoluent selon les trois temps pédagogiques, et que l'on peut noter selon une échelle à quatre niveaux allant de « -- » à « ++ ».

FIGURE 1 — Grille d'évaluation d'un jeu pédagogique virtuel par Mélanie Fenaert (2021)

Items CEPAJe	Critères d'évaluation	--	-	+	++
INTRODUCTION DU JEU					
Contexte	• introduction immersive, scénario explicite • lien avec thème demandé • cohérence du scénario en tant que fiction				
Pédagogie	• cohérence du scénario avec les objectifs pédagogiques				
Jeu	• règles / objectifs clairs • intro rapide et efficace				
DÉROULEMENT DU JEU					
Contexte	• décors et gameplay cohérents avec le scénario				
Pédagogie	• calibrage de la difficulté en fonction du public (niveau et nombre des activités ou énigmes, imbrication) • calibrage du temps				
Jeu	• activités / énigmes variées • qualité technique (navigation, déplacement, interactivité…)				
DÉBRIEFING DU JEU					
Contexte	• forme du débriefing cohérente avec le scénario				
Pédagogie	• pistes proposées pour le débrief en fin de jeu (pour réserver un retour collectif et favoriser les apprentissages individuels)				
Jeu	• éléments de feedback intégrés au jeu (coups de pouce…) ou à la fin du jeu				

Le modèle CEPAJe peut aussi faire l'objet d'ajouts de colonnes supplémentaires. C'est ainsi qu'Isabelle Motte et Pascal Vangrunderbeeck ont proposé une version enrichie et interactive de CEPAJe dès 2019 (figure 2). Il s'agit de guider le ludopédagogue sur la conception de sa séance ludopédagogique en amont, et d'évaluer l'ensemble de l'activité en aval.

FIGURE 2 — Déclinaison interactive du modèle CEPAJe par Isabelle Motte et Pascal Vangrunderbeeck (2019)

Guide pour la conception, l'animation et l'évaluation d'une activité ludo-pédagogique
inspiré du modèle CEPAJe de Julian Alvarez et Pascal Chaumette

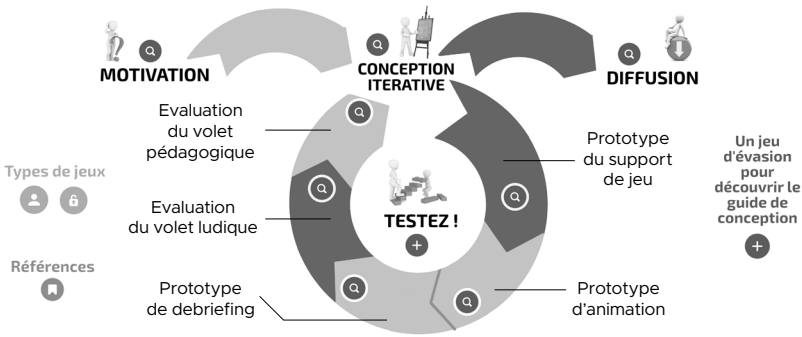
LOUVAIN LEARNING LAB — UCLouvain — par @vgpascal et @zmotte

	Conception du scénario	CONTEXTUALISATION ↓	**Introduction du jeu**	**Animation du jeu**	DÉCONTEXTUALISATION ↑	**+ Debriefing** Retour aux objectifs	**Evaluation** Apprentissages Jeu
Pédagogie	Définir les objectifs		Annoncer les objectifs	Activer les objectifs		Approfondir les objectifs	Objectifs acquis ? Jeu facilitant ?
Jeu	Scénariser le jeu		Présenter le jeu	Réguler le jeu		Rejouer les étapes clés	Adéquation entre apprentissages et jeu ?
Contexte	Penser à l'ergonomie et à la facilitation		Faire découvrir l'environnement de jeu	Stimuler et cadrer la progression		Revenir sur les difficultés (technos ntmt)	Adéquation entre apprentissages et environnement ?
Apprenant	Prendre en compte le profil des joueurs		Expliquer la dynamique entre joueurs	Stimuler et réguler les joueurs		Ecouter le vécu et les questions	Adéquation entre apprentissages et dynamique ?
Enseignant	Définir le rôle de l'animateur		Annoncer le rôle de l'animateur	Observer, animer		Atterrir en douceur	Adéquation entre apprentissages et animation ?

FIGURE 3 — Déclinaison interactive du modèle CEPAJe par Motte, Vangrunderbeeck, Guisset, Malcourant et Renson (2022)

Guide pour la conception d'une activité ludo-pédagogique
Isabelle Motte, Pascal Vangrunderbeeck, Manuela Guisset, Emilie Malcourant et Virginie Renson

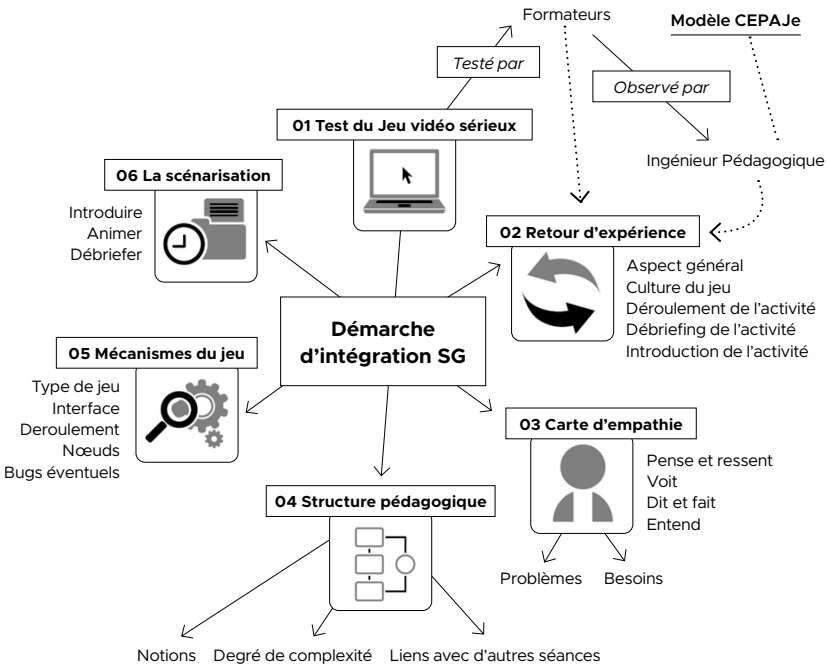
LOUVAIN LEARNING LAB — UCLouvain



Dans la proposition d'Isabelle Motte et Pascal Vangrunderbeeck de 2019, opérée sur Genealy, chaque case vient dévoiler une série de questions clés. Les deux auteurs ont également modifié l'ordre des différentes dimensions, et positionné des flèches de contextualisation et de décontextualisation pour guider au mieux l'utilisateur dans l'emploi du modèle. En 2022, Motte, Vangrunderbeeck, Guisset, Malcourant et Renson ont entièrement repensé le modèle pour séquencer son utilisation en phase de conception. De la partie motivation à la diffusion de la séance ludopédagogique, un ensemble de cases étapes (prototype de support de jeu, d'animation, de débriefing…) oriente le ludopédagogue tout en le renvoyant à des parties précises du modèle CEPAJe (figure 3). Cela permet ainsi de mieux guider le ludopédagogue dans la construction de sa séance ludopédagogique. Il est vrai qu'au regard des nombreuses cases qu'il présente, les premières itérations de CEPAJe (tableau 1 et figure 2) peuvent perdre des utilisateurs débutants.

À noter, enfin, que l'on peut recenser des approches visant à déconstruire le modèle CEPAJe pour le refondre dans des schémas mentaux comme celui de Frédéric Bouyssi (2018), en vue de proposer une démarche intégrative d'un serious game (figure 4).

FIGURE 4 — Approche intégrative d'un serious game par Frédéric Bouyssi (2018)



III. EXEMPLE D'ÉVALUATION AVEC CEPAJe : SERIOUS GAME HSE

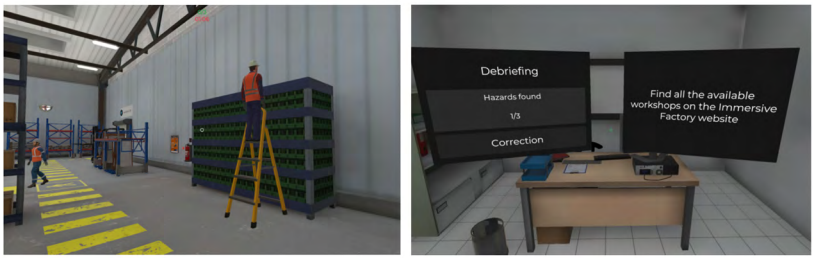
Une fois muni de son modèle CEPAJe, le ludopédagogue est théoriquement en mesure de préparer et d'évaluer l'activité de jeu sérieux. Pour illustrer la manière d'utiliser CEPAJe, nous allons l'appliquer à un exemple concret. Il s'agit de l'emploi du serious game numérique dédié au domaine HSE (hygiène, sécurité et environnement) et mobilisant la modalité « réalité virtuelle » (VR). Précisons que dans cet exemple, nous utiliserons le modèle CEPAJe tel que représenté en tableau 1.

1. CONTEXTE ET DESCRIPTION DU DISPOSITIF

La société Immersive Factory propose un ensemble de serious games numériques liés à la prévention de risques. Les formateurs de la société interviennent en présentiel en entreprise ou sur des sites industriels, dans le cadre de journées sécurité, pour animer des sessions ludopédagogiques. Parmi les applications proposées, nous recensons la chasse aux risques. Concrètement, il s'agit pour les participants, équipés d'un casque VR autonome, d'explorer, dans un temps imparti, un environnement virtuel (bureaux, chantiers, ateliers, entrepôts, chaînes de montage…), et de photographier un nombre défini d'anomalies qui pourraient donner lieu à des accidents, ou à des problèmes de santé comme des troubles musculosquelettiques (TMS). À noter que le nombre d'anomalies à trouver et la durée de jeu sont paramétrables par le formateur. En outre, des anomalies aléatoires sont introduites dans les différentes parties. Ce qui rend chaque expérience unique pour les participants. À l'issue de la partie, l'application fournit un score, et propose de passer en revue les anomalies non détectées (figure 5).

La session étudiée avec CEPAJe se déroule en novembre 2022 dans le nord de la France, sur un site industriel spécialisé dans la fabrication de molécules chimiques pour la production de médicaments. Durant cette journée, une trentaine de salariés ont été convoqués de manière obligatoire. On trouve des cadres, des techniciens et des ouvriers. Un manager en charge de la journée sécurité est à l'origine de la commande de formation. Il souhaite animer les séances ludopédagogiques pour se former à la pratique. L'animateur de la société Immersive Factory se met donc en retrait, pour observer et le seconder le cas échéant.

FIGURE 5 — Exemple d'une chasse aux risques dans un entrepôt



2. SÉQUENÇAGE

Avant de chausser le visiocasque, les participants bénéficient d'un briefing où sont exposés, durant une dizaine de minutes, les enjeux de la formation, la manière d'utiliser le casque de réalité virtuelle autonome, les manettes associées, les règles du jeu et la durée de l'expérience. Puis, les participants sont invités à passer pour jouer, chacun à leur tour. Le formateur se tient à proximité du participant pour prévenir des risques de chutes ou de chocs avec des murs ou des tables. Pendant ce temps, les autres participants patientent en regardant sur grand écran la partie en cours. Une fois le jeu terminé et corrigé, le participant retire son casque, et passe le relais à la personne suivante. Entre deux participants, le casque et les manettes sont nettoyés, pour des questions d'hygiène. Une fois que l'ensemble des participants ont joué, un débriefing général est opéré. Il s'agit de les questionner sur leurs ressentis, les apprentissages effectués et les transpositions avec le monde réel. Enfin, à l'issue de la formation, les participants sont questionnés sur ce que l'on pourrait améliorer dans la séance ludopédagogique proposée. Un questionnaire individuel leur est également proposé sur tablette, pour récolter un ensemble de pistes d'améliorations et des indicateurs destinés aux chercheurs.

3. RETOURS OBTENUS

Les retours recensés s'opèrent sur plusieurs plans, comme les aspects ressenti, ergonomie, technique, contenu, gameplay, pédagogie et recherche. Sur le plan du ressenti, on note que les participants font globalement état de retours positifs sur l'expérience proposée. Pour nombre de participants, c'est la première expérience en VR. Cela se traduit par un effet « waouh ». Cependant, sur une trentaine de participants, l'un d'entre eux a éprouvé de la cinétose. Cet effet de vertige, lié à l'emploi du casque VR, a pu être corrigé en lui proposant d'être assis durant la séance. D'autres ont fait part de vertiges et de maux de tête en regardant le grand écran permettant de voir les différentes parties en cours. Ce phénomène peut s'expliquer éventuellement par un manque d'habitude de la part de certains participants à

regarder ou à pratiquer des jeux vidéo en 3D en vue subjective. Il est possible que cela soit lié au fait que les participants étaient disposés en U autour de l'écran. Peut-être est-il plus confortable de voir de telles images en étant positionné face à l'écran ?

Sur les plans ergonomique et technique, les participants ont fait remonter la nécessité de prendre en main le dispositif VR avant de se lancer dans le serious game proposé. Ceci, dans le but d'être plus réceptif aux contenus proposés. En parallèle, les participants ont fait remarquer que certains boutons des manettes sont inactifs. Or, ils pourraient avantageusement faire office de doublons pour servir la prise de photographies. Enfin, certains participants suggèrent de bien expliquer durant le brief comment prendre des photographies dans le jeu. En effet, en l'état, les anomalies sont validées si elles sont bien centrées sur la photographie. Cela a engendré quelques frustrations chez des participants, stressés par le chronomètre, et qui devaient s'y reprendre à plusieurs fois pour valider un risque détecté.

Sur le plan des contenus, les participants font état de quelques améliorations à apporter sur les risques à proposer ou à modifier. Par exemple, dans un entrepôt, certains fûts ou bidons chimiques ne sont pas positionnés sur des bacs de rétention ; un cariste n'a pas de ceinture de sécurité ; certaines étiquettes sont mal positionnées sur quelques barils…

Sur les aspects du gameplay, d'aucuns ont remarqué que le cariste n'a pas de jambes. Ce détail graphique les a surpris. D'autres ont estimé qu'il y a trop d'anomalies à trouver pour le temps imparti dans le cadre d'une partie jouée sous la forme « session courte ». D'autres encore ont regretté que des lieux soient trop grands à parcourir au regard du temps alloué. Par ailleurs, des participants ont aussi noté que certains tiroirs ouverts comptent comme des anomalies, alors que d'autres non. Ce qui peut être source de confusion. Enfin, quelques participants ont émis le souhait de pouvoir agir davantage dans l'environnement virtuel, comme en conduisant un engin (transpalette, camion, tractopelle…) ou en s'investissant dans des tâches à effectuer dans les lieux visités.

Pour les aspects pédagogiques, des participants ont déclaré avoir appris des choses, par exemple, que l'empilement de palettes en bois ne peut pas excéder 1,80 mètre dans un entrepôt. Certains ont apporté des précisions sur des risques identifiés, ce qui a donné lieu à des échanges d'anecdotes réellement vécues au sein de l'entreprise. D'autres ont voulu pratiquer une chasse aux risques en situation réelle, dans un véritable environnement de travail de l'entreprise, pour vérifier leurs acquis.

Enfin, la majorité des participants ont répondu[45] au questionnaire dédié à la

45.https://docs.google.com/forms/d/e/1FAIpQLSfvSzsl2gtH1ROnEn_vKh4aLtbQxppJf8yyCX5MToVb-Jr6h4A/viewform?usp=sf_link

recherche. Mais il a été perçu comme trop long, en particulier par les techniciens et les ouvriers du site. Parmi eux, certains ont choisi de ne pas y répondre. Ceux qui ont complété le questionnaire ont été perturbés du fait qu'il est impossible de répondre « non » à quelques questions.

4. ANALYSES AVEC CEPAJe

Au regard de ces éléments, il importe désormais de renseigner le modèle CEPAJe pour analyser de manière plus fine ce qui pourrait être amélioré pour cette séance ludopédagogique. Le tableau 2 liste l'ensemble des réponses aux questions posées, en intégrant les éléments précédemment examinés. À cette fin, nous proposons de passer en revue les différentes questions du tableau 1 pour les consigner dans le tableau 2.

TABLEAU 2 — Serious game : la chasse aux risques sur site chimique

Critères évaluatifs/ dimensions	Contexte
Introduction de l'activité (briefing)	**Le contexte est-il propice à introduire l'activité ?** La séance de jeu sérieux est proposée dans le cadre d'une journée sécurité obligatoire. Cependant, certains participants font état de dossiers importants à gérer, et demandent à pouvoir s'en aller dès que possible. La salle proposée et le mobilier sont conformes aux attentes des formateurs pour permettre à la séance ludopédagogique de se dérouler dans de bonnes conditions.
Animation de l'activité	**Le contexte perturbe-t-il l'activité ?** L'écran géant qui transmet l'environnement virtuel du jeu présente par moment des sautes d'image, qui perturbent parfois certains spectateurs. En outre, la disposition en U des participants semble susciter un inconfort chez certains d'entre eux. **Y a-t-il des éléments de contexte qui viennent favoriser l'activité ?** La présence d'un manager qui se sent investi par la mission de sensibiliser les salariés aux risques industriels, par la réalité virtuelle, joue un rôle prépondérant pour susciter l'adhésion des participants. En outre, la salle où se déroule l'activité est insonorisée, et bien équipée d'un point de vue multimédia, avec une bonne connectivité internet.
Débriefing de l'activité	**Le contexte est-il propice à débriefer l'activité ?** Le manager présent induit auprès des participants une écoute attentive. Les participants ont compris que la formation proposée s'inscrit dans des enjeux importants pour chaque salarié, qui risque sa vie en travaillant sur le site industriel. On pourra noter quelques salariés qui sont tiraillés par le fait de boucler des dossiers urgents, et qui s'abstiennent de poser des questions pour ne pas trop retarder leur retour à leur poste de travail.

Critères évaluatifs/ dimensions	**Enseignant** (Tuteur, formateur, médiateur pédagogique…)
Introduction de l'activité (briefing)	**Les objectifs de l'activité sont-ils clairement présentés tant sur le plan du jeu que sur le plan utilitaire ?** Le manager du site connaît bien les aspects en lien avec la sécurité du site, mais découvre pour la première fois la manière d'animer une séance ludopédagogique. De son côté, l'animateur d'Immersive Factory connaît son serious game et la technologie VR. En revanche, il découvre pour la première fois le site industriel : il découvre ainsi l'écosystème, les produits développés et la nature des dangers. Les deux personnes se complètent. **Est-ce présenté de manière engageante ?** Les deux médiateurs cherchent leurs marques respectivement au début, mais finissent par trouver une manière d'opérer conjointement.
Animation de l'activité	**Habileté à animer le jeu et à accompagner les apprenants durant l'activité de jeu (aide à la lecture et à l'utilisation du jeu…)** L'animateur d'Immersive Factory a l'habitude d'assurer des séances ludopédagogiques. Il partage son savoir-faire avec le manager qui cherche à s'initier à la ludopédagogie, en lui confiant le rôle d'animateur de jeu et en le secondant le cas échéant. **Les stratégies pour tenter de motiver les apprenants sont-elles adaptées à leurs profils ?** Le manager connaît ses collaborateurs. Cela joue un rôle prépondérant pour les sensibiliser à l'importance de s'impliquer dans le jeu sérieux et la modalité VR proposés.
Débriefing de l'activité	**Le débriefing est-il bien conduit pour amener les apprenants à faire le lien entre jeu et apprentissages ?** Le débriefing est assuré par le manager. Il a appris le jour même quelles questions clés poser lors d'un débriefing lié à une session ludopédagogique. Il passe en revue le ressenti des participants, tâche d'opérer une distanciation et une conscientisation des apprentissages visés. **Éventuellement, demander comment améliorer l'activité pour la fois prochaine (méta-design).** Cela a été demandé. Il ressort des échanges qu'il faudrait proposer en amont une session de prise en main avant de se lancer dans le jeu, pour être plus réceptif aux erreurs et risques associés. Il est également demandé de bien préciser en amont comment photographier les anomalies.

Critères évaluatifs/ dimensions	**Pédagogie** (Scénario de l'activité prenant place dans le scénario pédagogique)
Introduction de l'activité (briefing)	**Le scénario propose-t-il d'associer de manière cohérente et équilibrée les aspects jeu et visées utilitaires ?** Le scénario pédagogique consiste à utiliser le serious game en autonomie, et à questionner les utilisateurs à l'issue de la partie pour les sensibiliser aux risques.
Animation de l'activité	**Le scénario d'utilisation propose-t-il des rôles et objectifs clairs pour le compte des apprenants et des enseignants ?** Le scénario d'utilisation est orienté uniquement vers l'apprenant qui doit faire usage du jeu. Il n'est pas prévu de scénario d'utilisation pour le médiateur. Ce dernier peut en revanche paramétrer la durée d'une partie, et le nombre d'anomalies à faire trouver. Mais il ne peut pas choisir la nature des anomalies à faire trouver. Ce qui peut empêcher de thématiser la nature des risques à traiter dans une formation.
Débriefing de l'activité	**Le scénario prévoit-il d'exploiter le jeu proposé pour permettre aux apprenants d'atteindre les objectifs utilitaires visés ?** Le débriefing final se construit partiellement autour des risques identifiés dans le serious game. Il s'agit plutôt d'un prétexte pour aborder les véritables risques pouvant être identifiés dans l'entreprise où se déroule l'activité. En outre, l'idée est aussi de capitaliser sur l'expérience VR pour susciter un engagement dans les échanges lors du débriefing.

Critères évaluatifs/ dimensions	**Apprenant** (Participants, utilisateurs, élèves, étudiants…)
Introduction de l'activité (briefing)	**Envie de s'engager dans l'activité de jeu sérieux proposée ?** Oui, c'est clairement observé. Aucun salarié n'a fait part d'une réserve, y compris la personne sujette à la cinétose. **Les objectifs sont-ils clairs pour les apprenants ?** L'objectif du jeu est clair. Le fait de faire passer à la suite les différents participants permet à tout un chacun de bien comprendre les objectifs et le mode de fonctionnement du jeu. En revanche, si le nombre de participants est trop élevé, cela peut induire une attente un peu longue pour jouer. **Recense-t-on des apprenants qui se mettent en retrait par rapport à l'activité ?** Cela n'a pas été constaté. En revanche, plusieurs participants demandent en amont à être rassurés quant à l'emploi du casque de réalité virtuelle et du jeu associé.
Animation de l'activité	**Les apprenants montrent-ils de la motivation plutôt intrinsèque ou extrinsèque ?** C'est extrinsèque, car ce sont les dimensions ludiques et la technologie VR proposées qui servent de levier motivationnel. **Recense-t-on de l'entraide ou de la rivalité entre pairs ?** Une entraide entre participants est constatée via l'emploi de l'écran géant, qui permet de guider les joueurs qui seraient bloqués ou qui passeraient à côté d'une anomalie à recenser. Dans ce contexte, les formateurs soit se mettent en retrait pour laisser les participants aider leurs collègues, soit interviennent pour réduire l'aide afin de laisser le joueur trouver par lui-même. **Recense-t-on des apprenants qui souhaitent modifier ou faire péricliter l'activité ?** Non.
Débriefing de l'activité	**Les apprenants manifestent-ils de la motivation à comprendre les aspects utilitaires de l'activité ?** L'expérience vécue est globalement positive. Cela donne lieu à du partage d'anecdotes et parfois de taquineries entre participants. En général, les jeunes sont mis en avant par leurs collègues pour leurs compétences à jouer. Les participants plus âgés voient l'intérêt d'une telle approche pour aborder la thématique de la sécurité pour les nouveaux arrivants ou les jeunes. En effet, les salariés plus âgés font état d'une expérience en matière de sécurité, qui peut donner le sentiment qu'ils connaissent mieux le sujet que les jeunes générations. Le questionnaire final n'est pas renseigné par certains participants, qui le trouvent trop long. **Les apprenants sont-ils à même de pouvoir aider les pairs à faire part de leurs ressentis, messages perçus, apprentissages, des suites concrètes à donner ou de pistes d'améliorations de l'activité ?** Lorsque le jeu propose d'identifier en fin de partie les anomalies non trouvées, certains participants aident systématiquement le joueur à comprendre la nature et la localisation de l'erreur. Certaines anomalies non trouvées font parfois état de contestations : c'est la faute de l'appareil photo, qui ne permet pas de prendre en compte facilement les risques détectés. Ou bien, certains participants déclarent avoir perdu du temps sur des anomalies que l'application n'intégrait pas. Enfin, des aspects du gameplay donnent lieu à des remises en question : mieux positionner les étiquettes sur les barils… Sinon, c'est plutôt le manager qui guide les participants pour faire le lien entre le jeu et la sécurité sur site. Dès lors, les participants font état d'anecdotes, et élargissent les échanges à d'autres types de risques, qui dépassent ce que le jeu a pu présenter. Globalement, les salariés ont complimenté le manager pour l'initiative d'employer de la VR et du jeu associé dans le cadre d'une journée sécurité.

Critères évaluatifs/ dimensions	**Jeu** (Artefact : jeu, jouet, simulateur, serious game ou serious toy)
Introduction de l'activité (briefing)	**Le jeu propose-t-il des leviers motivationnels donnant envie de s'engager ?** La modalité VR constitue le principal attrait du jeu. Pour une majorité de participants, c'est la première fois qu'ils font usage de cette technologie.
Animation de l'activité	**Le jeu propose-t-il des systèmes d'aide pour lire ou utiliser le jeu ? (Tutoriels, moyens de débloquer le joueur, ressources pédagogiques complémentaires…)** Le jeu sérieux commence par proposer un tutoriel pour la prise en main des commandes du jeu : déplacement et principes pour photographier des risques. Les participants ont souligné des difficultés ergonomiques : trouver sur la manette le bouton permettant de photographier. En outre, le jeu propose d'explorer des surfaces trop grandes au regard du temps imparti. Le facteur temps conjugué à une faible compétence à jouer a constitué pour certains participants un stress durant la partie. À noter que, pour des utilisateurs peu habitués à jouer, l'aspect chronomètre semble constituer une source de stress, qui les empêche de repérer certaines anomalies durant la partie. **L'accessibilité est-elle prévue ?** Il est proposé des versions anglaise et française du même serious game, pour s'adresser à des publics non francophones. Néanmoins, il n'est pas prévu de prendre en compte des formes de handicap dans le jeu : malvoyance, surdité ou dimension « dys ». **Le jeu propose-t-il de faire des liens entre le monde réel et sa diégèse (effet-V) ?** Non. Certains participants ont demandé qu'il soit proposé, à l'issue du jeu, de faire une activité similaire en grandeur nature, dans une véritable pièce de l'entreprise.
Débriefing de l'activité	**Un bilan est-il proposé au joueur, tant sur l'aspect jeu que sur les aspects utilitaires ?** Oui, l'application propose en fin de partie de passer en revue les anomalies non identifiées, et de montrer où elles se situaient. Il est noté des apprentissages issus du jeu : la hauteur d'empilement de palettes ne doit pas excéder 1,80 mètre.

IV. ANALYSE RÉFLEXIVE AVEC CEPAJe

À présent que le tableau 2 est renseigné, il importe d'aller au-delà des retours recensés auprès des participants, en opérant une analyse réflexive. Pour ce faire, dans un premier temps, il s'agit de recenser, pour l'ensemble des trois temps pédagogiques, les principales forces et faiblesses associées à chaque dimension. Et dans un second temps, de proposer le cas échéant des solutions adaptées, ou d'en tirer des enseignements.

Pour le critère « contexte », le principal élément capable de perturber l'activité est le fait que certains participants soient tiraillés par des dossiers ou projets à conduire en parallèle de la journée sécurité. La mise en tension vécue peut empêcher d'être pleinement impliqué dans l'activité de jeu sérieux proposée. Une solution pourrait être d'inviter, en amont, les managers des collaborateurs concernés à prendre en compte la journée sécurité dans

leurs plannings de production, et à aménager la charge de travail en conséquence. Peut-être conviendrait-il aussi de responsabiliser les managers ou collaborateurs récalcitrants sur les risques dans le domaine HSE, si les journées sécurité sont négligées ?

Pour le critère « enseignant », la présence de deux médiateurs, un animateur et un manager, semble avoir donné de bons résultats auprès des participants pendant l'animation de jeu et le débriefing. Cette situation a permis aux médiateurs de répondre à des questions spécifiques en lien avec la sécurité sur site. Il convient par conséquent de conserver cette piste pour de prochaines séances ludopédagogiques. Néanmoins, en termes de prérequis, il aurait été bienvenu de fournir au manager un livret pédagogique et une formation idoine, afin de lui permettre d'être encore plus opérationnel durant ses premiers briefings. Quant à l'animateur d'Immersive Factory, il aurait été judicieux qu'il puisse bénéficier d'une visite guidée du site industriel, pour mieux intégrer les aspects sécurité à prendre en compte durant la séance ludopédagogique.

Pour le critère « pédagogie », le scénario repose sur la médiation humaine pour maximiser les chances d'atteindre les objectifs utilitaires. Cependant, le formateur n'a pas la latitude pour paramétrer la nature des risques à faire travailler aux participants. Offrir un tel paramétrage dans le serious game permettrait de gagner en précision dans la mise en place du scénario pédagogique.

Pour le critère « apprenant », il ressort que seules des motivations d'ordre extrinsèque sont suscitées par le serious game et la modalité VR. Cela nous invite à réfléchir sur des stratégies à mettre en place pour faire en sorte que les utilisateurs puissent basculer de motivations extrinsèques à intrinsèques, quant à la prise en compte des risques dans le cadre du site industriel. Une piste évoquée est de proposer, en aval de la séance ludopédagogique, une version grandeur nature de la chasse aux risques sur site. Cette dernière pourrait être préparée par des salariés afin de les engager davantage dans la participation aux journées sécurité.

Pour le critère « jeu », des corrections sont à apporter sur le plan ergonomique (meilleure gestion de boutons de manettes et de l'appareil photo) et des contenus (anomalies à améliorer). Surtout, on constate une nécessité de s'interroger sur un gameplay avec chronomètre. Si proposer d'atteindre les objectifs du jeu dans un temps donné peut constituer un levier motivationnel engageant et un défi clair pour l'utilisateur, il faut cependant réfléchir à la mise en place d'autres types de gameplays, pour éviter le stress temporel qui nuit à la bonne perception de certains risques durant la partie… Les participants ont aussi ouvert la voie à plus d'interactions, en proposant par exemple de conduire des véhicules. Cette piste semble pertinente à explorer.

CONCLUSION

Dans le cadre de l'exemple étudié, nous avons pu illustrer une manière d'utiliser le modèle CEPAJe pour recenser et trier des retours liés à une séance ludopédagogique dans le but de l'évaluer. L'approche consistait ici à tenir compte des principales problématiques ou des progrès notables, dimension par dimension. Cette approche est une possibilité parmi d'autres. Nous aurions pu adopter une approche visant à lister et prioriser les problèmes rencontrés. Nous aurions également pu croiser les éléments issus des différentes dimensions pour établir des corrélations. Une autre possibilité aurait été de nous focaliser sur les différents temps pédagogiques. Ainsi, au même titre que CEPAJe est un outil personnalisable, son usage l'est également. Il appartient à chaque utilisateur de trouver comment l'exploiter au mieux, selon ses besoins. Les déclinaisons proposées par Fenaert, Motte, Vangrunderbeeck ou encore Bouyssi témoignent d'appropriations variées, qui induisent des usages multiples répondant à des besoins précis.

Même si le modèle CEPAJe peut faire l'objet d'évolutions dans sa structure et ses usages, l'objectif principal reste de fournir aux ludopédagogues de quoi opérer une analyse réflexive de leurs pratiques ludopédagogiques, et améliorer, par les itérations, leurs prochaines séances. L'exemple proposé témoigne de cette possibilité. En parallèle, l'approche évaluative proposée met en relief la complexité et l'exigence d'une séance ludopédagogique. On comprendra à présent qu'adresser des évaluations aux seuls apprenants est réducteur. En effet, cela ne concerne qu'une seule dimension. Il est possible que les apprenants aient été victimes d'éléments de contexte, d'un médiateur malhabile, ou d'un jeu aux feed-back ou objectifs abscons.

Mais que dire de CEPAJe lui-même ? Est-il efficient ? Pour répondre à ces questions, il conviendrait d'étudier quelles déclinaisons du modèle s'avèrent les plus à même d'être utilisées par les ludopédagogues pour améliorer leurs séances et pratiques ludopédagogiques. Une telle approche pourrait s'opérer concrètement via la mise en place d'indicateurs de performance associés à l'emploi de différentes versions de CEPAJe. L'exemple sur la prévention des risques en entreprise constitue une bonne illustration. Dans ce cas, l'indicateur de performance pourrait correspondre à une approche statistique basée sur le nombre d'accidents annuels. L'expérimentation visée serait alors, dans un premier temps, de vérifier si ce taux diminue significativement sur les sites où les séances ludopédagogiques associent CEPAJe. Puis, dans un second temps, nous pourrions comparer quelles déclinaisons de CEPAJe apportent les meilleurs scores. Nous disposerions ainsi d'une mise en abîme où les outils d'évaluation seraient eux-mêmes évalués pour en mesurer la robustesse.

Remerciements

Merci à Olivier Pierre, Bertrand Pierre, Olivier Chabiron, Nicolas Biais et Aurélie Popelier de la société Immersive Factory pour leurs relectures attentives et conseils.

Bibliographie

J. Alvarez, L. Druette, G. Melia & P. Staccini (2016), « Adaptation du modèle CEPAJe destiné à évaluer une activité ludopédagogique pour le domaine de la Santé », SEGAMED 2016, Proceedings, SEGAMED, NICE, (FRANCE), 6 pages.

J.Alvarez & P. Chaumette (2017). « Présentation d'un modèle dédié à l'évaluation d'activités ludo-pédagogiques et retours d'expériences », *Recherche et pratiques pédagogiques en langues de spécialité,* Vol 36:2, LYON, (FRANCE), 15 pages.

F. Bouyssi (2018). « Démarche d'intégration d'un serious game, ou jeu sérieux, dans une séquence pédagogique ».

S. De Freitas & M. Oliver (2006). « How Can Exploratory Learning with Games and Simulations Within the Curriculum Be Most Effectively Evaluated? », *Computers and Education,* vol. 46, n° 3, pp.249-264.

M. Fenaert (2021). « Évaluer un jeu d'évasion virtuel », S'CAPE, https://scape.enepe.fr/evaluer-jeu-virtuel.html (consulté le 4 avril 2023).

P. Laudati & S. Leleu-Merviel (2018). « De l'UXD (User eXperience Design) au LivXD (Living eXperience Design) : vers le concept d'expériences de vie et leur design » in S. Leleu-Merviel, P. Useille & D. Schmitt (2018). *De l'UXD au LivXD, le design des expériences de vie,* ISTE Editions, pp.253-278.

I. Motte & P. Vangrunderbeeck (2018 / 2022). « Concevoir et animer un escape game pédagogique », université catholique de Louvain, LOUVAIN-LA-NEUVE, (BELGIQUE).

N. Tremblay (2007). « Formation initiale des enseignants, médiation pédagogique et approche philosophique », in N. TREMBLAY (dir.), *Des pratiques philosophiques en communauté de recherche en France et au Québec,* Presses de l'université Laval (PUL), LAVAL, (CANADA), pp. 95-116.

Les biomarqueurs numériques sont-ils des outils puissants et agnostiques d'évaluation des dispositifs pédagogiques numériques ?

Théo Marchand, Laurence Vanin, Vanessa Douet-Vannucci

Résumé

Avec l'avènement des technologies numériques, les biomarqueurs numériques émergent. Il s'agit d'une nouvelle catégorie de données collectables massivement à partir des dispositifs numériques pédagogiques. Ils sont porteurs d'informations inédites, inférant les comportements et les processus cognitifs. Depuis quelques décennies, l'industrie du jeu vidéo a développé des méthodologies d'évaluation spécifiques pour maximiser l'expérience utilisateur. Une approche hybride d'évaluation voit aujourd'hui le jour. Elle combine jeu vidéo et biomarqueurs numériques, afin d'examiner l'usabilité des contenus pédagogiques, leur pertinence, ainsi que leur efficacité sur les performances de l'utilisateur.

INTRODUCTION

Depuis la fin du XXe siècle, les politiques éducatives convergent vers une plus grande utilisation des technologies numériques, malgré des particularités historiques, pour améliorer l'efficacité des processus d'apprentissage, et les rendre plus personnalisés. Cependant, l'intégration du numérique dans les organisations a fortement été freinée par la perception des risques qu'il représente. Conformément à l'incitation des pays de l'Union européenne (UE) et de l'Organisation de coopération et de développement économiques (OCDE), la France a lancé en 1997 un plan afin de raccorder les établissements scolaires à Internet, et d'ancrer de manière institutionnelle l'informatique en milieu scolaire. Ce plan, intitulé « nouvelles technologies de l'information et de la communication dans l'enseignement », a permis d'amorcer la transformation numérique du système éducatif français.

Cette transformation s'est accélérée avec la pandémie COVID-19, et la démocratisation massive de l'utilisation des appareils numériques ainsi que des plateformes d'apprentissage en ligne pour l'éducation. Pour continuer à enseigner à distance, les écoles, universités et autres établissements d'enseignement ont dû s'adapter rapidement à l'utilisation d'outils numériques, tels que les vidéoconférences, les classes virtuelles et les manuels numériques. Pour autant, ce brutal passage à l'apprentissage à distance a finalement été perçu comme facilitateur de la communication, de la collaboration et de l'enseignement.

Aujourd'hui, le système éducatif continue sa numérisation à un rythme rapide. Celle-ci est poussée par l'expansion croissante des technologies numériques, comme Internet, qui s'immiscent dans tous les secteurs de notre société, transformant inéluctablement les relations entre les êtres et les objets. La plupart des étudiants et des enseignants considèrent désormais l'éducation et son accès comme étroitement liés aux technologies et dispositifs numériques. Les ordinateurs, les tablettes et les smartphones deviennent des outils pédagogiques complètement intégrés dans les processus d'enseignement et d'apprentissage. De même, les programmes d'enseignement, les outils collaboratifs en ligne simulant les conditions d'enseignement en groupe ou de classe, et les MOOC (Massive Open Online Course) sont de plus en plus populaires. Tous ces dispositifs numériques pédagogiques présentent des opportunités et avantages non négligeables, comme la mutualisation des connaissances.

Les « techno-prophètes » présageaient que la numérisation aiderait à améliorer l'éducation des futures générations en facilitant l'accès et la flexibilité de l'apprentissage. Or, un individu sur six est en situation d'illectronisme, c'est-à-dire en incapacité ou en difficulté pour utiliser les outils numériques, par ignorance totale ou partielle de leur fonctionnement. Cela n'épargne pas les populations nées avec le numérique, laissant entrevoir une fracture

éducative et sociale forte. Cette condition est aggravée si on considère que la plupart des utilisateurs méconnaissent les algorithmes qui les régissent et sous-tendent l'ensemble des interactions, qui ne sont connus que de l'ingénieur sachant. La croissance exponentielle des dispositifs pédagogiques numériques pose alors la question de leur évaluation attentive, validée et protocolisée par les pairs.

Ces différentes évolutions des systèmes éducationnels et pédagogiques imposent donc d'identifier et de caractériser de nouveaux indicateurs et/ou métriques, afin de construire une méthodologie d'évaluation de ces dispositifs pédagogiques. Après une introduction sur les dispositifs pédagogiques numériques, cet article s'attache à décrire les potentialités d'une approche nouvelle, hybridant les méthodologies d'évaluation établies par le monde du jeu vidéo, et les biomarqueurs numériques pour l'évaluation de ces dispositifs.

I. LES DISPOSITIFS PÉDAGOGIQUES NUMÉRIQUES

Les technologies numériques s'invitent dans tous les environnements de la société. L'enfant d'aujourd'hui naît avec les appareils numériques, et son empreinte numérique est si importante que l'individu peut être considéré comme vivant désormais en quasi-symbiose avec ces technologies. Aussi, les systèmes éducatifs se sont attachés à intégrer les utilisations de dispositifs numériques afin de favoriser la pédagogie et l'apprentissage des plus jeunes (loi d'orientation du 08/07/2013). Quelques exemples de dispositifs parmi les plus usités sont proposés dans le tableau 1 ci-après.

TABLEAU 1 — Fonctionnalités des dispositifs pédagogiques numériques

Type de dispositif pédagogique numérique	Descriptif	Exemples
Systèmes de gestion d'apprentissage / Learning machine systems (LMS) (Bradley, 2021)	Créer et partager des supports de cours, des devoirs et des évaluations en ligne avec les étudiants.	Sakai, Moodle
« Jeux sérieux / serious games », environnements / plateformes d'apprentissage (Laamarti, 2014)	Transmettre des informations, développer des compétences ou sensibiliser les joueurs à des thèmes importants, tout en offrant une expérience ludique et engageante. À rapprocher des environnements et plateformes d'apprentissage par le jeu.	Scratch, BrainPOP
Tableaux blancs interactifs (Drigas, 2014)	Afficher du contenu multimédia et annoter des documents pendant le cours.	Sharp, Samsung
Applications de réalité virtuelle et augmentée (Kavanagh, 2017)	Créer des expériences d'apprentissage immersives et des simulations. La réalité virtuelle (VR) peut être utilisée pour créer des visualisations 3D interactives de concepts scientifiques et mathématiques, ce qui les rend plus attrayants et accessibles aux apprenants.	Go For Launch: Mercury, Toti Submarine-VR Experience
Systèmes de tutorat basés sur l'intelligence artificielle (Steenbergen-Hu, 2014)	Créer des expériences d'apprentissage, fournir un retour et un soutien personnalisés aux étudiants en fonction de leurs performances.	MathSpring

L'intégration du numérique dans l'éducation présente plusieurs avantages, dont l'accessibilité à l'information, et la flexibilité de l'enseignement. Les apprenants peuvent en effet adapter leur enseignement à leur rythme, à leurs besoins, ou au contexte géographique (urbain versus rural), social, et personnel (situation de handicap).

Cette adaptabilité se retrouve également au niveau de la communication et de la collaboration à plusieurs niveaux, qui s'auto-enrichissent : inter-apprenants et inter-enseignants, mais également dans une relation mixte entre apprenants et enseignants. Il s'agira alors de repenser les rôles de chacun de ces acteurs, et de transformer l'environnement d'apprentissage.

De plus, comme les nouvelles technologies permettent d'intégrer les stratégies d'apprentissage individuelles, pour le développement personnalisé des connaissances et l'apprentissage autodirigé, il faudra également stimuler et accélérer les systèmes éducatifs, qui se développent naturellement plus rationnellement et lentement.

La technologie numérique encourage l'interactivité et l'engagement de l'utilisateur, grâce à l'ajout de principes de jeu, de multimédia, et de technologies toujours plus attrayantes, telles que la réalité virtuelle ou le métaverse (création de mondes virtuels sur Internet), ayant un impact significatif sur la pédagogie. Cette interactivité, induisant une attention soutenue de l'utilisateur, se révèle attractive et ludique.

Cependant, les bénéfices qu'apportent le numérique et les nouvelles interfaces aux dispositifs pédagogiques modifient les limites sémantiques de la définition du dispositif pédagogique versus dispositif de divertissement. Il est donc important de trouver et de protocoliser la ou les méthodes d'évaluation critique des dispositifs numériques pédagogiques, afin de réaliser cette transition numérique.

II. LES MÉTHODOLOGIES D'ÉVALUATION DU MONDE DU JEU VIDÉO

L'industrie du jeu vidéo analyse le ressenti de l'utilisateur face à un jeu pour améliorer celui-ci. L'expérience utilisateur (UX) et l'interface utilisateur (UI) sont utilisées pour mesurer le niveau attentionnel, l'engagement, la frustration des joueurs pendant qu'ils performent. L'UX se focalise sur la manière dont les joueurs interagissent avec le jeu, tandis que l'UI se concentre sur la manière dont les joueurs interagissent avec l'interface du jeu. L'UX et l'UI impliquent une ouverture phénoménologique à la captation des postures de l'utilisateur, qui devient sujet et objet de ses relations aux interfaces.

L'UI est donc une composante de l'UX générale. Elle étudie toute interaction de l'utilisateur avec l'interface pour la rendre agréable, pratique, voire source d'émotion. Cette interface est l'environnement graphique dans lequel évolue l'utilisateur d'un logiciel, d'un site web ou d'une application. Son évaluation fait appel à l'UI design, qui regroupe un ensemble d'éléments visuels : les boutons, les textes, les images, les couleurs, la typographie, le défilement des pages, les animations. Les développeurs mesurent ainsi la conception des applications, le design des interfaces, et leurs impacts sur l'engagement ou la frustration suscités, en examinant le temps passé sur chaque écran, les interactions avec les boutons et les menus, et les erreurs d'utilisation.

L'UX s'évalue par des méthodes d'analyse quantitative, qui nécessitent des connaissances en statistiques et des méthodologies scientifiques (protocole, échantillonnage, représentativité, biais, validité des outils de mesure, interprétation, etc.). Les méthodes statistiques sont alors utilisées pour décrire les phénomènes observés, et faire des estimations sur la population totale. Elles combinent ainsi « recherche et études de terrain » (observations, puis définition de groupes de types d'utilisateurs cibles, appelés les personae) avec « conception et évaluation » (tests utilisateurs en laboratoire, réalisés sur un petit échantillon, réitération).

Dans un jeu vidéo, cette analyse quantitative concerne les actions des joueurs, comme le nombre de clics, les mouvements de la souris, les entrées de clavier, les temps de réalisation des tâches. À ces indicateurs quantitatifs s'ajoutent des mesures qualitatives d'utilisabilité, d'utilité, d'esthétique, qui varient selon le type de dispositif concerné, et les critères ou ressentis humains (usages, émotions, pensées, etc.). Ces aspects hédoniques et pragmatiques des systèmes interactifs sont principalement évalués grâce à des questionnaires standards tels que AttrakDiff[46]. Les concepteurs peuvent donc mesurer la vitesse à laquelle les joueurs progressent dans le jeu, la durée des sessions de jeu, la difficulté perçue ou le degré d'intérêt, afin d'identifier les moments cruciaux en rapport avec l'usabilité et l'acceptation des apprenants. Les données, récoltées directement dans le jeu, servent aussi à examiner l'engagement, l'état attentionnel, la frustration, et la satisfaction des joueurs à différents moments de leur expérience de jeu (tableau 2).

46. Méthode de mesure de l'utilisabilité d'un système interactif basée sur la perception de l'utilisateur.

TABLEAU 2 — Les différentes métriques d'évaluation lors de la conception et du développement de jeux vidéo

Métriques	Descriptif	Références
Temps de jeu	• Durée pendant laquelle un joueur joue • Mesure l'engagement des joueurs	(Boyle et al., 2012)
Taux de rétention	• Proportion de joueurs qui continuent à jouer au jeu après un certain temps • Mesure l'amusement et l'engagement des joueurs Identification des moments où les joueurs ne sont plus engagés ou perdent l'intérêt	(Callaghan et al., 2014)
Score	• Données sur le niveau de performance (global ou spécifique) et/ou la progression d'un joueur dans le jeu • Lié à la concentration, à la motivation, aux compétences du joueur, à la vitesse de traitement, à la fatigue mentale…	(Boot et al., 2008)
Taux de complétion	• Nombre de joueurs qui ont terminé le jeu par rapport au nombre total de joueurs • Mesure l'engagement des joueurs, la difficulté du jeu	(McAllister et al., 2015)

Les aspects « émotionnels » sont une composante forte dans l'évaluation d'un jeu vidéo. Ils trouvent des métriques spécifiques comme l'analyse des expressions faciales et gestuelles, ou techniques psychophysiologiques, de type EEG (électroencéphalographique), rythme cardiaque, ou suivi des mouvements oculaires. Avec la démocratisation des nouvelles technologies, les réponses physiologiques sont devenues un moyen d'obtenir des mesures objectivées de l'expérience utilisateur avec une technologie de divertissement.

Avec l'avènement des réseaux sociaux et des jeux multijoueurs, une nouvelle dimension se retrouve dans le jeu : la composante sociale. Les joueurs peuvent parler à des partenaires de jeu, à des amis via des plateformes dédiées, des groupes, des blogs (Reddit, Discord, Steam, YouTube, Twitch ou Facebook) selon les besoins spécifiques des communautés de joueurs. Lors d'entretiens, ces interactions ont été attentivement étudiées par les concepteurs de jeux pour moduler les fonctionnalités et spécificités de jeu. Les discussions et commentaires deviennent des éléments de mesure de performances de jeu.

La méthodologie d'évaluation séquentielle du monde du jeu vidéo est donc un atout de l'évaluation de dispositifs pédagogiques numériques, où chaque moment pédagogique clé, les données de progression, de temps, de score et de difficulté, sont un reflet de la performance de l'apprenant. Les méthodes qualitatives et quantitatives complémentaires s'adressent alors autant au « pourquoi » qu'au « comment » de l'interaction, et permettent d'appréhender cette dernière dans sa globalité. Cela contribue à l'amélioration permanente des dispositifs, en fonction des données d'usage récoltées, mais aussi des perceptions associées, qui reposent sur des éléments subjectifs souvent difficiles à identifier.

La combinaison de ces méthodologies donne une image globale de l'expérience utilisateur transposable à l'évaluation des dispositifs pédagogiques numériques. Le parcours de l'apprenant est analysé ainsi étape par étape. On analyse le comportement de l'individu, et le cheminement qu'il adopte pour réaliser une action donnée, ces résultats permettant d'améliorer la qualité de l'expérience utilisateur et, par conséquent, d'augmenter l'efficacité de l'apprentissage.

Cependant, une différence structurelle existe entre les motivations des utilisateurs jouant aux jeux vidéo et celles des apprenants utilisant des dispositifs pédagogiques. Une adaptation des méthodologies de l'industrie du jeu sera nécessaire pour être efficient dans les dispositifs pédagogiques, avec des études rigoureuses, notamment sous le prisme de tests d'usabilité, pour identifier les fonctions transférables.

III. LES BIOMARQUEURS NUMÉRIQUES

Les biomarqueurs sont des indicateurs mesurables de l'activité cérébrale, de la physiologie ou de l'émotion, qui peuvent être utilisés pour évaluer les dispositifs pédagogiques. Ces biomarqueurs ont été identifiés et caractérisés selon différents types de technologies, comme le suivi des mouvements oculaires, l'EEG, l'imagerie de résonance magnétique fonctionnelle (IRMf), la mesure de la fréquence cardiaque, ou encore les capteurs. Les biomarqueurs numériques proviennent de la collecte d'informations depuis des capteurs, associée à des outils de calcul comme les algorithmes sur plusieurs couches d'outils matériels et de logiciels connectés. De cette nouvelle façon de collecter de l'information est née l'épidémiologie numérique. Elle désigne l'utilisation des technologies et des données numériques pour, d'une part, étudier le fonctionnement des individus et les facteurs de vulnérabilité au sein d'une population, et d'autre part, élaborer des interventions visant à prévenir ou à contrôler ces vulnérabilités.

Les biomarqueurs numériques permettent le phénotypage numérique selon deux techniques complémentaires de collecte de données. Le phénotypage numérique est dit « passif » lorsque les données sont engendrées simplement par les activités quotidiennes normales de l'utilisateur, ou par l'utilisation de l'appareil. La personne n'est pas obligée de faire quoi que ce soit. Lorsque la personne doit accomplir une tâche, par exemple répondre à une enquête ou réaliser une activité, le phénotypage est dit « actif ». Les données engendrées par les activités effectuées sur des smartphones ou d'autres appareils numériques peuvent être analysées pour comprendre les modèles de parole, les rythmes circadiens, les mouvements oculaires, les expressions faciales, l'activité du clavier, l'auto-évaluation, la vitesse de traitement et les réponses aux stimuli (tableau 3).

TABLEAU 3 — Description des différents types de biomarqueurs digitaux

Métriques	**Descriptif**	**Références**
Les appareils connectés	• Mesure l'activité quotidienne de l'usager de façon passive ou active, pour améliorer l'expérience et la personnalisation de l'apprenant • Mesure des changements dans les marqueurs physiologiques tels que la fréquence cardiaque, la conductance de la peau et la respiration, pendant les tâches d'apprentissage, pour évaluer les réponses physiologiques de l'apprenant, ou l'état de stress ou de relaxation	Dagum et al. 2018 ; R Wang et al. 2018
Suivi oculaire	• Mesure les trajectoires, les points de fixation et temps de fixation du regard de l'utilisateur lors d'une tâche • Identifier automatiquement les événements d'échange entre l'interface et l'utilisateur, afin d'évaluer de façon systématisée les points d'attention et d'intérêt • Évaluer les éléments qui causent de la confusion ou de la distraction pour les utilisateurs, et donc mesurer les changements dans l'attention visuelle et les processus cognitifs associés à la lecture et à l'apprentissage • Examiner l'efficacité de différents matériels de lecture et stratégies d'enseignement pour les enfants	Kruger and Steyn, 2014 ; Vajs et al. 2022 ; Chaudhuri et al. 2022

Suivi du doigt	• Identifier les actions des utilisateurs qui sont courantes, et celles qui sont moins courantes dans l'interaction avec un écran tactile • Corréler les trajectoires digitales avec l'état émotionnel de l'apprenant, autant qu'automatiser les méthodes de quotation des performances	Lio et al. 2019
Électroencéphalographie	• Identifier des réactions émotionnelles et cognitives des utilisateurs, pour une meilleure compréhension des ressentis des utilisateurs aux différents éléments de l'interface, comme les textes, les images et les vidéos • Analyse des émotions et de la charge mentale pour évaluer un dispositif	Pi et al. 2021 ; AlShorman et al. 2020

Ces données ont été très utilisées par le marketing de précision et les sciences cognitives pour mieux étudier et comprendre les comportements des consommateurs. Aujourd'hui, l'épidémiologie numérique arrive à fournir des indications prédictives des comportements humains comme des processus cognitifs, tels que la mémoire, l'attention, le raisonnement. Elles permettent de mieux appréhender les potentialités individuelles, et de rendre compte de logiques algorithmiques qui traduisent les comportements humains.

Ces biomarqueurs numériques appliqués à l'évaluation des dispositifs pédagogiques pourraient devenir de véritables indicateurs de l'engagement dans l'activité, de l'état de compréhension des apprenants, mais également de leur état mental et psychologique. Avec la puissance de calcul et de modélisation des dernières technologies, le résultat des performances peut être obtenu en temps quasi réel ; et permettre ainsi une modification corrective immédiate de l'enseignant sur sa méthodologie pédagogique, et/ou d'utilisation de son dispositif auprès des apprenants. Enfin, ces biomarqueurs sont considérés comme objectifs, puisqu'ils sont le reflet mesurable des comportements non conscientisés de l'apprenant lors de l'utilisation d'un dispositif pédagogique numérique. De plus, ils sont agnostiques de la méthode d'enseignement. Ils ouvriraient ainsi la voie à une évaluation comparative de plusieurs dispositifs pédagogiques, quels que soient leurs type, public, etc.

Cette méthodologie d'évaluation faciliterait l'inférence de processus cognitifs, conduisant à la personnalisation d'un dispositif pédagogique, grâce à la numérisation de contenus modulaires, à la flexibilité des stratégies d'apprentissage, et aux retours d'expérience adaptés aux besoins de chaque

apprenant. Cette personnalisation, à travers le numérique, vise à adapter la distribution, la fréquence des contenus, et les méthodes d'enseignement, pour répondre aux besoins individuels des apprenants en maintenant un parcours d'apprentissage commun. La personnalisation ne doit pas être confondue avec l'individualisation, qui procure à chaque élève un parcours d'apprentissage totalement distinct. Cette attention, portée sur le déroulement temporel et situé des activités, fait apparaître des phénomènes qui, traditionnellement, échappent aux recherches en sciences humaines, fréquemment plus globalisantes.

Les biomarqueurs numériques permettraient donc d'observer et de comprendre la façon dont les concepteurs scénarisent les pratiques d'évaluation avec la technologie, ainsi que la manière dont les usagers s'en saisissent et se les approprient. L'instrumentation des processus d'évaluation répond à des besoins dans différents contextes sommatifs comme formatifs. Les pratiques se stabilisent progressivement, alors que la recherche dans ce domaine n'en est qu'à ses prémisses.

CONCLUSION

Le processus d'évaluation est une démarche visant à mesurer, quantifier ou caractériser une action, des savoirs ou savoir-faire souvent complexes. De définition variable selon les domaines, il est souvent considéré comme une pratique pédagogique entre un apprenant et son enseignant. Il s'identifie par une séquence question/réponse/évaluation aux dimensions socio-psychologiques.

La finalité de la méthodologie d'évaluation des dispositifs pédagogiques numériques n'est pas de substituer totalement une technologie à l'évaluateur, mais de le soutenir dans sa démarche pédagogique et d'évaluation de l'apprenant, pour mieux appréhender les capacités de ce dernier, et élaborer des stratégies plus adaptées, voire personnalisées. Les méthodologies issues du monde des jeux vidéo et les biomarqueurs numériques sont une brique complémentaire pour affiner ce processus d'évaluation.

L'évaluation des apprentissages assistée par ordinateur deviendrait plus efficiente, pratique, objective, et accessible. Mais, elle s'inscrit encore trop souvent dans la continuité de pratiques existantes. Elle est donc limitée à la reproduction de connaissances déclaratives et de procédures fixes. Aujourd'hui encore, le travail d'évaluation d'apprenants est un processus cognitif, interactionnel, social et pédagogique complexe, qui implique les deux parties pour une efficacité maximale. L'instrumentation de ce travail par les technologies numériques impose une rupture dans les pratiques d'évalua-

tion, qui ne pourrait être ultimement acceptée et adoptée que si un travail de collaboration se tisse entre concepteurs et acteurs de l'éducation. Dans cette boucle vertueuse, une appréciation de l'appropriation des technologies par les apprenants est indispensable pour maximiser leurs capacités et méthodes de travail.

Enfin, la numérisation engendre un accès et une diffusion ultra-rapides des informations, voire leur production automatisée. C'est le cas notamment avec ChatGPT. L'accumulation des connaissances de l'apprenant et leur restitution ne semblent plus aussi cruciales qu'autrefois. Les capacités des apprenants doivent désormais résider dans leur aptitude à rechercher des informations, et à développer leur esprit critique face à cette somme croissante d'informations. L'évaluation des apprentissages ne pourra donc plus se limiter à mesurer la capacité des apprenants à mémoriser et reproduire des informations. Elle devra désormais s'aligner sur les nouveaux objectifs d'une éducation préparant les futures générations à affronter les défis d'un monde globalisé et connecté.

Immergées au cœur du processus d'innovation des technologies d'évaluations pédagogiques, l'expérimentation et la méthodologie d'observations restent indispensables pour étudier pleinement les interactions entre tous les acteurs, et révéler les jalons de développement. L'immersion sur le terrain et la constitution progressive de réseaux sociotechniques joueront un rôle déterminant d'évaluation pertinente et complète, pour une meilleure acceptation des nouveaux dispositifs pédagogiques. En bref, le numérique bouleverse et transforme le rôle, les rituels et la forme des évaluations, dans leur dimension à la fois institutionnelle, économique et sociale.

Bibliographie

Aude Inaudi (2017). « École et numérique : une histoire pour préparer demain. », in *Hermès,* 2, pp. 72-79.

Christine Barats (2006). « Les mythes du supérieur à l'heure des TIC, Analyse de la rhétorique ministérielle pour l'intégration des TIC (négociation, promesse et avantages) », in *Enjeux et usages des TIC*. Reliance sociale et insertion professionnelle.

Shivangi Dhawan (2020). « Online Learning: A Panacea in the Time of COVID-19 Crisis », in *Journal of Educational Technology Systems,* 49(1), p.5–22.

Christopher Brinton et al. (2016). « Mining MOOC clickstreams: Video-watching behavior vs. in-video quiz performance. » *in IEEE Transactions on Signal Processing,* 64(14) : 3677-3692.

Larry Cuban, Petar Jandrić (2015). « The dubious promise of educational technologies: Historical patterns and future challenges. », in *E-learning and Digital Media,* 12(3-4), pp.425-439.

Stéphane Legleye, Annaïck Rolland (2019). « Une personne sur six n'utilise pas Inter-

net, plus d'un usager sur trois manque de compétences numériques de base », in *Insee Première,* n° 1780, octobre.

Donell Holloway, Lelia Green, Sonia Livingstone (2013). « Zero to eight : Young children and their internet use. ».

Lydia Plowman (2015). « Researching young children's everyday uses of technology in the family home », in *Interacting with Computers,* 27(1), pp.36-46.

Laurent Tessier (2019), *Éduquer au numérique ? Un changement de paradigme.* MkF Éditions.

Steven Higgins, ZhiMin Xiao, Maria Katsipataki (2012). « The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation. Full Report. », in *Education Endowment Foundation.*

Carine Lallemand, Guillaume Gronier, Vincent Koenig (2013). « L'expérience utilisateur : un concept sans consensus ? Enquête sur le point de vue des professionnels. », in *Activités humaines, Technologies et Bien-être.*

Jeff Sauro, James R. Lewis (2016). *Quantifying the user experience: Practical statistics for user research.* Morgan Kaufmann.

Marc Hassenzahl, Kai Eckoldt, Sarah Diefenbach, *et al.* (2013), « Designing moments of meaning and pleasure. Experience design and happiness. », in *International journal of design,* 7(3), pp.21-31.

Elizabeth A. Boyle, et al. (2012), « Engagement in digital entertainment games: A systematic review » in *Computers in human behavior,* 28.3, pp.771-780.

Michael James Callaghan, Niall McShane, A. Gomez Eguiluz (2014). « Using game analytics to measure student engagement/retention for engineering education. », in *11th International Conference on Remote Engineering and Virtual Instrumentation (REV),* IEEE, 2014, pp.297-302.

Walter R. Boot and al (2008). « The effects of video game playing on attention, memory, and executive control», in *Acta psychologica,* 129(3), pp.387-398.

Graham McAllister, Gareth R. White (2015). « Video game development and user experience. », in *Game user experience evaluation,* pp.11-35.

Regan L Mandryk, Kori M. Inkpen, Thomas W. Calvert (2006). « Using psychophysiological techniques to measure user experience with entertainment technologies. », in *Behaviour & information technology,* 25(2), pp.141-158.

Christian Montag, Daniel S. Quintana (2023). « Digital phenotyping in molecular psychiatry—a missed opportunity?. », in *Molecular Psychiatry,* 28.1, pp.6-9.

Paul Dagum (2018). « Digital biomarkers of cognitive function. », in *NPJ digital medicine* 1(1), p10.

Rui Wang and al (2018). « Tracking depression dynamics in college students using mobile phone and wearable sensing. », in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies,* 2(1), pp.1-26.

Jan-Louis Kruger, Faans Steyn (2014). « Subtitles and eye tracking: Reading and performance. », in *Reading Research Quarterly,* 49(1), pp.105-120.

Ivan Vajs, and al (2022), « Spatiotemporal eye-tracking feature set for improved recognition of dyslexic reading patterns in children. », in *Sensors* 22(13):4900.

Saswati Chaudhuri and al (2022), « Teachers' focus of attention in first-grade classrooms: exploring teachers experiencing less and more stress using mobile eye-tracking. », in *Scandinavian Journal of Educational Research*, 66(6), pp.1076-1092.

Guillaume Lio and al (2019). « Digit-tracking as a new tactile interface for visual perception analysis. », in *Nature Communications,* 10(1):5392.

Zhongling Pi and al (2021). « Learning by explaining to oneself and a peer enhances learners' theta and alpha oscillations while watching video lectures. », in *British Journal of Educational Technology,* 52(2), pp.659-679.

Omar AlShorman and al (2020). « The effects of emotional stress on learning and memory cognitive functions: an EEG review study in education. », in 2020 *Sixth International Conference on e-Learning* (econf), IEEE, pp. 177-182.

Ferdousi Sabera Rawnaque, Khandoker Mahmudur Rahman, Syed Ferhat Anwar and al. (2020). « Technological advancements and opportunities in Neuromarketing: a systematic review. », in *Brain Informatics,* 7, pp.1-19.

Shen Ba Xiao Hu (2023), « Measuring emotions in education using wearable devices: A systematic review. », in *Computers & Education,* pp.104797.

Thomas R Insel (2017), « Digital phenotyping: technology for a new science of behavior. », in *Jama* 318(13), pp.1215-1216.

Jukka-Pekka Onnela (2021). « Opportunities and challenges in the collection and analysis of digital phenotyping data. », in *Neuropsychopharmacology*, 46(1), pp.45-54.

Bernie Trilling, Charles Fadel (2009). *21st century skills: Learning for life in our times.* John Wiley & Sons.

PARTIE III

Les dispositifs

Création et évaluation d'un programme pédagogique inclusif basé sur l'activité physique et le jeu sportif à destination d'enfants ayant un trouble du spectre de l'autisme

Olivia Collet, Geneviève Cabagno, Sandrine Le Sourn-Bissaoui

Résumé

Alors même que l'activité physique semble constituer un moyen efficace pour améliorer les compétences sociales des enfants présentant un trouble du spectre de l'autisme (TSA), les professionnels ne proposent pas toujours les activités physiques avec cet objectif. Par ailleurs, l'évaluation de l'effet de telles activités est rarement étudiée. Cet article vise à présenter la démarche de cocréation d'un programme pédagogique inclusif, basé sur l'activité physique et le jeu sportif, à destination d'enfants présentant un TSA. Il s'agira par ailleurs d'évaluer l'effet de ce programme aussi bien pour les enfants avec un TSA que pour les enfants neurotypiques, leurs parents, et les professionnels impliqués dans le dispositif.

INTRODUCTION

La pratique d'activité physique, aujourd'hui reconnue comme bénéfique pour des personnes en situation de handicap (INSERM, 2019), reste peu étudiée dans le champ des sciences humaines en ce qui concerne le trouble du spectre de l'autisme (TSA). En effet, le TSA est un trouble du neurodéveloppement caractérisé par un déficit persistant des interactions sociales et de la communication, observé dans des contextes variés, ainsi que par des intérêts, activités et comportements restreints et répétitifs (DSM 5[47], 2013), pouvant conduire à une exclusion sociale. L'intensité de ces troubles varie d'une personne à l'autre, permettant de définir la sévérité du TSA allant du « niveau 1 » (nécessitant un soutien) au « niveau 3 » (nécessitant un soutien très important). Pourtant, pour cette population particulière, la Haute Autorité de santé considère que les activités physiques peuvent participer à l'épanouissement personnel et social, à la condition d'un accompagnement spécifique. Par ailleurs, la « Stratégie nationale pour l'autisme 2018-2022 » préconise de soutenir l'inclusion à tous les âges de la vie.

Selon la littérature scientifique centrée sur le TSA, la pratique d'activités physiques pourrait conduire à une diminution des comportements extériorisés (crises de colère, agressivité, automutilation) et stéréotypés (balancement, tournoiement, écholalie, etc.), ainsi qu'à une augmentation des comportements dirigés vers l'action (écoute les consignes, effectue ce qui lui est demandé, etc.). De plus, plusieurs études ont montré que la pratique d'activités physiques contribuerait aussi au développement des compétences sociales des enfants présentant ce trouble, et à leur développement social en général. En permettant notamment un développement des capacités communicatives, une amélioration de la participation sociale et de leurs relations interpersonnelles, l'activité physique induirait un effet positif sur les interactions sociales. Elle contribuerait ainsi à créer un cadre social optimal, favorisant les interactions, les relations entre participants et les occasions de communiquer, dans un environnement naturel. En effet, les situations de jeu et les activités sportives, par exemple celles de nature coopérative, offriraient des opportunités positives de socialisation à l'enfant, lui permettant d'interagir avec ses pairs, d'expérimenter différents rôles et situations sociales, de partager des intérêts communs, d'apprendre à respecter des règles et à coopérer.

Cependant, l'analyse de la littérature scientifique centrée sur le TSA et l'activité physique met en évidence que les études menées s'inscrivent, jusqu'à récemment, de manière prédominante dans des considérations intra-personnelles centrées de manière ciblée sur les compétences des enfants. Elles ne prennent pas suffisamment en compte son environnement social (pairs, parents) et l'expérience des professionnels engagés dans leur accompa-

47. Diagnostic and Statistical Manual of Mental Disorders, cinquième édition.

gnement. Or, une analyse uniquement intra-personnelle semble insuffisante pour évaluer l'impact d'un programme d'éducation par l'activité physique sur les compétences socio-comportementales d'enfants avec un TSA, dès lors que l'enjeu prioritaire repose sur l'inclusion de ces enfants avec des enfants neurotypiques. Cette inclusion repose sur l'adaptation du programme à l'ensemble des enfants, qui, de fait, leur permet de participer conjointement et de développer des interactions entre eux, nécessitant des compétences socio-comportementales. De plus, à notre connaissance, peu d'études se fixent comme objectif de tester réellement la pertinence et l'efficacité de ces programmes proposant de l'activité physique en situation d'inclusion, au regard d'autres modalités d'accompagnement, et sur plusieurs mois. Enfin, au regard de l'enquête menée auprès des professionnels accompagnant ce public spécifique, il semble que l'accompagnement proposé soit largement axé sur le développement général de l'enfant, et de façon plus marginale sur le développement d'habiletés sociales ciblées, reposant sur des objectifs d'apprentissage spécifiques en situation d'inclusion avec des enfants au développement typique et mobilisant de l'activité physique. Ceci pourrait expliquer en partie pourquoi les professionnels accompagnant ces enfants TSA n'évaluent pas (ou peu), de manière rigoureuse et la plus objective possible, les effets de leurs pratiques sur les compétences socio-comportementales des enfants.

Face à ces constats, une démarche de création d'un programme pédagogique inclusif basé sur l'activité physique et le jeu sportif semble pertinente, dans l'objectif de développer certaines habiletés sociales de manière ciblée chez des enfants ayant un TSA. Par ailleurs, il convient de mettre en place une procédure d'évaluation des impacts de ce programme pédagogique, sur les habiletés sociales d'une part, mais également sur l'inclusion des enfants d'autre part.

I. CONCEPTION D'UN PROGRAMME PÉDAGOGIQUE INCLUSIF INNOVANT

Afin de concevoir un programme pédagogique le plus pertinent possible visant à développer des habiletés sociales ciblées chez les enfants avec TSA, une analyse approfondie de la littérature consacrée aux habiletés sociales et à l'intérêt du jeu et de l'activité physique comme leviers d'apprentissage a été réalisée. De plus, un diagnostic, à la fois quantitatif et qualitatif, a été effectué auprès de professionnels accompagnant ce public spécifique. Les résultats obtenus ont montré que l'utilisation de l'activité physique semble instituée dans l'accompagnement de ce public à raison d'une fois par semaine minimum, mais que les professionnels qui la mobilisent sont rarement formés à la fois à l'activité physique et au TSA. De

même que le choix du type d'activité dépend principalement des objectifs généraux d'accompagnement fixés, ou des programmes scolaires, mais ils ne répondent pas toujours aux besoins spécifiques de chaque enfant. Bien que le type d'activité physique reste majoritairement individuel, les séances sont tout de même réalisées en groupe, mais rarement en situation d'inclusion avec des personnes au développement typique. Enfin, plusieurs difficultés rencontrées par les accompagnants ont été identifiées : difficultés liées aux problématiques du TSA, manque de personnel encadrant et de formation, ou encore difficultés d'accès aux locaux et au matériel. L'ensemble de ces éléments nous ont permis de définir les orientations pour la conception de notre programme pédagogique innovant.

Pour que celui-ci soit le plus adapté possible aux difficultés rencontrées par les enfants, et le plus pertinent possible au regard des objectifs poursuivis, une démarche participative a été adoptée. Un collectif d'acteurs se composant de quatre professionnelles accompagnant des enfants ayant un TSA, de deux expertes en didactique du sport, et d'une doctorante spécialiste en activités physiques adaptées a été constitué, avec supervision de deux experts universitaires. Cette démarche de conception, appelée « cocréation », permet, en intégrant différents acteurs, d'enrichir les connaissances et de partager des savoirs issus de différents secteurs, aboutissant ainsi à la mise au point de nouveaux outils utiles à l'intervention. Elle permet également la confrontation des intérêts de chacun, et facilite la recherche de consensus. De plus, en intégrant le point de vue de ces différents acteurs, elle permet de répondre au mieux à leurs besoins et à ceux des enfants qu'ils accompagnent.

1. DÉMARCHE DE COCRÉATION DU PROGRAMME PÉDAGOGIQUE

La construction du programme pédagogique inclusif a été effectuée au cours de six séances de cocréation d'une durée de quatre heures, à raison d'environ une après-midi par mois. L'objectif de la première séance vise à présenter l'objectif de la démarche de cocréation à l'ensemble des participantes, ainsi que son déroulement, qui s'est étendu sur cinq mois. Au cours de cette première séance, le programme inclusif à construire et les habiletés sociales faisant l'objet des apprentissages visés ont été présentés. Il s'agit de l'imitation et l'attention conjointe, la communication verbale et non verbale, et la gestion des émotions. Ces habiletés sociales ont été identifiées d'une part sur la base de l'analyse de la littérature scientifique, et d'autre part à l'aune des résultats obtenus avec le diagnostic précédemment réalisé, qui souligne que les intervenants considèrent ces habiletés comme particulièrement susceptibles d'être améliorées grâce à l'activité physique. Chacune de ces habiletés a ensuite fait l'objet d'une séance à part entière, ayant abouti à la création d'un jeu sportif spécifique. La cinquième séance de cocréation a pour but de créer une « situation de référence ». Cette dernière correspond à une séance de jeu sportif, qui nous conduira, au cours de la phase suivante de l'étude, à évaluer les effets du programme auprès des

enfants. Enfin, la sixième et dernière séance de cocréation a permis de faire un bilan de l'ensemble du programme créé, et d'établir un guide de recommandations générales sur l'ensemble du dispositif proposé, et plus spécifiques à chacune des séquences.

Chaque séance de cocréation est animée par les deux expertes en didactique du sport, et se déroule de la manière suivante :

Cadrage de la session (30 min)

Cette séance consiste à présenter la thématique à travailler par le biais du jeu sportif à créer, ainsi que les objectifs. Une présentation des étapes suivantes de la séance est ensuite effectuée, et des feuilles permettant une prise de notes pour chacune des étapes sont distribuées.

Identification des notions clés (20 min)

L'objectif de cette étape est de définir les éléments ou messages à transmettre à travers le jeu créé, tels que des émotions ou ressentis, des apprentissages spécifiques, ou encore les liens avec le quotidien (ex. : travailler les formules de politesse). Pour cela, un temps de réflexion individuel de dix minutes est octroyé, suivi d'un temps de mise en commun de dix minutes également.

Choix des leviers pédagogiques d'introduction de la thématique (50 min)

Pour introduire la thématique dans un jeu et y intégrer les notions clés préalablement identifiées, des leviers pédagogiques peuvent être utilisés. Les deux expertes en didactique du sport en ont recensé cinq, à savoir : « renommer l'espace, le matériel, les statuts », « modifier une ou plusieurs règles », « utiliser des annexes didactiques », « les enfants adaptent leurs objectifs dans le jeu et/ou créent de nouvelles règles », « jeu de rôle ou simulation d'émotion ». Après une présentation de chaque levier pédagogique, un nouveau temps de réflexion individuel d'une dizaine de minutes est proposé, afin que chacun identifie les deux formes de symboliques les plus pertinentes au regard de la thématique à travailler et des notions clés définies précédemment. Par exemple, pour la thématique « communication verbale et non verbale », mobiliser le levier pédagogique « utiliser des annexes didactiques » permet l'utilisation de pictogrammes pour communiquer. Puis, une mise en commun autorise chaque personne à présenter et justifier ses choix. Les deux formes les plus choisies sont retenues pour la suite de la cocréation, sous réserve d'un consensus.

Vérification de la cohérence (20 min)

Il s'agit ici de récapituler l'ensemble des éléments recensés précédemment (la thématique, les objectifs, les notions clés, les formes d'introduction des jeux sportifs) pour vérifier leur adéquation entre eux, et les faire valider par l'ensemble du groupe. Ce temps offre l'occasion aux membres de ce groupe de poser des questions si certaines subsistent.

Choix des formes de jeu (50 min)

Cette étape a pour intention de déterminer les formes du jeu les plus appropriées en fonction des notions clés et des leviers pédagogiques préalablement sélectionnés. Pour démarrer, un temps de réflexion de dix minutes est accordé aux participants pour compléter une fiche permettant de guider leur réflexion et de prendre des notes. Cette fiche comporte plusieurs questions visant à amener les participants à visualiser progressivement les types et formes de jeu les plus pertinents (ex. : jeu collectif/individuel ; jeu compétitif/coopératif ; nombre de séances sur ce jeu évolutif ; les principales règles de ce jeu ; le matériel/les annexes). Un temps de mise en commun de dix minutes est ensuite alloué pour recenser les différents choix et identifier les plus récurrents, permettant ainsi d'obtenir les contours du jeu. La dernière phase de cette étape a pour ambition de formaliser la séance. Pour cela, pendant trente minutes, l'ensemble des participants verbalisent l'idée (ou les idées) de jeu qu'ils imaginent, permettant à chacun de rebondir sur la proposition des autres. Ces idées sont notées sur un tableau blanc, sous forme de mots et de dessins, aboutissant progressivement à un jeu sportif concret.

Validation de la séance (30 min)

La séance de cocréation se clôture par trente minutes d'échanges reprenant le cheminement ayant permis la création du jeu. L'objectif est alors de vérifier que le jeu et les éléments qui le constituent sont bien en accord avec la thématique et les objectifs, présentés au début de la séance. Si des questions demeurent de la part des participants, elles sont passées en revue. Avant de se quitter, les feuilles de prise de notes distribuées au début de la séance sont récupérées par les animatrices expertes en didactique du sport.

À l'issue de chaque séance de cocréation, ces deux expertes ainsi que la doctorante travaillent sur la formalisation du jeu, et sur sa déclinaison en quatre à cinq séances. Un recensement des recommandations énoncées lors de chaque séance de cocréation est enfin formalisé dans un document appelé « guide des recommandations ».

2. PRÉSENTATION DU PROGRAMME PÉDAGOGIQUE

Les six séances de « cocréation » ont permis d'aboutir à un programme pédagogique inclusif composé de trois jeux sportifs dont les contenus sont évolutifs ; d'une séance « situation de référence » ayant pour but l'évaluation des habiletés ; et d'un guide de recommandations, dont le but est de fournir des indications aux professionnels qui mobiliseront ce programme. La mise en œuvre de ce programme est effectuée sur une durée de seize semaines. À raison d'une séance par semaine, il se déroule comme présenté dans le tableau 1.

Le programme pédagogique conçu a été élaboré sur la base des trois habiletés sociales citées précédemment, constituant les socles thématiques respectifs des jeux sportifs cocréés. Chaque séquence thématique se découpe en quatre à cinq séances évolutives. Deux objectifs généraux ont été identifiés pour chaque séquence thématique. Ces deux objectifs sont déclinés en plusieurs objectifs spécifiques de séance (1 à 2 par séance) qui évoluent au fil de l'avancée du programme. Les séances se déroulent selon le format suivant : échauffement (5 min), jeu sportif (15-20 min), temps calme (5 min).

Les jeux sportifs développés dans le cadre de ce programme sont conçus pour favoriser l'inclusion des enfants présentant un TSA. Leur mise en place prévoit donc la participation conjointe d'enfants présentant un TSA (deux à trois enfants) et d'enfants au développement typique (deux à trois enfants). Les règles de ces jeux sont adaptées à l'ensemble des enfants, et visent à créer de l'interaction entre eux.

TABLEAU 1 — Calendrier de la mise en œuvre du programme

Semaine 1	Semaines 2 à 5	Semaine 6	Semaines 7 à 10	Semaine 11	Semaines 12 à 15	Semaine 16
Évaluation initiale	Séquence thématique 1	Évaluation intermédiaire 1	Séquence thématique 2	Évaluation intermédiaire 2	Séquence thématique 3	Évaluation finale
Situation de référence	Séances de jeu sportif sur la thématique « Imitation et attention conjointe »	Situation de référence	Séances de jeu sportif sur la thématique « Communication verbale et non verbale »	Situation de référence	Séances de jeu sportif sur la thématique « Gestion des émotions »	Situation de référence

Un temps d'information auprès des professionnels intégrant ce programme est effectué au préalable afin d'en expliquer le déroulement et leur implication.

II. ÉVALUATION DU PROGRAMME PÉDAGOGIQUE INCLUSIF BASÉ SUR L'ACTIVITÉ PHYSIQUE ET LE JEU SPORTIF

L'ambition de notre étude est de déterminer si la mise en place de ce programme, auprès d'enfants ayant un TSA, conduit à une amélioration significative de leurs habiletés sociales et participe à une meilleure inclusion. Il convient ainsi d'évaluer les impacts de ce programme sous différents angles : d'abord, au niveau intra-personnel, en observant comment les habiletés sociales des enfants atteints de TSA évoluent tout au long du programme ; puis, en comparant les effets de ce programme axé sur l'activité physique et le jeu sportif avec d'autres types d'accompagnements pédagogiques ; enfin, en analysant les réactions affectives, cognitives et comportementales des individus dans leur environnement social comprenant les pairs, les professionnels qui les accompagnent, et leur famille. Cette étude a été approuvée par le comité d'éthique de l'université Rennes 2 (numéro de validation : 2022-028). À la suite de la présentation d'une notice d'information, les consentements éclairés ont été obtenus pour les enfants, les parents et les professionnels.

Pour évaluer notre dispositif pédagogique innovant, différentes options ont été adoptées : une approche comparative, une approche longitudinale, et une approche par triangulation de données. Sont ainsi impliqués dans cette évaluation des enfants avec un TSA, des enfants neurotypiques, des parents, et des professionnels ayant mis en œuvre le dispositif. L'ensemble de ces éléments sont détaillés dans la suite de cet article.

1. APPROCHE COMPARATIVE

Les études adoptant une approche comparative conduisent à tester l'efficacité d'un programme par rapport à celle d'un autre programme déjà en place. Dans cette perspective, la méthodologie utilisée implique la constitution d'au moins un groupe dit « expérimental » et d'un autre groupe appelé « témoin ». Dans la littérature, la plupart des études cherchant à analyser les effets de l'activité physique sur le plan social des enfants ayant un TSA sont des études dites « randomisées » (avec une répartition aléatoire des participants entre différentes modalités ou différents groupes expérimentaux). Ces études comparent un groupe qui suit le programme d'activité physique évalué à un groupe qui réalise la même activité, mais dans des conditions différentes, ou à un groupe qui effectue une autre activité, voire à un groupe n'ayant reçu aucune intervention proposée.

Dans le cadre de notre étude, pour des raisons éthiques, le choix a été fait de comparer notre groupe expérimental à un groupe témoin réalisant une autre activité mais de même nature, à savoir un programme pédagogique.

L'évaluation de l'efficacité du programme proposé sera donc réalisée en comparant le groupe expérimental participant à notre programme pédagogique inclusif basé sur l'activité physique et le jeu sportif, avec un groupe témoin qui suit un programme pédagogique inclusif mobilisant des activités manuelles pour améliorer les compétences sociales des enfants ayant un TSA. Chacun des groupes est composé de deux enfants atteints d'un TSA (deux garçons pour un groupe ; un garçon et une fille pour l'autre groupe) et de deux enfants au développement typique (deux filles pour chaque groupe), tous âgés de 8 à 10 ans. Les enfants avec TSA inclus dans cette étude sont accueillis au sein d'une unité d'accompagnement et d'apprentissage (U2A). Ils présentent de grandes difficultés de communication (non verbale, imitation verbale ou écholalie) et des troubles du comportement. Les séances se déroulent au sein de l'U2A, à raison d'une fois par semaine.

Notre choix d'activités pour le programme pédagogique du groupe témoin s'est porté sur des activités manuelles. En effet, ces dernières peuvent être réalisées dans les mêmes conditions que celles du groupe expérimental. C'est-à-dire en situation d'inclusion avec des enfants au développement typique, dans un espace clos ou ouvert, sur une durée de trente minutes. Ces activités sont déjà connues des enfants, qu'ils aient un TSA ou un développement typique, puisqu'elles sont déjà mises en place soit dans le cadre de leur scolarité, soit dans leur accompagnement quand celui-ci est effectué en structure.

L'utilisation d'une telle approche comparative pour évaluer l'efficacité d'un dispositif, si elle est pertinente au plan théorique, engendre toutefois un certain nombre de difficultés et de contraintes, d'autant plus importantes dès lors que les groupes constitués s'appuient sur des publics vulnérables. En effet, pour des résultats significatifs, une telle approche nécessite généralement la sollicitation d'un nombre important de participants. Or, dans le cadre de notre étude, il est souvent difficile de solliciter un grand nombre d'enfants ayant un diagnostic de TSA, assez proches géographiquement, et appartenant à une même tranche d'âge, afin que ces éléments ne constituent pas des biais explicatifs de phénomènes identifiés. De surcroît, le public TSA présente une importante diversité de profils de développement, qui dépendent notamment de la sévérité des symptômes autistiques, du niveau de développement et de l'âge chronologique. À cela s'ajoute une hétérogénéité développementale intra-individuelle importante, qui se caractérise par des niveaux différents entre les capacités cognitives, langagières, motrices et adaptatives. Il convient aussi de noter que les enfants présentant un niveau de sévérité de TSA assez important ont généralement peu, voire pas de temps prévu en inclusion avec des enfants au développement typique. Cela induit alors une difficulté supplémentaire à la mise en place de notre protocole d'étude. Enfin, les professionnels accompagnant ce public ne sont pas toujours disponibles pour pouvoir encadrer de nouvelles activités qui viennent s'ajouter à leurs activités professionnelles habituelles, donnant lieu à des difficultés lorsque l'on souhaite valider de nouveaux dispositifs, notamment en termes de sécurité.

Pour surmonter ces obstacles, il est souvent nécessaire d'adapter le protocole d'expérimentation (conçu de manière « théorique ») aux contraintes liées au terrain. Cette approche peut engendrer différentes actions, telles que la réduction du nombre d'enfants intégrés à l'étude, un plus grand investissement en termes de temps pour la sélection des participants et de leur établissement d'accompagnement, voire l'identification des enfants en amont, en fonction de leur profil de TSA, pour garantir une population d'étude plus homogène. Il s'agit dès lors de prendre en compte ces adaptations dans l'analyse des données recueillies.

2. APPROCHE LONGITUDINALE

Selon Safi (2011, p. 161-162), « on peut qualifier de longitudinal l'ensemble des dispositifs empiriques permettant d'intégrer la dimension temporelle dans l'analyse des faits sociaux. Une définition globale reposerait ainsi sur le principe de répéter un protocole d'observation dans le temps ». Une telle approche permet donc de suivre l'évolution dans le temps des thématiques au cœur de nos évaluations, chez un même sujet. Les études longitudinales sont donc des instruments privilégiés pour observer les changements dans le développement des enfants présentant un TSA, et évaluer l'efficacité d'un dispositif au niveau individuel.

Dans le cadre de notre étude, l'utilisation de cette approche va nous permettre de tester l'hypothèse selon laquelle le programme innovant basé sur l'activité physique et le jeu sportif que nous avons créé permet une amélioration des habiletés sociales des enfants ayant un TSA entre le début et la fin de sa mise en place. Cela nécessite que le programme élaboré soit expérimenté sur une période suffisamment longue pour que des transformations cognitives, affectives et comportementales soient possibles. Dans la littérature scientifique, les études proposent des périodes d'expérimentation allant de six à vingt semaines, suivant la fréquence et la durée des séances proposées dans la semaine. Dans le cadre de notre programme, celui-ci sera expérimenté sur une durée de seize semaines.

L'évaluation de la transformation des habiletés sociales des enfants nécessite que celles-ci soient mobilisées lors de situations comparables entre elles. À ces fins, une situation dite « référence » a été spécialement élaborée lors de l'étape de cocréation. Cette situation référence a pour objectif de mobiliser l'ensemble des habiletés sociales ciblées par le programme pédagogique. Elle est donc mise en place au début du programme et à la fin de celui-ci, afin d'effectuer une évaluation du dispositif dans son ensemble ; mais également pour des évaluations intermédiaires, après chaque cycle thématique, pour évaluer de manière plus fine les habiletés concernées par le cycle en question.

Les évaluations sont effectuées sous la forme d'observations écologiques retranscrites dans une grille spécifiquement réalisée pour cette étude. La

grille d'observation a été élaborée en tenant compte des différentes habiletés sociales identifiées dans les objectifs de séquences thématiques. Dans la littérature, la plupart des études longitudinales évaluent les compétences socio-communicatives par le biais d'entretiens ou de questionnaires complétés par les parents ou les professionnels accompagnants. Pour autant, l'activité physique et le jeu sportif mobilisés dans le cadre de notre étude favorisent les interactions, et celles-ci s'avèrent difficiles à évaluer par le biais de ces méthodologies, qui produisent des évaluations moins objectives. L'influence du contexte sur les compétences sociales d'un public présentant un TSA, et les interactions produites justifient l'utilisation d'observations écologiques pour les évaluer. Afin de renforcer la fiabilité des données, des enregistrements audio et vidéo des séquences sont assurés, permettant d'être exhaustif et précis dans les évaluations, et de procéder à des observations plurielles par plusieurs observateurs (calcul d'accords inter-observateurs).

Mobiliser une approche longitudinale dans le cadre de notre étude semble pertinent au vu de ce que nous venons de présenter. Cependant, cela ne va pas sans poser certaines difficultés. Tout d'abord, afin de limiter le biais lié à la présence du chercheur lors des séances d'évaluation, il est nécessaire que celui-ci soit présent à l'ensemble des séances mises en place auprès des enfants, imposant le suivi d'un effectif d'enfants restreint. De plus, si on combine la durée de l'étude à l'effectif restreint, nous pouvons être confrontés à l'abandon de quelques participants. Pour répondre à ces limites, il semblerait nécessaire, quand cela est possible, d'adosser plusieurs chercheurs à une même étude, permettant ainsi de suivre un plus grand nombre d'enfants.

3. TRIANGULATION DE DONNÉES

Comme indiqué précédemment, évaluer l'impact d'un programme pédagogique inclusif nécessite non seulement d'analyser son effet sur le jeune lui-même, mais également auprès de sa famille, et des autres acteurs engagés dans les activités proposées, tels que les professionnels et les pairs. Effectuer une triangulation des données a pour ambition d'évaluer au mieux les effets de ce programme auprès de l'ensemble des bénéficiaires, en recueillant des données complémentaires portant sur un même objet. Le croisement des points de vue des acteurs et des outils de recueil de données met en exergue la concomitance des résultats, ou la nécessité de les approfondir. Cela donne la possibilité de confronter les données issues de différents points de vue, et cela dans le but d'identifier les convergences ou divergences, proposant ainsi une analyse plus qualitative des données recueillies. Dans le cadre de notre étude, les comportements adaptatifs des enfants ayant un TSA sont donc complémentairement évalués par le biais de la version française du VABS-II (Vineland Adaptative Behavior Scales, seconde édition) aussi appelée Vineland-II. Ces évaluations sont réalisées par les professionnels intégrant notre étude, en amont et à la fin du programme. Par ailleurs, les compétences socio-émotionnelles sont évaluées par le biais du SSIS-SEL (Social Skills Improvement System – Social Emotio-

nal Learning edition, CASEL[48], 2015). Celui-ci est administré aux parents des enfants concernés, au début et à la fin du dispositif.

En complément des données portant sur les enfants ayant un TSA, il est impératif de considérer les différents acteurs qui prennent part à notre étude, afin d'évaluer l'impact plus global du programme pédagogique proposé. Cette démarche s'ancre dans une approche dite « socio-écologique », qui prône une analyse à plusieurs échelles[49]. De ce fait, nous interrogerons les enfants au développement typique, les professionnels, ainsi que les parents des enfants TSA intégrant l'étude. Plus précisément, des entretiens collectifs sont conduits avec les quatre enfants au développement typique participant à l'étude (programme d'activités physiques et programme d'activités manuelles), en amont de la mise en place des activités, pour identifier leurs attitudes et leurs croyances envers le TSA et l'inclusion des enfants qui en sont atteints. De nouveaux entretiens conduits à l'issue du programme permettent d'évaluer si le dispositif mis en place a engendré des évolutions à ce propos, et de recueillir le ressenti global de ces enfants sur leur participation à ces activités, afin d'identifier les points forts et les points d'amélioration de ces dernières.

De plus, un entretien semi-directif est réalisé à l'issue de la mise en place des activités, avec les parents de chaque enfant ayant un TSA, et avec les deux professionnels ayant intégré notre étude. L'objectif est de recenser leur perception concernant l'impact du programme sur les enfants ayant un TSA, et sur les enfants au développement typique (pour les professionnels). L'entretien avec les professionnels constitue aussi une occasion pour obtenir leur témoignage sur leur participation à ces activités (ressenti, points forts et points d'amélioration, difficultés rencontrées, etc.).

Les données recueillies grâce à ces différents outils nous donnent la possibilité de réaliser une évaluation approfondie des enfants atteints d'un TSA, par la triangulation des données provenant des chercheurs, des enfants neurotypiques, des parents et des professionnels. Elles permettent, par la suite, de faire évoluer le dispositif (contenu du programme basé sur les activités physiques et le jeu sportif, durée des séquences, fréquence des séquences, habiletés retenues, etc.), afin de produire un protocole répondant de façon la plus pertinente possible aux objectifs établis à l'origine du projet, à savoir permettre à des enfants ayant un TSA de développer leurs compétences sociales pour favoriser leur inclusion.

48. CASEL : Collaborative for Academic, Social, and Emotional Learning (groupe de recherche et d'influence sur l'apprentissage social et émotionnel et son implication sur les performances scolaires).

49. Les différentes échelles : ontosystème ; microsystème ; mésosystème ; exosystème ; macrosystème.

CONCLUSION

Mobiliser une approche comparative et longitudinale pour évaluer l'impact du programme pédagogique innovant, en effectuant par ailleurs une triangulation des données avec les différents acteurs, nous permet d'évaluer finement l'efficacité du dispositif à différentes échelles : au niveau individuel (démarche longitudinale avec plusieurs échéances intermédiaires et progression), au niveau collectif (comparaison de groupes avec programmes différents), et par le biais de points de vue différents (pairs, parents, professionnels). Toutefois, en dépit de la pertinence de cette approche, la difficulté principale réside dans la faible ampleur des effectifs pouvant être sollicités, nous amenant à être très prudents dans les conclusions qui seront avancées. De nouvelles mises en œuvre, tenant compte des éléments de discussion identifiés, doivent être proposées, avant de réaliser l'objectif, à terme, de déploiement de ce programme innovant au sein de différentes structures accueillant des enfants ayant un TSA, à l'échelle nationale.

Bibliographie

Cay Anderson-Hanley, Kimberly Tureck, Robyn L. Schneiderman (2011). « Autism and exergaming: effects on repetitive behaviors and cognition. », in *Psychology research and behavior management,* pp. 129-137.

Christopher Petrus, Sarah R. Adamson, Laurie Block, Sarah J. Einarson, Maryam Sharifnejad, Susan R. Harris (2008). « Effects of exercise interventions on stereotypic behaviours in children with autism spectrum disorder. », *Physiothérapie Canada,* 60(2), pp. 134-145.

Jenna Lequia, Wendy Machalicek, Mandy J. Rispoli (2012). « Effects of activity schedules on challenging behavior exhibited in children with autism spectrum disorders: A systematic review », in *Research in Autism Spectrum Disorders,* 6(1), pp. 480-492.

Fatimah Bahrami, Ahmadreza Movahedi, Sayed Mohammad Marandi, Carl Sorensen (2016). « The effect of karate techniques training on communication deficit of children with autism spectrum disorders », in *J Autism Dev Disord,* 46(3), pp. 978-986.

Mahboubeh Ghayour Najafabadi, Mahmoud Sheikh, Rasoul Hemayattalab, Amir-Hossein Memari, Maryam Rezaii Aderyani, Sina Hafizi (2018). « The effect of spark on social and motor skills of children with autism », in *Pediatrics & Neonatology,* 59(5), pp. 481-487.

Sean Healy, Adam Nacario, Rock E. Braithwaite, Chris Hopper (2018). « The effect of physical activity interventions on youth with autism spectrum disorder: A meta-analysis », in *Autism Research,* 11(6), pp. 818-833.

Michelle Sowa, Ruud Meulenbroek (2012). « Effects of physical exercise on autism spectrum disorders: A meta-analysis », in *Research in Autism Spectrum Disorders,* 6(1), pp. 46-57.

Mengxian Zhao, Shihui Chen (2018). « The effects of structured physical activity program on social interaction and communication for children with autism », in *BioMed Research International,* pp. 1-13.

Janette Hynes, Martin Bloc (2022). « Effects of Physical Activity on Social, Behavioral, and Cognitive Skills in Children and Young Adults with Autism Spectrum Disorder: a Systematic Review of the Literature », in *Review Journal of Autism and Developmental Disorders*.

Margaret M. Bass, Catherine A. Duchowny, Maria M. Llabre (2009). « The effect of therapeutic horseback riding on social functioning in children with autism », in *J Autism Dev Disord,* 39(9), pp. 1261-1267

Maël Le Paven, Sophie Roesslé, Emmanuelle Roncin, Monique Loquet, Yvon Léziart (2007). « La dévolution dans les activités physiques sportives et artistiques non scolaires », in *Éducation et didactique,* 1-3, pp. 9-29.

Andrew M. Colombo-Dougovito, Jihyun Lee (2021). « Social skill outcomes following physical activity–based interventions for individuals on the autism spectrum: A scoping review spanning young childhood through young adulthood », in *Adapted Physical Activity Quarterly,* 38(1), pp. 138-69.

Crystal F. Branta, Jacqueline D. Goodway (1996). « Facilitating social skills in urban school children through physical education. », in *Peace and Conflict: Journal of Peace Psychology,* 2(4), pp. 305-319.

Elaine McHugh (1995). « Going 'beyond the physical': Social skills and physical education », in *Journal of Physical Education, Recreation & Dance,* 66(4), pp. 18-21.

Martin Gendron, Egide Royer, Richard Bertrand, Pierre Potvin (2006). « Les troubles du comportement, la compétence sociale et la pratique d'activités physiques chez les adolescents », in *Revue des sciences de l'éducation,* 31(1), pp. 211-233.

Olivia Collet, Sandrine Le Sourn-Bissaoui, Geneviève Cabagno (2022). « Regards croisés sur la place de l'activité physique dans l'accompagnement des enfants ayant un TSA », *1er colloque international du Groupement national des centres de ressources Autisme* (GNCRA), TSA développer, partager, innover !, Lyon, 19-20 mai 2022.

Bérangère Storup, Claudia Neubauer, Glen Millot (2012). « La recherche participative comme mode de production de savoirs : un état des lieux des pratiques en France. », in *Fondation Sciences Citoyennes.*

Pierre Fournier, Manon Ménard, Elisa Wrembel (2021). « Expérimenter la recherche-projet en doctorat : applications dans les champs de la santé et du handicap », in *Études de communication,* 56, pp. 55-76.

Sophie Genelot, Bruno Suchaut (2022). « L'évaluation des innovations pédagogiques : quelles modalités de coopération entre les différents acteurs ? », in *Mesure et évaluation en éducation,* 23(1), pp. 23-42.

Chia-Hua Chu, Chien-Yu Pan (2012). « The effect of peer- and sibling-assisted aquatic program on interaction behaviors and aquatic skills of children with autism spectrum disorders and their peers/siblings », in *Research in Autism Spectrum Disorders,* 6 (3), pp. 1211-23.

Leah Ketcheson, Janet Hauck, Dale Ulrich (2017). « The effects of an early motor skill intervention on motor skills, levels of physical activity, and socialization in young children with autism spectrum disorder: A pilot study », in *Autism*, 21(4), pp. 481-492.

M.A. Bernard, Erice Thiébaut, Camilla Mazetto, et al. (2016). « L'hétérogénéité du développement cognitif et socio-émotionnel d'enfants atteints de trouble du spectre de

l'autisme en lien avec la sévérité des troubles », in *Neuropsychiatrie de l'enfance et de l'adolescence,* 64(6), pp. 376-382.

Mirna Safi (2011). « Chapitre 10. L'analyse longitudinale données et méthodes », in *La France dans les comparaisons internationales*, Presses de Sciences Po, pp. 161-172.

Amaria Baghdadli, Charles Aussilloux (2002). « Intérêts et limites des études longitudinales dans la mesure du changement chez la personne autiste », in *Enfance*, 54(1), p. 40.

Urie Bronfenbrenner (1979). *The Ecology of Human Development: Experiments by Nature and Design.* Harvard University Press.

Séverine Ferrière, Fabien Bacro, Agnès Florin, Philippe Guimard (2016). « Le bien-être en contexte scolaire : Intérêt d'une approche par triangulation méthodologique. », in *Les Cahiers internationaux de psychologie sociale,* 111(3), pp. 341-365.

Sara S. Sparrow, Domenic V. Cicchetti, David A. Balla (2005). « Vineland Adaptive Behavior Scales–Second Edition (Vineland–II) », in *Circle Pines,* MN: American Guidance Service.

Utilisation d'un jeu en pédagogie universitaire ; évaluation de l'attention des étudiants

François Lecellier

Résumé

Au sein de l'université de Poitiers, le jeu vidéo Minecraft est utilisé depuis de nombreuses années à différents niveaux. L'évaluation de cette utilisation a été réalisée de trois manières distinctes : 1) au moyen des notes obtenues par les étudiants traduisant leur acquisition de compétences ; 2) par le biais de questionnaires anonymes transmis aux étudiants ; 3) au travers de la mesure de l'attention visuelle des étudiants lors des séances. L'ensemble des trois évaluations montrent une augmentation de l'attention et de l'engagement des étudiants lors de l'utilisation du jeu, par rapport à des séances plus traditionnelles.

INTRODUCTION

En 2015, l'université de Poitiers a entamé la transformation pédagogique d'un module de « gestion de projet » au sein du département « génie électrique et informatique industrielle » de l'IUT de Poitiers-Niort-Châtellerault. Ce module comprend une étude bibliographique et la réalisation d'un mini-projet technique grâce à l'utilisation du jeu Minecraft. L'objectif est clair : à travers cette transformation, il s'agit d'utiliser le jeu vidéo pour différencier les compétences de gestion de projet, des compétences techniques en cours d'acquisition.

Cette modification de la pédagogie a entraîné de nombreux questionnements sur la mesure de l'efficacité de cette nouvelle pratique pédagogique. En effet, le jeu vidéo dans un sens pédagogique nécessite une appropriation par l'équipe enseignante, mais aussi par les étudiants. Ces derniers connaissent, en règle générale, le domaine du jeu vidéo, mais ne sont pas habitués à l'utiliser dans un contexte scolaire, et encore moins universitaire. C'est donc tout naturellement que la question de l'évaluation est apparue pour permettre de déterminer les points aussi bien positifs que négatifs d'une telle approche.

Afin d'évaluer le dispositif pédagogique mis en œuvre à l'université de Poitiers, il a été décidé de s'appuyer sur trois indicateurs : 1) les notes traduisant l'acquisition de compétences liées à la gestion de projet ; 2) les retours sur questionnaires d'évaluation ; 3) l'évaluation de l'attention visuelle des étudiants lors des séances. Ces trois indicateurs déterminent de manière globale l'attention, l'intérêt et l'acquisition des compétences nécessaires à la validation de l'approche pédagogique innovante.

Dans le cadre de cet article, nous éclairerons la manière dont les trois indicateurs ont été collectés, et comment ils soulignent la pertinence de l'intégration du jeu vidéo dans l'enseignement de la gestion de projet pour les étudiants en premier cycle universitaire. Après avoir contextualisé la transformation pédagogique, nous présenterons les deux premiers indicateurs, avant de nous pencher sur l'expérimentation de la mesure de l'attention visuelle, pour finalement conclure.

I. CONTEXTE

Dans cette première partie, nous détaillerons les éléments contextuels pour déterminer les modalités pédagogiques utilisées et leur évaluation. Le module qui exploite le jeu vidéo à des fins pédagogiques est intitulé « conduite

de projet » dans l'ancien programme pédagogique national du diplôme de DUT « génie électrique et informatique industrielle » (GEII). Dans le cadre de la réforme du bachelor universitaire de technologie (BUT), ce module a été transformé. Il s'intitule désormais « vie de l'entreprise ». Ce qui élargit les compétences et connaissances à enseigner. Pour la suite de cet article, nous parlerons de gestion de projet pour couvrir les deux modules, et la notion de manière plus générale.

Le module arrive très rapidement dans le parcours universitaire des étudiants, puisqu'il est enseigné dès le premier semestre universitaire. L'enseignement de la gestion de projet se révèle déjà complexe pour les étudiants en fin de licence ou en master, comme le soulignent Thomas et Mengel (2008). Ces auteurs insistent notamment sur le fait que, pour gérer efficacement un projet, les managers doivent apprendre à composer avec l'incertitude. En 2013, les travaux de Shariff, Johan et Jamil (2013) mettent en lumière que les compétences en gestion de projet sont plus facilement développées lorsque les étudiants sont en mesure de s'organiser eux-mêmes dans la conception, la définition et la réalisation de leurs projets.

Pour ces différentes raisons, l'enseignement de la gestion de projet pour des étudiants de première année est un véritable défi pédagogique. D'un côté, nous ne pouvons pas leur demander de mener à bien un projet technique dans le domaine du génie électrique ou de l'informatique industrielle, sans les guider profondément. D'un autre côté, l'aspect pratique de la gestion de projet est capital pour leur permettre d'acquérir les compétences visées.

Dans la perspective de résoudre ce challenge, nous avons décidé d'utiliser le jeu vidéo afin de passer outre les compétences techniques, et de pouvoir évaluer réellement les compétences de gestion de projet.

1. PRÉSENTATION DU MODULE DE « GESTION DE PROJET »

Pratiquement, le module de « gestion de projet » est organisé de la manière suivante :

- présentation des objectifs du module et de la méthodologie générale de gestion de projet en cours magistral ;
- séances sur le jeu vidéo en travaux dirigés en présence d'un enseignant ;
- travail personnel en dehors de séances, depuis l'université ou à distance.

Les sessions sur le jeu vidéo sont composées de quatre séances de deux heures de travaux dirigés. Durant ces séances, les étudiants sont répartis en groupes de six individus. Leur tâche consiste à gérer le projet de bout en bout. À cet effet, ils sont appelés à :

- décider de manière collégiale comment répondre à la consigne générale donnée (créer une construction au sein du jeu Minecraft) ;
- réaliser leur construction en fonction de la connaissance du jeu de chaque membre du groupe ;
- respecter les contraintes de temps (un mois et demi) et d'espace (zone de construction prédéfinie au sein de l'univers virtuel) ;
- présenter leur construction et la manière dont ils ont réparti leur travail.

L'ensemble de ces étapes permet de valider les compétences étudiées et mises en pratique en gestion de projet.

2. MÉTHODES D'ÉVALUATION

Dans le cadre du module, nous proposons une évaluation formelle des compétences de gestion de projet. Pour cela, les étudiants reçoivent une note qui représente leur acquisition de compétences. Cette notation ne se focalise pas sur la réalisation, mais sur la méthodologie de réalisation. C'est-à-dire la coopération entre les membres du groupe, la communication interne et externe, et le respect du cahier des charges. Si cette méthode d'évaluation est importante dans un contexte d'étude, elle ne reflète pas l'engagement des étudiants dans le module.

Pour évaluer cet engagement, nous avons tout d'abord invité les étudiants à compléter un questionnaire visant à identifier les points positifs et négatifs de l'enseignement. Cette évaluation permet de déterminer le degré d'acceptation et d'engagement général des étudiants dans le module. Toutefois, il convient d'admettre qu'elle est subjective, et ne permet pas de mesurer objectivement l'attention des étudiants dans le cadre du cours.

Pour y parvenir, nous avons procédé à une dernière évaluation au travers de mesures d'oculométrie. Celles-ci permettent de saisir l'attention visuelle et les points de focalisation des utilisateurs dans tout contexte. Dans ce cas, nous avons choisi de réaliser la mesure de l'attention visuelle des étudiants lors des séances de travaux dirigés.

II. ÉTAT DE L'ART EN MATIÈRE D'ÉVALUATION DE L'ATTENTION VISUELLE

En ce qui concerne l'évaluation de l'attention visuelle des étudiants lors de séquences pédagogiques, Jarodzka, Skuballa et Gruber (2021) observent

qu'il est complexe de comparer les résultats entre plusieurs étudiants, du fait de la variabilité des séances, des acteurs et du contexte. En effet, il n'est pas possible de proposer des séquences identiques lors d'une session de cours. Les auteurs constatent, cependant, que cette variabilité n'est pas problématique si l'on réalise les mesures pendant des activités pédagogiques similaires. Seules quelques études ont mesuré l'attention visuelle des étudiants lors de séquences habituelles d'apprentissage. En règle générale, il s'agit d'expérimentations en laboratoire (Przybyło, Kantoch, & Augustyniak, 2019) ou en vidéo (Wang, Antonenko, & Dawson., 2020). Pour autant, certaines conclusions s'imposent, y compris dans le cadre de ces expérimentations. En particulier, les émotions ressenties par un étudiant et son niveau de stress ont un impact direct sur son attention visuelle. Le stress diminue de manière sensible le nombre de saccades et de fixations. Durant un cours en vidéo, la présence ou non de l'enseignant à l'écran change du tout au tout l'attention visuelle. Lorsque l'enseignant est présent à l'écran, l'attention visuelle et l'acquisition de connaissances augmentent.

D'autres auteurs ont mené des études sur l'attention visuelle en classe. Par exemple, les travaux de Yang, Chang, Chien, Chien, & Tseng (2013), ou encore ceux de Jarodzka, Skuballa et Gruber (2020) mettent en évidence que l'attention visuelle des étudiants est corrélée à différents facteurs : 1) l'engagement de l'étudiant dans le cadre de l'activité pédagogique ; 2) la présence et les interactions avec l'enseignant ; 3) et bien entendu la présence et les interactions possibles avec les autres étudiants. Un autre point à noter est que l'obtention de résultats dans un environnement non standardisé (en dehors du laboratoire) demeure rare. Cependant, les études révèlent que l'attention visuelle représente un phénomène complexe. Elle ne traduit pas, à elle seule, la qualité ou la bonne acquisition de connaissances au sein d'un processus d'enseignement. De plus, les expérimentations réalisées en laboratoire montrent leurs limites pendant le passage à l'échelle et en conditions réelles d'enseignement.

Par ailleurs, la réglementation en termes de données personnelles s'avère compliquée en raison de la capacité des oculomètres à capturer en temps réel un volume important de données personnelles. À titre d'exemple, toutes les personnes présentes dans la pièce peuvent être vues pendant les enregistrements. Par conséquent, obtenir le consentement de chaque individu, y compris les étudiants et l'enseignant, est indispensable. Cela se complexifie si un des étudiants est mineur. L'ensemble de ces difficultés se traduit par un nombre faible d'études sur l'attention visuelle en situation pédagogique réelle. Idéalement, il serait souhaitable de recueillir des données pour chaque individu présent dans la salle au cours de la même séquence pédagogique. Mais, cette perspective se heurte à des contraintes financières et logistiques. Il est impossible d'équiper chaque personne d'un eye-tracker dans une salle de classe. La raison ? Cet outil est trop coûteux et difficile à mettre en œuvre.

III. L'ÉVALUATION DU MODULE DE GESTION DE PROJET

Comme nous l'avons indiqué dans la présentation du contexte, nous avons décidé d'utiliser trois mesures différentes pour évaluer la transformation pédagogique du module de gestion de projet : 1) une évaluation classique au travers de notation ; 2) une évaluation subjective par les étudiants ; 3) une évaluation de l'attention visuelle.
Ce sont ces trois mesures que nous allons détailler maintenant.

1. ÉVALUATION STANDARD DU MODULE

Cette évaluation est standard dans un contexte d'enseignement. La notation reflète le niveau d'acquisition de connaissances ou de compétences selon une grille d'évaluation spécifique. Dans notre cas, l'objectif de l'enseignement est clair. Quant à la grille d'évaluation, elle comprend cinq points :

- communication interne dans le groupe ;
- participation de l'ensemble des membres du groupe à la réalisation finale ;
- organisation du travail entre les étudiants familiers ou moins familiers du jeu utilisé ;
- réalisation conforme au cahier des charges initial ;
- présentation finale / communication externe.

Nous avons comparé les notes obtenues par l'ensemble des étudiants de première année sur le module, soit un total de 65 étudiants. En comparant les notes obtenues par les étudiants pendant l'utilisation de Minecraft et hors utilisation de Minecraft, on se rend compte d'une progression des notes de 20 % (trois points) en moyenne avec l'utilisation du jeu. Cette augmentation témoigne que les étudiants acquièrent bien les compétences que nous souhaitons développer.

2. ÉVALUATION PAR LES ÉTUDIANTS

Cette seconde évaluation est certainement la plus subjective. En effet, les étudiants vont évaluer non seulement le module et les compétences qu'ils acquièrent, mais aussi le logiciel, les méthodes d'enseignement, les interactions avec le professeur… Toutefois, en cas d'évaluation comparée entre le cours sans le jeu et le cours avec le jeu, on peut considérer que l'impact des différents biais est presque le même.

Dans ce cas, nous avons demandé aux étudiants de noter sur une échelle de un à sept leur satisfaction sur le module, puis de répondre à quatre questions :

- Quels sont les points forts du module ?
- Quels sont les points faibles du module ?
- Quelles améliorations proposer ?
- Avez-vous des suggestions complémentaires ?

Cette évaluation n'étant pas obligatoire, tous les étudiants n'y ont pas pris part. Seuls 43 ont répondu sur les 65 concernés par le module. Les résultats sont clairs : les notes données par les étudiants au module augmentent de 26 % en fonction de l'utilisation, ou non, du logiciel. De manière plus détaillée, les étudiants considèrent que s'approprier les compétences de manière ludique les aide à s'impliquer dans la réalisation de leur projet. La grande majorité des commentaires sont positifs, et attestent de l'investissement des étudiants dans le module (encadré 1).

ENCADRÉ 1 — Commentaires d'étudiants

« L'approche pédagogique est très abordable par tout le monde pour avoir une gestion du groupe cohérente. »
« Projet sur un logiciel intéressant et plaisant (Minecraft). Bonne ambiance. Apprends à travailler en groupe. »
« Le module a permis de travailler en groupe plus que d'habitude. Étant donné que c'était des groupes de cinq à six personnes, il a fallu se mettre d'accord sur beaucoup de choses pour pouvoir avancer. J'ai trouvé ça très intéressant. »
« Un cours intéressant, compte tenu de l'importance du travail d'équipe dans les domaines vers lesquels nous dirige la formation GEII. »
« Il serait bon de donner des objectifs pour chaque séance et de demander plus d'organisation dans les groupes. »

À la lecture des commentaires des étudiants, il est clair que la majorité d'entre eux apprécient le module. Cet intérêt peut contribuer à l'engagement des étudiants. Certains, en revanche, affirment ne pas être intéressés par cet exercice. Et d'autres avouent ne pas comprendre l'intérêt d'utiliser le jeu pour la gestion de projet (encadré 2).

ENCADRÉ 2 — Remarques négatives d'étudiants

« Apprendre à gérer des projets avec un jeu vidéo qui n'est pas maîtrisé de la même façon par tous les élèves, c'est nul et en plus moi, perso, je déteste Minecraft. »

« Les travaux dirigés ne sont pas pour moi assez représentatifs de ce que l'on attend en conduite de projet. »

« Certains étudiants peuvent se poser des questions concernant l'utilité de travailler sur Minecraft. »

3. ÉVALUATION DE L'ATTENTION VISUELLE

Afin d'évaluer de manière objective l'attention des étudiants dans ce contexte spécifique d'apprentissage, nous avons mis en place un protocole expérimental pour mesurer leur attention visuelle durant les séances. L'attention visuelle est mesurée au moyen d'un oculomètre qui permet de déterminer les points de fixation oculaires d'un utilisateur au fil du temps. Il existe plusieurs modèles d'oculomètres : certains sont fixes sur l'écran d'un ordinateur, et d'autres portés sous forme de lunettes. C'est ce second modèle que nous avons utilisé.

Le protocole expérimental est le suivant :

- l'étudiant est équipé d'un oculomètre durant 30 minutes en travaux dirigés ;
- l'oculomètre utilisé est celui proposé par la société Tobii sous la référence Glasses 3, et permet d'enregistrer le champ visuel de l'utilisateur avec une résolution spatiale de 1 920 x 1 080 pixels, et une résolution temporelle de 60 Hz ;
- deux types de travaux dirigés sont analysés, des séances sur le jeu Minecraft, et des séances plus traditionnelles avec papier/crayon ;
- pour limiter les biais de l'étude, nous n'analyserons les séquences qu'à partir de cinq minutes d'enregistrement et pour une durée de 20 minutes.

Ce protocole a été mis en place pour permettre une mesure la plus fiable et reproductible possible.

4. MESURES RÉALISÉES

Nous avons réalisé les mesures auprès de quatorze étudiants de première année universitaire en séquences de travaux dirigés. Pour sept d'entre eux, il s'agissait d'une séquence traditionnelle, et pour les sept autres d'une séquence sur Minecraft. Dans la séquence traditionnelle, les étudiants participent aux travaux dirigés d'informatique industrielle, durant lesquels ils sont chargés de concevoir des algorithmes correspondants aux problèmes po-

sés. Avec Minecraft, les étudiants se concentrent sur leur construction, s'interrogent sur le sujet, et travaillent en utilisant des feuilles distribuées ainsi que l'aide apportée par le professeur au tableau.

Les enregistrements contiennent un taux d'acquisition oculométrique supérieur à 80 %. L'utilisation des 20 minutes centrales permet de ne pas introduire de biais liés à la mise en place ou au retrait des lunettes. En effet, les étudiants mettent un peu de temps à s'habituer à porter les lunettes ; et ils ont tendance à bouger plus rapidement, ou inversement, plus lentement en début d'acquisition, et vers la fin de cette dernière.

Au cours de ces séances, nous avons entrepris de mesurer les fixations visuelles. Ces mouvements oculaires spécifiques se caractérisent par un maintien du regard, relativement fixe, pendant une durée supérieure à 50 millisecondes. Nous avons enregistré le nombre de fixations, la durée de ces dernières, ainsi que les cartes de chaleur qui en résultent. Par la suite, nous avons calculé les moyennes de ces trois mesures pour pouvoir comparer les résultats selon les deux modalités d'enseignement. Dans cette perspective, nous avons cherché à distinguer les zones d'intérêt des autres zones sur les enregistrements. Nous avons décidé de considérer comme zone d'intérêt les zones suivantes :

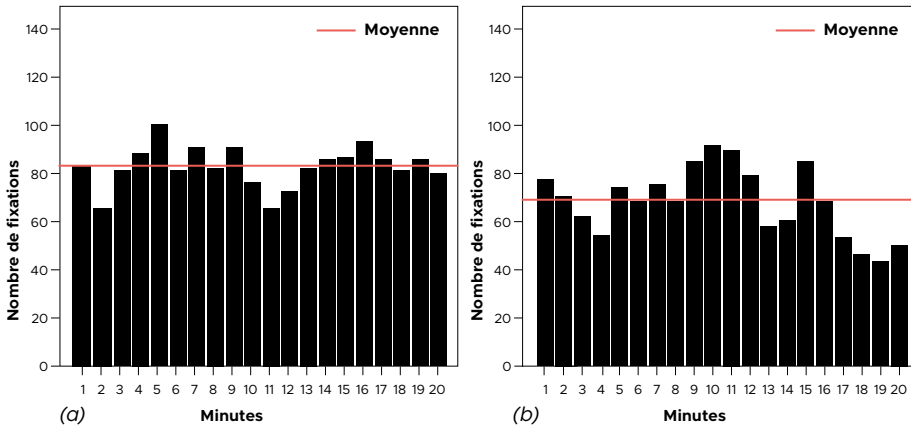
- l'ordinateur, le tableau et le professeur dans le cadre d'une séquence sous Minecraft ;
- le papier/crayon, le tableau et le professeur dans le cadre d'une séquence plus traditionnelle.

Le contenu situé en dehors de ces zones est identifié comme étant « hors zone d'intérêt ». Il importe de préciser que pendant les séquences avec Minecraft, si l'ordinateur n'est pas utilisé pour le jeu, les fixations ne sont pas considérées comme « en zone d'intérêt ». Enfin, pour pouvoir agréger au mieux les différentes mesures, nous avons effectué la moyenne de chaque paramètre sur 20 périodes d'une minute pendant les 20 minutes analysées. Cette approche permet de visualiser de manière plus synthétique les résultats expérimentaux.

5. RÉSULTATS EXPÉRIMENTAUX

Si l'on commence par analyser le nombre moyen de fixations dans des zones d'intérêt, on constate que ce nombre varie de manière importante selon le type de séquence. Lors d'une séquence standard, 64,5 % des fixations se font dans des zones d'intérêt, alors qu'avec Minecraft, ce sont plus de 86 % des fixations qui sont focalisées dans ces zones. De manière plus spécifique, l'évolution du nombre de fixations au fil du temps est beaucoup plus régulière avec Minecraft que dans une séance traditionnelle (figure 1).

FIGURE 1 — Nombre moyen de fixations par minute pour (a) une séance sous Minecraft et (b) une séance traditionnelle



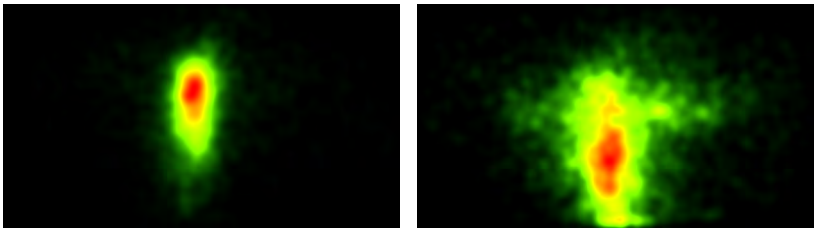
Cette évolution se traduit par une différence dans l'attention des étudiants entre les deux types de séquences. Dans la séquence traditionnelle, on observe une diminution significative de l'attention au fil du temps. Ce qui s'explique par un relâchement de l'attention des étudiants en cours de la séance. À l'inverse, lors des séquences avec Minecraft, l'implication des étudiants est plus régulière. Ce qui les maintient actifs en permanence. Cette diminution de l'attention, en fonction du temps, est un phénomène classique lors des mesures de la concentration ou de l'attention, qu'elles soient visuelles ou non, comme l'indique Neil A. Bradbury (2016). Lorsque l'on compare la durée des fixations plutôt que leur fréquence, l'écart se révèle tout aussi notable. Les étudiants consacrent plus de 92 % de leur temps à des zones d'intérêt dans le cadre des séquences avec Minecraft, tandis que dans le contexte plus traditionnel, cette proportion s'élève à près de 75 %.

L'analyse des fixations visuelles montre que le regard des étudiants se focalise sur les zones d'intérêt pendant la majeure partie du temps, quelle que soit l'activité pédagogique. Mais, lorsqu'ils utilisent Minecraft, cette attention augmente, atteignant des niveaux élevés (plus de 80 %). Cela explique en grande partie la réception positive de Minecraft par la plupart des étudiants, et pourquoi ce dispositif ludique favorise une meilleure acquisition des connaissances.

Enfin, nous avons étudié les cartes de chaleur, qui illustrent le positionnement des points de fixation dans le champ visuel de l'observateur. La zone où les fixations sont plus fréquentes est représentée en rouge, tandis que celle avec moins de fixations apparaît en vert ou en noir. En comparant ces cartes de chaleur (figure 2), plusieurs différences émergent. Tout d'abord, lorsque les étudiants utilisent Minecraft, leur regard se concentre davantage vers le centre de leur champ visuel, contrairement à la séance traditionnelle. De plus, le regard des participants se concentre dans une région plutôt restreinte sous

Minecraft par rapport aux séances avec papier/crayon. La focalisation du regard sur le centre du champ visuel est habituelle et totalement normale pour les humains. Il s'agit du « biais centré », comme l'ont expliqué Zhao et Koch (2013).

FIGURE 2 — Cartes de chaleur pour (a) une séance sous Minecraft et (b) une séance traditionnelle



(a) *(b)*

La différence observée peut être attribuée à l'activité menée par les étudiants. L'usage de Minecraft induit une position plus élevée de l'écran de l'ordinateur, par rapport à celle d'une feuille de papier lorsqu'ils travaillent avec un crayon. En outre, l'attention est plus facilement perturbée par les autres étudiants en séance traditionnelle. Par ailleurs, avec Minecraft, l'enseignant a tendance à donner moins d'explications au tableau qu'en séance traditionnelle. L'ensemble de ces éléments met en évidence une nette différence dans l'attention des étudiants en fonction du type de séquence pédagogique. L'objectif sous-jacent à la comparaison des cartes de chaleur n'est pas de déterminer quelle modalité pédagogique est la plus performante, mais plutôt de souligner que l'attention visuelle des étudiants diffère fondamentalement en fonction de l'activité pédagogique à laquelle ils participent.

CONCLUSION

Cet article présente plusieurs méthodes d'évaluation d'un dispositif pédagogique innovant en pédagogie universitaire : l'utilisation du jeu Minecraft pour la gestion de projet. Les mesures réalisées soulignent une meilleure attention visuelle, ainsi qu'une amélioration de l'acquisition des compétences, de l'appropriation et de l'engagement des étudiants pendant les séances. Le recours à trois stratégies d'évaluation du dispositif permet de le valider de manière fiable et objective. Il est bien entendu possible de considérer les variations sur les notes ou l'évaluation par les étudiants comme des mesures relativement subjectives. Toutefois, la mesure objective de l'attention visuelle des étudiants corrobore la plus-value de l'approche pédagogique dans ce contexte particulier.

Le travail avec Minecraft a été étendu au-delà du cadre universitaire. Il a été utilisé auprès des élèves de primaire, de collège et de lycée. Dans chaque cas, les équipes enseignantes ayant expérimenté cette pédagogie ont constaté une augmentation de l'attention des apprenants. Pour autant, contrairement aux étudiants, aucune évaluation n'a été mise en place pour ces publics.

À ce jour, nous envisageons de réaliser de nouvelles séquences de mesures oculométriques, pour évaluer l'attention des apprenants dans d'autres contextes, et relier attention visuelle, acquisition de compétences et engagement des apprenants. Cela pourrait se réaliser en ajoutant à la mesure d'attention visuelle une analyse du comportement, pour mesurer l'attention globale des apprenants. Pour cela, il serait nécessaire de réaliser des captations vidéo des séances pédagogiques, en plus des acquisitions oculométriques et de tests de positionnement.

Bibliographie

N. A. Bradbury (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education.*

H. Jarodzka, I. Skuballa & H. Gruber (2020). Eye-Tracking in Educational Practice: Investigating Visual Perception Underlying Teaching and Learning in the Classroom. *Educational Psychology Review.*

H.Jarodzka, I. Skuballa & H. Gruber (2021). Eye-Tracking in Educational Practice : Investigating Visual Perception Underlying Teaching and Learning in the Classroom. *Educational Psychology Review.*

J. Przybyło, E. Kantoch & P. Augustyniak (2019). Eyetracking-based assessment of affect-related decay of human performance in visual tasks. Future Generation Computer Systems.

S. M. Shariff, Z. J. Johan & N. A. Jamil (2013). Assessment of Project Management Skills and Learning Outcomes in Students' Projects. *Procedia – Social and Behavioral Sciences,* 90.

J. Thomas & T. Mengel (2008). Preparing project managers to deal with complexity - Advanced project management education. *International journal of Project Management,* 26.

J. Wang, P. Antonenko & K. Dawson (2020). Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. *Computers & Education.*

F.-Y. Yang, C.-Y. Chang, W.-R. Chien, Y.-T. Chien & Y.-H. Tseng (2013). Tracking learners' visual attention during a multimedia presentation in a real classroom. *Computer & Education.*

Q. Zhao & C. Koch (2013). Learning saliency-based visual attention: A review. *Signal Processing.*

La place des jeux au sein de la formation

Élodie Delahaye, Jean Marc Ferrandi, Lionel Muniglia

Résumé

À partir d'une enquête auprès de 832 apprenants et 170 formateurs, nous caractérisons la place des dispositifs ludopédagogiques au sein de la formation ingénieur agroalimentaire. L'objectif est de comprendre les motivations, attentes et craintes sous-jacentes différenciées des deux types d'acteurs. Notre étude souligne toute l'importance de proposer une évaluation adaptée des dispositifs selon l'acteur. Elle permet de poser les bases des indicateurs à intégrer pour faciliter l'adhésion aux dispositifs ludiques, et évaluer leur efficacité en termes d'apprentissage.

INTRODUCTION

La ludification constitue une vraie source de motivation et de plaisir pour apprendre (Nemessany, 2020). Mais pour quels apprentissages serait-elle la plus bénéfique ? Ceux-ci varient-ils selon la cible choisie, formateurs ou apprenants ? Selon le contenu de la formation en elle-même ?

Au sein du programme d'investissement d'avenir 3 HILL (Hybrid Innovative Learning Lab), qui regroupe 25 partenaires, dont 16 établissements de l'enseignement supérieur français formant à l'éco-innovation alimentaire, nous avons conduit une enquête auprès de 832 apprenants et 170 formateurs en septembre 2021. Les objectifs : caractériser les freins et les moteurs à la ludification des apprentissages (façons de s'informer, de communiquer, de [se] former) ; et dresser un état des lieux des pratiques et attentes suite aux changements occasionnés par le virus du COVID-19 (et ses confinements) dans nos vies, ainsi que dans notre façon d'appréhender et de vivre la pédagogie. L'étude porte sur trois axes : notre familiarité avec les outils numériques, notre rapport au jeu, et notre désir d'intégration du jeu dans la formation. Les profils types dressés se différencient selon les apprentissages à ludifier, leurs leviers de motivation, les freins et les difficultés rencontrés par les acteurs.

Sur cette base, nous définirons les questions pertinentes à se poser pour développer des indicateurs, susceptibles de mesurer les facteurs favorables à l'acceptabilité des dispositifs ludiques, et d'évaluer la satisfaction à leur égard et leur efficacité.

I. LES FACTEURS FAVORABLES À L'ACCEPTABILITÉ DES DISPOSITIFS PÉDAGOGIQUES LUDIQUES

Nous avons tout d'abord identifié les facteurs qui facilitent l'acceptabilité des dispositifs pédagogiques. Ils sont à prendre en compte lors de leur conception, pour favoriser l'adhésion des apprenants et des formateurs impliqués. La mesure d'indicateurs dans ce sens durant la conception permet de prédire l'efficacité du dispositif. Pour ce faire, nous nous sommes appuyés sur le modèle CEPAJe, qui décrit des critères d'évaluation d'activités ludopédagogiques (Alvarez et Chaumette, 2017). Ce modèle est le fruit de diverses réflexions sur l'évaluation de tels dispositifs. De plus, il donne des pistes pour faciliter la conception ou l'animation des séquences ludopédagogiques. Tel est aussi le cas du guide élaboré par Motte *et al.* (2022), qui vise à élaborer et évaluer les activités ludopédagogiques. L'acteur, formateur ou apprenant, et son rapport au jeu se retrouvent au cœur du dispositif. C'est pourquoi

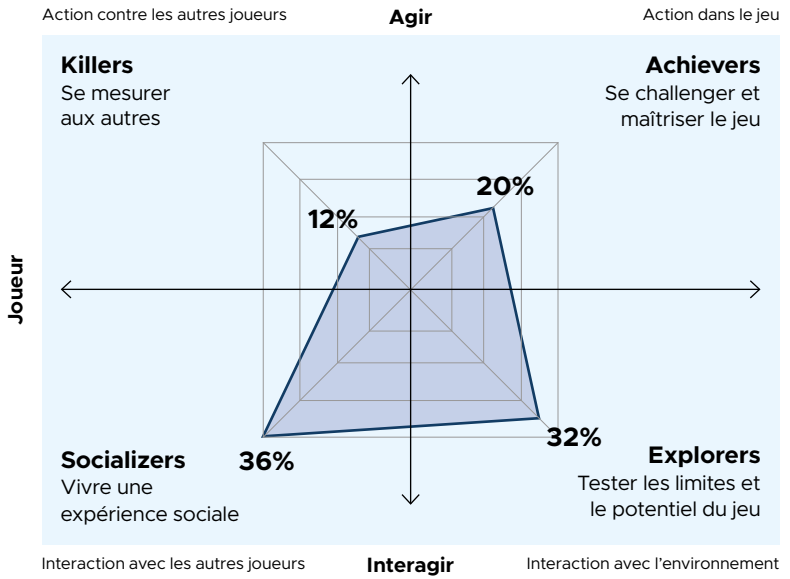
nous avons caractérisé leurs profils afin d'identifier des indicateurs qui leur sont propres.

1. FACE AU JEU, UNE DIVERSITÉ DE PROFILS

Les joueurs

Pour saisir les motivations au jeu, les apprenants et formateurs ont été segmentés selon le modèle de Bartle (1996), qui définit quatre profils autour de deux axes : « action-interaction » et « joueur-monde » (figure 1). Les deux groupes les plus fréquents sont les « sociaux » et les « explorateurs ». Ils sont tournés vers l'interaction avec les autres joueurs ou l'environnement proposé par le jeu. Pouvoir interagir avec les autres et l'environnement du jeu lors de la proposition d'une activité ludopédagogique est donc à intégrer, tout en n'oubliant pas qu'une partie des acteurs aiment des mécaniques d'action.

FIGURE 1 : Le profil joueur des répondants selon la typologie de Bartle (1996)



Cette typologie est construite sur un jeu particulier, le « multi-user dungeon ». Comment, alors, construire un jeu pédagogique prenant en compte les attentes et les motivations des quatre profils, dans la mesure où ils sont nécessaires à son bon fonctionnement ?

Des motivations variées

Désirant proposer des dispositifs adaptés à la diversité de profils des apprenants et des formateurs, nous nous sommes basés sur les modèles développés par Quantic Foudry et Marczewski (2015). Ces derniers permettent de saisir les motivations à pratiquer une activité ludique. Nous y avons ajouté l'apprentissage comme source de motivation afin de comprendre sa perception.

Les deux principales catégories de motivation à jouer (tableau 1) sont liées au divertissement, à l'expérientiel, à l'exploration, au partage et à l'interaction avec les autres. Ces résultats sont conformes à l'étude de Médiamétrie (2021) relative aux motivations des joueurs de jeux vidéo. La performance est secondaire, et l'apprentissage n'est qu'un incitateur mineur. Lors du développement d'un jeu pédagogique, savoir répondre aux différents moteurs de l'apprenant est un moyen de l'impliquer. Des univers de jeu bien construits et des mécaniques de jeu, qui font « sortir de la réalité » faciliteront l'acceptabilité des jeux. En outre, il existe une grande variété de pratiques ludopédagogiques, et avec elle, la capacité de s'adapter aux besoins et contraintes. À titre d'exemple, Pellon *et al.* (2020) présentent les dispositifs mis en place à l'université de Louvain. Ils sont appréciés par les apprenants et les formateurs, en raison de leur efficacité pour favoriser l'apprentissage. En un mot, la diversité de profils, en particulier si elle est mesurée dans la population visée, permet de proposer des outils adaptés, suscitant à la fois adhésion et motivation.

TABLEAU 1 — Les principales motivations à jouer

Motivation à jouer	**%**
Divertissement – Expérientiel Me divertir Explorer Passer le temps	86,3 28 24
Communication – Interaction avec les autres Être avec des amis Jouer avec d'autres personnes	62,3 22,3
Performance Maîtriser le jeu Gagner	24,6 10
Formation Apprendre	15,3

2. UN RAPPORT AU JEU DIFFÉRENT ENTRE LES APPRENANTS ET LES FORMATEURS

Les jeux coopératifs traditionnels sont préférés par les apprenants et les formateurs (tableau 2). En particulier, les apprenants, plus à l'aise avec les jeux numériques, préfèrent les utiliser par rapport aux formateurs (51 *vs* 46 %).

TABLEAU 2 — Les types de jeux préférés des apprenants et des formateurs (en %)

Préférence	**Jouer seul**	**Jouer en coopération**	**Jouer en compétition**
Apprenant	22	49	29
Formateur	19	59	22
Préférence	**Jeu numérique**	**Jeu traditionnel**	**Les deux**
Apprenant	12	49	39
Formateur	9	64	37

Au-delà de la diversité des profils, deux analyses en composantes multiples (figures 2 et 3) ont mis en évidence les rapports, préférences et motivations au jeu différenciés des apprenants et des formateurs. Les formateurs ont un profil de Bartle plutôt tourné vers le partage d'expérience (socializers). Ils préfèrent les jeux coopératifs (figure 2). Les dispositifs ludiques sont perçus comme un moyen de passer le temps. À l'inverse, les apprenants apprécient et sont à l'aise avec tous les types de jeux. Les explorateurs (explorers) et challengeurs (achievers) montrent une préférence pour les jeux en compétition. Ils trouvent leur motivation dans le divertissement et l'interaction sociale. Enfin, à la différence des apprenants, les formateurs sont plus familiers avec le jeu sérieux.

3. VERS DES INDICATEURS SPÉCIFIQUES AUX DISPOSITIFS NUMÉRIQUES

La compréhension de l'usage des supports et médias numériques complète nos clés de lecture pour mettre en place des dispositifs ludopédagogiques numériques. Les apprenants sont plus de 30 minutes par jour sur les médias numériques, dans plus d'activités et sur des supports différents. Ils passent plus de temps sur les réseaux sociaux, à regarder des vidéos en ligne, écouter de la musique ou communiquer, alors que les formateurs pratiquent plus la visioconférence et la recherche d'information (figure 3). En outre, les apprenants vont sur ces médias principalement pour se divertir et communiquer avec leurs proches, là où les formateurs les emploient pour communi-

quer avec d'autres personnes que leurs proches et pour s'informer. L'usage des médias numériques plus vaste chez les apprenants est tourné vers soi et ses proches. Il est plus utilitaire pour les formateurs.

4. LES BASES DES INDICATEURS À INTÉGRER POUR FACILITER L'ADHÉSION

Les points précédents nous invitent à poser plusieurs bases auxquelles pourraient répondre les indicateurs à intégrer pour faciliter l'adhésion des acteurs :

- prendre en compte la diversité des profils dans la proposition ludique grâce à la mesure de ces profils et ainsi, offrir des dispositifs accessibles au plus grand nombre ;
- proposer aux apprenants des jeux qui permettent de « sortir de la réalité » et « d'interagir avec les autres et l'environnement ». Donnons-nous suffisamment de place à l'expérientiel et à l'interaction ? Comment créons-nous de l'interaction entre les apprenants (joueurs) en favorisant une bonne entente entre eux (en imaginant des activités brise-glace préalables par exemple, s'ils ne se connaissent pas initialement) ? Comment apportons-nous du challenge pour davantage les engager grâce aux mécaniques du jeu, et pas uniquement par une simple victoire ? ;
- combiner coopération et compétition en proposant par exemple différents modes – compétition, entraînement, collaboration –, de façon à intégrer les différences d'attente entre les apprenants-joueurs ;
- adapter les indicateurs au type d'acteur, en prenant en compte les différences d'usage et d'acceptabilité du numérique entre apprenants et formateurs pour les jeux numériques ;
- mesurer l'adéquation du média numérique utilisé par rapport aux interfaces et habitudes, comme l'aisance perçue avec un casque de réalité virtuelle.

FIGURE 2 : Les préférences des apprenants et des formateurs

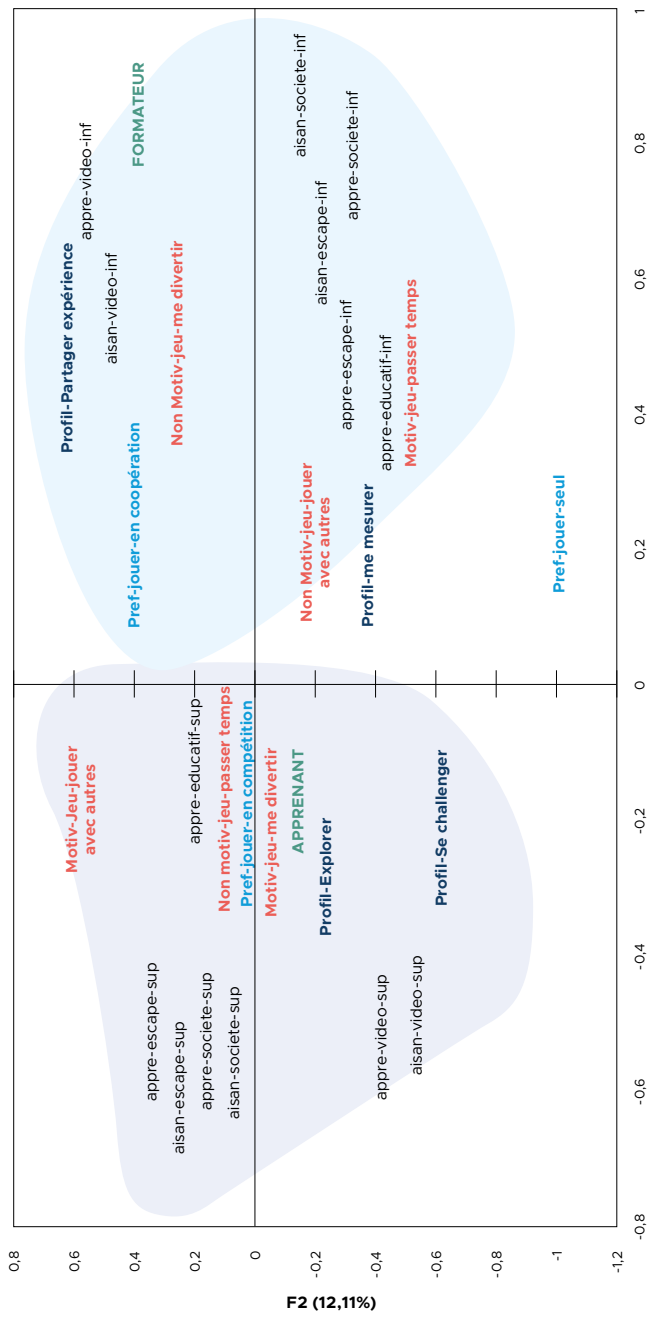
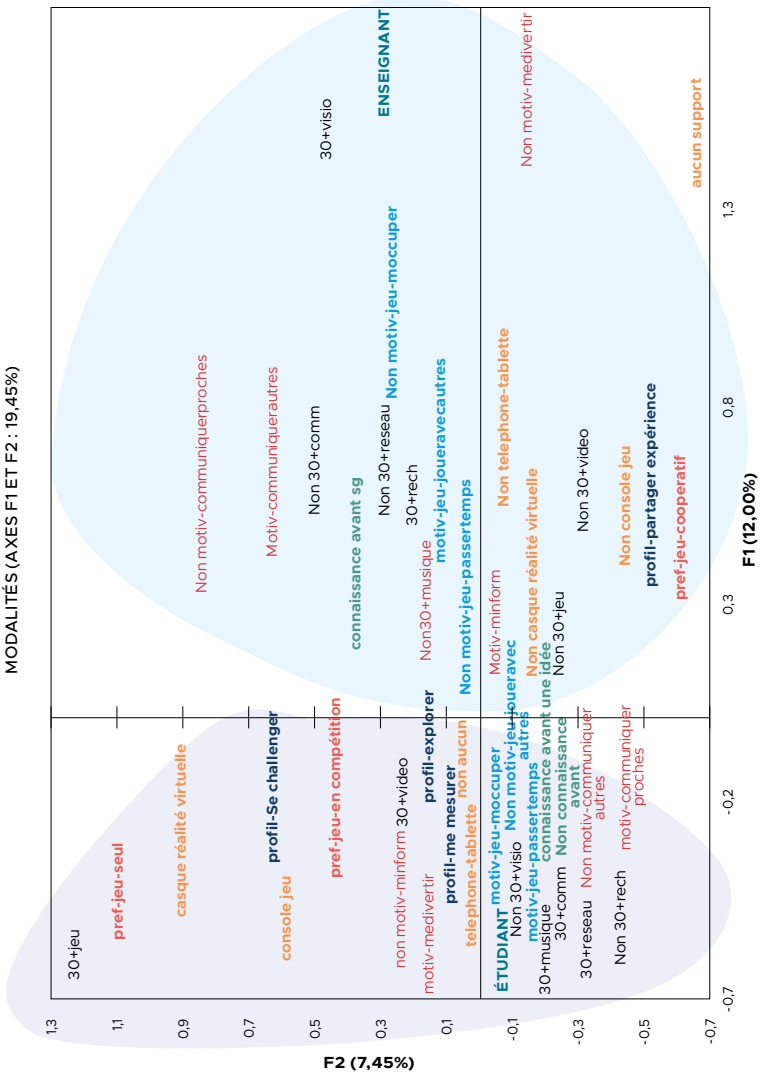


FIGURE 3 : Les profils des apprenants et des formateurs



Dans cette première partie, nous avons identifié un ensemble de préférences à ne pas négliger lors de la mise en place de systèmes ludopédagogiques, afin de faciliter l'adhésion des acteurs à ces dispositifs. Il s'agit alors de définir le degré d'accompagnement des cibles à ces changements. **Quelles capacités – opportunités – motivations à pratiquer et à adopter** (Michie *et al.,* 2014) – devons-nous mettre en œuvre pour que chacun accepte de manière durable des dispositifs pédagogiques novateurs ? Comment lever les résistances au changement ?

II. L'ÉVALUATION DE L'EFFICACITÉ DES DISPOSITIFS LUDIQUES

Dans un second temps, nous nous sommes penchés sur la pertinence de la ludification des dispositifs dans les différents champs d'application qui touchent les apprenants et les formateurs.

1. DE LA PERTINENCE ET DE L'APPRÉCIATION DE DISPOSITIFS LUDOPÉDAGOGIQUES

L'aspect ludique permet le développement d'habilités de coopération, de communication et de relations humaines, la motivation à l'apprentissage, la structuration des compétences, l'intégration de connaissances, et le développement d'habilités en résolution de problèmes (Sauvé *et al.,* 2007). Nemessany (2020) abonde en ce sens : « le jeu est un facteur de motivation et un vecteur de transmission de savoirs et de compétences ». Le média ludique a en effet le pouvoir de se transposer au monde réel, tout en proposant un environnement sécure pour le joueur (droit à l'erreur) et propice à l'apprentissage.

Les apprenants apprécient les jeux sérieux auxquels ils ont joué, et les formateurs, ceux qu'ils ont animés. Tous sont favorables à leur intégration plus large dans la formation. Ils apprécient grandement tous les types de jeux, et ils se sentent à l'aise (tableau 3). Cependant, les jeux vidéo requièrent un plus grand accompagnement des formateurs pour limiter leur appréhension.

TABLEAU 3 — Appréciation et aisance selon les types de jeu (en %)

Moyenne	Type de jeu	Jeux vidéo	Jeux de société, de plateau, de cartes	Escape games, jeux d'énigme et de logique	Jeux éducatifs pour apprendre, se former
Apprenant	**Appréciation**	63,2	87,5	82,8	69,3
	Aisance	63	87,8	81,3	75,8
Formateur	**Appréciation**	42,8	79,3	69,5	65,3
	Aisance	46,25	80,5	70,5	74,5

Pour apprendre et se former, nous ne notons pas de réticence ni d'appréhension pour les jeux éducatifs. La dimension ludique est un élément positif. Elle ne met pas en difficulté les deux publics. Toutefois, elle n'est pas forcément adaptée à tous les apprentissages. Quels seraient les domaines qui bénéficieraient le plus à une ludification de la formation ?

2. DES DOMAINES PERTINENTS POUR L'APPRENTISSAGE

Pour nous focaliser sur les apprentissages qui bénéficieraient le plus d'une ludification lors de la création de nouveaux outils, nous avons caractérisé les formes d'apprentissage qui posent le plus de difficultés aux apprenants et aux formateurs (tableau 4).

TABLEAU 4 — Difficultés perçues d'apprentissage des apprenants (en %) et intérêt de la ludification pour l'apprentissage (en %)

Apprenants			Formateurs		
	Difficultés	**Intérêt**		**Difficultés**	**Intérêt**
Appliquer des concepts à des cas concrets	**48,4**	52,6	Appliquer des concepts à des cas concrets	**63,7**	**54,4**
Autoévaluer ses apprentissages		**39,5**	Autoévaluer ses apprentissages		**36,3**
Savoir prendre des décisions dans des environnements complexes	**44**	**35,8**	Comprendre l'utilité des apprentissages	**52,1**	**29,8**
Comprendre l'utilité des apprentissages	**40,9**	**31,3**	Mobiliser des compétences transversales	**43,3**	**43,3**
Acquérir des connaissances de cours	30,2	36,4	Savoir prendre des décisions dans des environnements complexes	**42,7**	**41,5**
Mobiliser des compétences transversales	**26**	**33,3**	Acquérir des connaissances de cours	28,1	15,2
Appliquer des concepts dans des exercices	23,3	24,4	Appliquer des concepts dans des exercices	27,5	15,8
Évaluer les apprentissages		8,1	Évaluer les apprentissages		8,2
Aucun intérêt		2,4	Aucun intérêt		5,3

Selon les apprenants et les formateurs, les apprenants éprouvent des difficultés d'apprentissage, même si ce n'est pas toujours avec la même intensité. Par exemple, « savoir prendre des décisions dans des environnements complexes » semble aussi difficile pour les apprenants que pour les formateurs. Alors que les trois autres difficultés principales sont largement plus mises en valeur par les formateurs. Globalement, les principales difficultés d'apprentissage portent sur la concrétisation de concepts abstraits. Comment alors réduire la distance psychologique perçue par l'apprenant entre la connaissance et sa mise en application ?

De manière complémentaire, nous avons identifié les domaines perçus comme étant pertinents à ludifier. Les difficultés principales ressenties par les apprenants et les formateurs se retrouvent dans les facteurs favorables à la ludification. L'auto-évaluation des apprentissages est aussi un facteur à fort potentiel d'intérêt. Les propositions d'intérêt englobent donc les principales difficultés identifiées :

- appliquer des concepts à des cas concrets ;
- prendre des décisions dans des environnements complexes ;
- comprendre l'utilité des apprentissages ;
- mobiliser des compétences transversales.

Au final, le jeu pourrait être une réponse adaptée et appréciée à la réduction de la distance psychologique entre l'apprenant et la réalité de ses apprentissages, en facilitant la concrétisation des concepts abstraits (Trope et Liberman, 2003). Pour autant, il reste à s'assurer que l'apport du jeu n'est pas terni par les contraintes qui entourent l'apprentissage. C'est pourquoi nous nous sommes intéressés aux freins/craintes à la ludification.

3. DES CRAINTES À SURVEILLER

La principale crainte perçue quant à l'utilisation de jeux sérieux dans la formation, pour près de 50 % des acteurs, porte sur la non-acquisition ou la non-transmission des compétences, connaissances. En d'autres termes, des apprentissages (tableau 5). Pour réduire ce risque de détournement de l'objectif principal de la formation, il semble nécessaire de bien poser les objectifs pédagogiques du jeu, et de construire l'outil autour de ces objectifs. Il s'agit d'assurer la bonne transmission/acquisition des apprentissages. De plus, une étape de débriefing semble centrale pour contrôler le bon déroulement et l'atteinte des objectifs fixés.

La seconde crainte, pour près du tiers des acteurs, est liée aux problèmes techniques anticipés. Ce risque peut être réduit en testant en amont le matériel, en démontrant la fiabilité du système, et en créant un guide utilisateur. Ensuite, les craintes des apprenants et formateurs divergent. Les appre-

nants ont peur de se retrouver en situation d'échec, ne pas réussir, ou ne pas comprendre le jeu. Réduire la distance psychologique sociale (les autres apprenants réussissent et comprennent) par l'exemple pourrait constituer un canal de rassurance, tout en apportant des explications claires, faciles, non ambiguës. Enfin, pour les formateurs, deux autres craintes sont significatives : perdre son temps à mettre en place un outil, et produire un outil pas assez durable. Ceci invite à s'interroger sur la manière d'accompagner le formateur dans la mise en place du dispositif, en démontrant tous ses bénéfices face aux éventuels coûts (obsolescence, maintenance par exemple).

TABLEAU 5 — Craintes à l'utilisation des jeux sérieux dans la formation (en %)

Apprenants		Formateurs	
Ne pas acquérir les compétences/connaissances	**48,1**	Ne pas transmettre les apprentissages	**49,1**
Ne pas réussir	**33,2**	Perdre du temps (mise en place)	**48,5**
Problèmes techniques	**31,4**	Problèmes techniques	**43,9**
Ne pas comprendre	**31,1**	Outil peu durable	**30,4**
Perdre son temps	27,3	Perdre du temps (en classe)	27,5
Ne pas trouver	25,5	Activité perçue non sérieuse	16,4
Ne pas être à l'aise	24,8	Pas de crainte	9,9
Pas de craintes	9,4	Perdre le contrôle de l'outil	9,4
Autre	2	Autre	5,9

4. LES QUESTIONS À SE POSER POUR DÉFINIR LES INDICATEURS À INTÉGRER POUR ÉVALUER L'EFFICACITÉ DES DISPOSITIFS LUDIQUES

Différentes questions préalables à la définition des indicateurs d'évaluation de l'efficacité des dispositifs ludopédagogiques peuvent être soulevées :

- L'apprentissage bénéficierait-il d'une ludification ? La ludification est-elle adaptée à l'acquisition des connaissances – compétences en jeu ? Pour y répondre, une analyse coût/bénéfice de la ludification semble un prérequis.
- Comment réussir à réduire la distance psychologique envers la ludification (Trope et Liberman, 2010) ? Quels moyens sont mis en œuvre

pour faciliter la concrétisation de concepts et/ou connaissances abstraits ?

- Comment réduire les freins : risques de problèmes techniques, coûts temporel et financier perçus, difficulté d'appréhension à cause de l'interface ou du matériel, durabilité de l'outil… ?
- Les apprentissages sont-ils bien transmis ? Les dispositifs ont-ils permis une meilleure réussite aux examens ?
- Quels accompagnements des acteurs dans la mise en œuvre et dans les mises à jour des dispositifs ont été prévus ?

Ces questions peuvent être posées soit pour caractériser la qualité d'un dispositif existant ; soit lors de la conception d'un dispositif, afin de prendre en compte le maximum de facteurs d'amélioration.

CONCLUSION

Plusieurs constats sont à retenir :

- La littérature souligne tout l'intérêt de varier les dispositifs pédagogiques pour s'adapter aux différents profils, notamment d'apprenants, et « raccrocher » ceux qui seraient en échec ou en difficulté.
- Cet article appréhende les bases et les questions à se poser pour évaluer l'efficience de dispositifs ludopédagogiques, et pour faciliter leur acceptabilité par les formateurs et les apprenants.
- Au terme d'une étude menée auprès de 1 002 formateurs et apprenants, la ludification est sans ambiguïté bien acceptée. Des facteurs favorables au développement des outils ludopédagogiques sont présents. Apprenants et formateurs montrent une maîtrise certaine des outils et des médias, et de réelles motivations à leur adoption. Toutefois, les moteurs et les freins ne sont pas toujours similaires entre les deux populations, et doivent être intégrés dès la conception des dispositifs. En particulier, est mise en lumière la nécessité d'un accompagnement plus grand des formateurs, surtout dans le cas de développement de dispositifs numériques, et pour lever les craintes d'une obsolescence trop rapide. Ceci demande de faire preuve de « pédagogie », de démonstration, en adoptant une approche différenciée entre formateurs et apprenants pour les persuader. Afin d'assurer ce changement de perspective, nous avons à nous interroger sur les capacités/opportunités à mettre en œuvre, et sur les motivations à susciter (Michie *et al.,* 2014), à même de réduire la dis-

tance psychologique entre l'acteur et l'objet de sa formation (Trope et Liberman, 2010), pour qu'il soit capable de mettre en œuvre les connaissances acquises.

Bibliographie

Andrzej Marczewski (2015), « User Types HEXAD », in *Even Ninja Monkeys Like to Play: Gamification, Game Thinking and Motivational Design,* 65-80.

Gaëlle Pellon, Benoit Raucent, Thibault Philipette, Céline Mathelart, Julian Alvarez, Nicolas Kervyn de Meerandré, Fanny Cambier, Pascal Vangrunderbeeck, Isabelle Motte, Emilie Malcourant & Virginie Renson (2020), *Jouer pour apprendre dans l'enseignement supérieur,* Cahiers de l'université catholique de Louvain, 8.

Isabelle Motte, Pascal Vangrunderbeeck, Manuela Guisset, Emilie Malcourant & Virginie Renson, *Guide pour la conception d'une activité ludopédagogique.*

Julian Alvarez, Pascal Chaumette (2017), « Présentation d'un modèle dédié à l'évaluation d'activités ludopédagogiques et retours d'expériences », in *Recherche et pratiques pédagogiques en langues de spécialité* [Online], 36, 2.

Louise Sauvé, Lise Renaud, Mathieu Gauvin (2007), « Une analyse des écrits sur les impacts du jeu sur l'apprentissage », *Revue des sciences de l'éducation,* 33, 1, 89-107.

Médiamétrie, Syndicat des éditeurs de logiciels de loisirs, *L'Essentiel du jeu vidéo,* Novembre 2021.

Richard Bartle (1996), «Hearts, clubs, diamonds, spades: players who suit MUDs», *Journal of MUD Research*, 19.

Susan Michie, Robert West, Rona Campbell, Jamie Brown, Heather Gainforth (2014), *ABC of Behavior Change Theories,* Silverback Publishing.

Violette Nemessany (2020), « La pédagogie du jeu : comment remettre les apprentissages en jeu ? », *Rapport de synthèse, ANRT,* Cahiers Futuris.

Yaacov Trope, Nira Liberman, Temporal construal (2003), *Psychological Review,* 110, 3, pp. 403–421.

Yaacov Trope, Nira Liberman (2010), « Construal-level theory of psychological distance », *Psychological Review,* 117, 2, pp. 440–463.

« Sciado Tacker » : un module pour le suivi des apprenants dans les simulateurs

Nesrine Rahmouni, Laurent Lavogiez

Résumé

Les simulateurs sont des environnements prometteurs pour l'apprentissage grâce à leurs nombreux avantages, tels que la visualisation interactive, la simulation de situations réelles, et l'apprentissage par la pratique. Cependant, pour une meilleure utilité, il est nécessaire de travailler sur la personnalisation de l'apprentissage pour répondre aux besoins individuels de chaque apprenant. Dans cet article, nous avons proposé un module pour permettre aux formateurs d'obtenir des informations sur le processus d'apprentissage de chaque apprenant, afin de lui offrir une expérience adaptée. Ce module a fait l'objet d'une première évaluation. Celle-ci a montré des résultats satisfaisants et encourageants.

INTRODUCTION

Dans des domaines tels que l'ingénierie, la gestion, la finance et la médecine, les apprenants sont fréquemment confrontés à des systèmes complexes de la vie réelle. Il peut s'agir de systèmes d'infrastructure, de systèmes informatiques, de réseaux ou de chaînes d'approvisionnement, qui impliquent des milliers de variables et d'interactions complexes. Pour pouvoir prendre des décisions éclairées qui affectent ces systèmes, il est crucial que les apprenants reçoivent une formation spécialisée. Cependant, les systèmes de la vie réelle sont souvent soumis à l'incertitude et au dynamisme, ce qui les rend difficiles à modéliser avec les méthodologies d'apprentissage traditionnelles. C'est pourquoi, pour prendre en compte ces facteurs, il est souvent nécessaire d'utiliser des méthodologies basées sur la simulation, en coordination avec des techniques d'apprentissage automatique. Les méthodes d'apprentissage basées sur la simulation permettent de refléter des situations réelles avec une grande précision, stimulant ainsi les compétences de l'apprenant et sa compréhension du système, en montrant les conséquences de ses actions et décisions. De plus, ces actions et décisions peuvent être testées sans crainte d'échec.

Dans notre travail de recherche, nous nous intéressons particulièrement au domaine du management et de la gestion d'entreprise. Dans ce cadre, les simulateurs d'entreprise, appelés aussi jeux d'entreprise, constituent une méthode d'apprentissage en ligne pertinente. Ils ont connu, à l'échelle mondiale, un développement progressif dans les écoles de commerce, les universités et les associations professionnelles. Les simulateurs en ligne sont une alternative efficace aux méthodes d'enseignement traditionnelles. En effet, ils relient les concepts abstraits aux réalités du monde, offrent une dimension pratique supplémentaire à l'apprentissage, et donnent aux apprenants l'opportunité de se perfectionner dans la prise de décision dans des situations concrètes.

Toutefois, comme dans la plupart des environnements d'apprentissage en ligne, les simulateurs s'adressent à des apprenants dont le rythme d'apprentissage, les besoins ou les motivations ne sont pas similaires. Pour permettre à chaque apprenant de développer progressivement ses compétences à un rythme qui lui convient, il est nécessaire de personnaliser le parcours d'apprentissage proposé. Cela peut être accompli en adaptant à la fois les éléments pédagogiques et les supports, par exemple, les commentaires des formateurs et les débriefings. Afin de garantir la pertinence pédagogique des activités et des supports, il semble important de fournir aux formateurs les outils nécessaires pour suivre les progrès des apprenants dans le simulateur, et pour visualiser en temps réel leurs besoins, leurs lacunes et leurs acquis.

Au sein de l'entreprise Sciado Partenaires, nous travaillons pour développer des simulateurs pédagogiques dans divers domaines en lien avec l'entreprise. Dans cet article, nous présentons notre simulateur « Appro », conçu pour aider les apprenants à améliorer leurs compétences en gestion des approvisionnements de stock. Nous avons intégré à ce simulateur un module appelé « Sciado Tracker », qui permet aux formateurs de suivre en temps réel les activités de chaque apprenant. Nous décrirons la conception de ce module en détail plus tard dans l'article. Dans la section suivante, nous présentons les simulateurs les plus utilisés dans ce domaine, en mettant en lumière leurs atouts et leurs faiblesses respectifs. Il convient de préciser que dans la suite de cet article, nous avons choisi d'utiliser le terme « simulateur pédagogique » pour désigner les simulateurs en lien avec l'entreprise.

I. TRAVAUX CONNEXES

Dans le champ de la gestion de chaîne d'approvisionnement, les simulateurs sont généralement conçus pour aborder les aspects de planification, production, inventaire et contrôle des coûts. La planification de la production vise à garantir que l'entreprise simulée fournit les produits nécessaires en quantité suffisante. Ceci nécessite la formation des apprenants à élaborer des plans de production appropriés, qui prennent en compte les incertitudes inhérentes à cette discipline. Dans le paragraphe suivant, nous présentons certains des simulateurs les plus cités dans ce domaine.

1. DES SIMULATEURS POUR LA GESTION DES CHAÎNES D'APPROVISIONNEMENT

Le jeu « Beer Game » est un ancien jeu de simulation très populaire. Il a d'abord été présenté sous la forme d'un jeu de société non informatisé. Par la suite, il est devenu un jeu de simulation numérique pour ordinateur, utilisé pour explorer le processus de la chaîne d'approvisionnement, et former les joueurs à gérer les différents problèmes, comme le manque de partage d'informations entre les parties de la chaîne, et la mauvaise planification. Le succès de ce jeu dans l'enseignement a donné lieu à de nouvelles versions modifiées, telles que « Beer Game 2.0 ». Ce dernier inclut deux grossistes, pour rendre le jeu plus complexe. Bien que le Beer Game soit un outil d'apprentissage utile pour comprendre les principes de base de la gestion de la chaîne d'approvisionnement, il présente comme limite son manque d'adaptabilité. En effet, le format de ce jeu est rigide. Il ne peut pas être facilement adapté aux situations spécifiques de la chaîne d'approvisionnement, ou aux besoins d'apprentissage individuels.

Il existe un autre simulateur, similaire à « Beer Game », intitulé « LOGI-Game ». Ce dernier simule un réseau de distribution de différents produits, permettant aux participants de gérer une entreprise faisant partie de ce réseau. Le jeu permet aux apprenants de s'entraîner à faire face à des événements imprévus qui peuvent survenir pendant la partie, comme des grèves, des fluctuations monétaires internationales, des retards de traitement de certaines commandes. Ce jeu est basé sur Excel, ce qui le rend très facile à modifier ou à mettre à jour. En revanche, LOGI-Game est conçu pour enseigner à un groupe d'apprenants. Il ne peut pas être facilement modifié pour répondre aux besoins spécifiques de chaque individu.

Sur le même principe qu'une feuille de calcul Excel, citons les simulateurs « SBELP » et « RSS-POD ». « SBELP » simule une chaîne d'approvisionnement où le joueur est un fabricant d'équipements électroniques. Même avec ses nombreuses caractéristiques, dont la gestion des stocks et des commandes en attente, son objectif principal vise à illustrer pour les joueurs l'impact de l'effet coup de fouet dans la chaîne d'approvisionnement. En ce qui concerne « RSS-POD », il recrée un inventaire des médicaments, attribuant au joueur un rôle de gestionnaire de stocks d'articles médicaux. Le but est de gérer la distribution de ces articles médicaux dans le cas d'une urgence sanitaire, tout en minimisant l'attente des clients, et en allouant les articles à plusieurs sites d'utilisation.

Les simulateurs en ligne pour l'apprentissage offrent plusieurs avantages, dont notamment la capacité à simuler des situations complexes, et à permettre aux apprenants de mettre en pratique les compétences acquises. Il n'en demeure pas moins que la personnalisation de l'apprentissage est peu présente dans ces environnements. Les simulateurs sont destinés à des apprenants présentant diverses caractéristiques, incluant des rythmes d'apprentissage variés ainsi que des besoins et motivations différents. Cette diversité peut parfois rendre le parcours d'apprentissage proposé peu adapté aux besoins spécifiques de certains apprenants. Afin de garantir un apprentissage optimal pour chacun, il est nécessaire de personnaliser le parcours d'apprentissage en adaptant les éléments pédagogiques et les supports en fonction des besoins, acquis et motivations de chacun. En effet, le manque de personnalisation peut rendre l'expérience d'apprentissage moins efficace pour certains apprenants, et peut également dégrader leur engagement.

2. NOTRE PROPOSITION

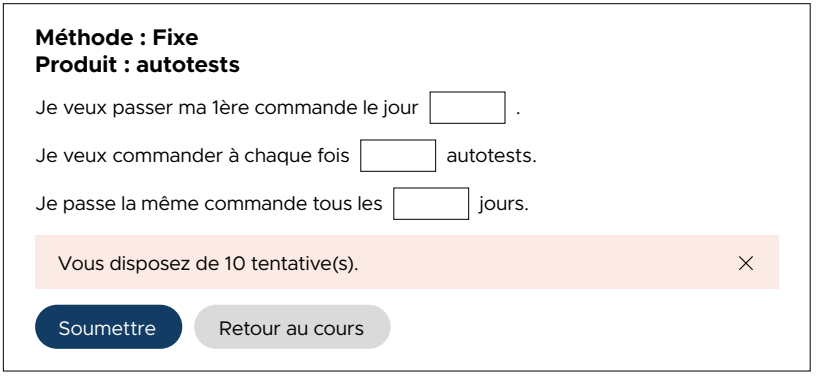
Notre travail de recherche se concentre sur la création de simulateurs pédagogiques visant à renforcer les compétences dans divers domaines, tels que la gestion d'entreprise, le marketing, le management ou la logistique. Dans cet article, nous présentons notre simulateur « Appro ». Il permet de simuler l'approvisionnement de différents produits, par exemple, des équipements médicaux, des produits alimentaires et des équipements informatiques.

Ce simulateur vise à permettre aux utilisateurs de développer des compétences en matière d'optimisation de l'approvisionnement. Nous avons identifié trois compétences susceptibles d'être évaluées :

- **approvisionner par la méthode fixe (C1) :** cette méthode consiste à commander un nombre constant de produits à intervalles réguliers, indépendamment de la demande réelle ;
- **approvisionner par la méthode du point de commande (C2) :** elle consiste à définir un seuil de stock minimum à partir duquel une commande doit être passée ;
- **approvisionner par la méthode de recomplètement (C3) :** qui implique de commander suffisamment de produits pour atteindre un niveau de stock maximum prédéfini.

Les utilisateurs peuvent faire l'expérience de plusieurs situations professionnelles (parcours d'activités), qui les amènent à passer des commandes de produits tout en prenant en compte différents paramètres, en fonction de la méthode d'approvisionnement choisie. Par exemple, pour la méthode fixe, chaque apprenant doit préciser dans sa commande, en plus de la quantité de produit commandé, le premier jour de la commande et la durée entre chaque commande (figure 1).

FIGURE 1 — Activité d'approvisionnement par la méthode fixe



À la fin de chaque décision, les joueurs peuvent voir les conséquences de leurs choix sur la chaîne d'approvisionnement en termes de coûts, tels que :

- **Le coût de possession :** le coût pour l'entreprise pour posséder et maintenir un stock de produits ou de matières premières,
- **Le coût de rupture :** le coût associé à la rupture d'un stock ou à l'indisponibilité d'un produit ou d'une matière première dans la chaîne d'approvisionnement,

- **Le coût de lancement :** le coût associé au lancement d'une commande à un fournisseur dans la chaîne d'approvisionnement (traitement administratif).

À la fin de chaque activité, l'apprenant obtient un score sous forme de pourcentage, qui représente la performance de l'apprenant dans une activité par rapport à une valeur optimale définie. Il est calculé en s'appuyant sur les trois coûts : les coûts de possession, de rupture et de lancement.

Un débriefing est effectué avec tous les joueurs à la fin de chaque simulation, pour analyser les différentes décisions prises, et corriger les perceptions erronées. Cependant, pour aider les apprenants à développer leurs compétences et à renforcer leurs connaissances, il est important d'aller au-delà d'un simple débriefing. Nous cherchons à accompagner chaque apprenant dans son parcours en l'adaptant à son rythme d'apprentissage, ses acquis et ses besoins, tout en favorisant son autonomie.

Pour ce faire, nous avons mis en place un module de suivi de l'activité de l'apprenant dans le simulateur. Ce module permet aux formateurs de mesurer l'acquisition des compétences de chaque apprenant, de comprendre son rythme d'apprentissage, de visualiser ses besoins et ses difficultés à l'aide d'un tableau de bord, et ainsi d'adapter le parcours pour une meilleure expérience d'apprentissage.

Dans la suite de cet article, nous détaillons le processus de conception de ce module de suivi, son fonctionnement, les informations visualisées par ce module, et enfin, les résultats de l'expérimentation réalisée pour son évaluation.

3. LE MODULE « SCIADO TRACKER »

Ce module consiste à suivre l'activité des apprenants dans les simulateurs. Tout d'abord, il est conçu pour les formateurs. C'est pourquoi le processus de conception s'appuie sur une collaboration étroite entre les formateurs, les développeurs et les pédagogues. Nous résumons ce processus comme suit :

Identification des besoins des utilisateurs

L'identification des besoins des utilisateurs est une étape cruciale dans la conception de « Sciado Tracker ». Il s'agit de déterminer, clarifier, affiner et hiérarchiser les besoins des utilisateurs du module (à savoir, dans ce cas, les formateurs). Pour atteindre cet objectif, nous avons mené des réunions avec des experts pour discuter de leurs besoins et souhaits. Nous avons également utilisé des questionnaires pour obtenir des informations plus détaillées. L'objectif principal de cette étape est de déterminer les traces pertinentes de l'apprenant dans le simulateur, afin que les formateurs puissent mieux les comprendre, et les accompagner de manière adaptée. Par exemple, tous les

experts ont montré le besoin d'obtenir des informations sur les apprenants en difficulté afin de les aider à progresser, ce qui nous a amenés à intégrer cette fonctionnalité dans le module.

Formalisation des besoins des utilisateurs

Il s'agit de formaliser et de modéliser les besoins des formateurs, en utilisant des maquettes et des diagrammes pour une compréhension plus claire et concise de leurs attentes. Cette étape de formalisation est indispensable pour garantir que la solution développée répond réellement à leurs besoins et satisfait leurs attentes. Elle permet également de communiquer de manière claire les fonctionnalités proposées, les processus impliqués et les objectifs à atteindre.

Développement de la solution

Cette phase consiste à traduire les maquettes et les diagrammes formalisant les besoins en un logiciel fonctionnel. La première étape de ce développement vise à collecter les traces des apprenants dans le simulateur. Pour ce faire, nous avons défini des méthodes de collecte automatisées de ces traces à partir de différentes sources, comme la plateforme « Moodle ». Cette dernière nous permet de suivre les activités et le parcours des apprenants de manière centralisée et structurée, ou à travers des capteurs intégrés au simulateur. Après la collecte des traces de l'apprenant dans le simulateur, nous avons commencé la phase d'analyse. Dans cette phase, nous avons développé des fonctions permettant de déterminer les informations sur le parcours de l'apprenant dans le simulateur, que nous décrirons dans les sous-sections suivantes. Nous avons choisi de les présenter aux formateurs sous forme de tableau de bord. Cette approche aide les formateurs à mieux comprendre les besoins de l'apprenant, et à l'accompagner de manière plus adaptée. Par exemple, si la représentation de la progression d'un apprenant montre régulièrement une faible acquisition pour une compétence donnée, les formateurs pourront utiliser cette information pour adapter leur approche pédagogique en conséquence. Dans ce cas, ils pourront proposer des exercices de renforcement sur cette compétence spécifique, ou encore mettre en place des rappels réguliers pour que l'apprenant puisse s'exercer davantage.

Évaluation de la solution

Il s'agit de l'évaluation de la solution proposée pour s'assurer de son efficacité et de sa pertinence. Cette phase consiste à mettre en œuvre la solution à travers des expérimentations, et à collecter des données pour évaluer ses performances. En outre, l'objectif est de déterminer si la solution répond aux besoins des utilisateurs, et s'il y a des parties à améliorer.

Nos experts considèrent l'évaluation de l'état des compétences et la détection des difficultés de l'apprenant comme étant des informations cruciales

pour l'adaptation personnalisée de son parcours d'apprentissage. Dans la section suivante, nous allons détailler le processus de notre module « Sciado Tracker » pour identifier ces informations importantes.

II. REPRÉSENTATION DES INFORMATIONS AUX FORMATEURS

Les informations déterminées par ce module sont présentées aux formateurs via un tableau de bord en temps réel par différentes représentations graphiques (figure 2).

Ce tableau de bord comporte plusieurs vues, selon les informations représentées :

- **vue générale :** elle regroupe des informations générales sur un groupe d'apprenants, telles que le taux de réussite dans le parcours, le taux de réussite pour chaque activité ou chaque compétence visée, ainsi que les informations sur la corrélation entre le taux de réussite et les différents objets pédagogiques proposés.
- **progression de l'apprenant :** cette vue présente des informations sur le parcours de chaque apprenant dans le simulateur, comme son taux de réussite, l'évolution de ses compétences, etc.
- **apprenant en difficulté :** cette vue affiche la liste des apprenants en difficulté dans le parcours, en précisant le type de difficulté détecté, et d'autres informations supplémentaires, telles que la date de la détection et les compétences concernées par cette difficulté.

Il est important de noter que le choix des représentations graphiques est basé sur le tableau des suggestions graphiques « A Thought Starter », qui permet de choisir la représentation graphique appropriée pour chaque information en fonction de l'objectif.

Le tableau de bord présente des informations détaillées sur les divers parcours ainsi que les compétences cibles. Grâce à la vue référentiel, les formateurs peuvent visualiser les compétences que nous visons pour les apprenants après l'utilisation du simulateur.

FIGURE 2 : la vue générale du tableau de bord

1. DÉTERMINATION DE L'ÉTAT DE COMPÉTENCES DE L'APPRENANT

Dans notre proposition, nous avons mis en place un système de représentation des compétences visées par le simulateur (C1, C2, C3) en utilisant un référentiel de compétences développé par des experts du domaine et des pédagogues. Ce référentiel a été élaboré pour couvrir les compétences nécessaires pour exceller dans le domaine de la gestion des approvisionnements. La conception de ce référentiel s'appuie sur la taxonomie de Bloom. Chaque compétence du référentiel est mise en pratique à travers une ou plusieurs activités pédagogiques. Le score obtenu à la fin de chaque activité nous fournit une indication sur son état de compétence, qui est classé en quatre niveaux : 1) non acquis ; 2) en cours d'acquisition ; 3) acquis ; 4) maîtrisé. Les informations ainsi obtenues permettent de suivre l'évolution de la compétence de chaque apprenant au fil du temps. Cela est indispensable pour adapter le parcours d'apprentissage à ses besoins, et lui permettre de progresser efficacement.

Pour améliorer la précision de ces informations, nous avons développé une fonction qui détermine l'état de chaque compétence, ainsi qu'un indice de certitude associé à cet état. Ce processus s'appuie sur l'analyse de différentes traces de l'apprenant dont, notamment :

- le score obtenu à la fin de l'activité ;
- le nombre de tentatives réalisées pour l'activité ;
- la progression de la performance de l'apprenant entre les tentatives, qui peut être :

  - positive : renvoie au cas d'une progression positive entre les scores de performance,
  - négative : renvoie au cas d'une progression négative entre les scores de performance,
  - aléatoire : renvoie au cas d'une progression positive et négative entre les scores de performance ;

- la progression de la performance de l'apprenant entre les activités qui mettent en pratique la même compétence. Cette progression peut être aussi :
  - positive : renvoie au cas d'une progression positive entre les scores de l'apprenant entre les activités concernées,
  - négative : renvoie au cas d'une progression négative entre les scores de l'apprenant entre les activités concernées,
  - aléatoire : renvoie au cas d'une progression positive et négative entre les scores de l'apprenant entre les activités concernées ;

- le nombre et la durée de consultation des cours pendant l'activité associée.

En utilisant ces informations, nous sommes en mesure de déterminer avec plus de précision l'état de compétence de chaque apprenant (figure 3).

FIGURE 3 : Partie de la fonction de détermination de l'état de compétences de l'apprenant

Fonction **(Détermination de l'état de compétence de l'apprenant)**
Etat de compétence {Non acquise « NA », Acquise « A », En cours d'acquisition « ECA », Maîtrisée « M »)
Score [0, 100]
Nombre des tentatives (NT) [0, 20]
Nombre des activités traitées (NA) [1, 3]
Progression entre les tentatives (PT) [positive, négative, aléatoire]
Progression entre les activités (PA) [positive, négative, aléatoire]
Certitude [faible, moyenne, forte]
Début:
Si **(Score > 60)** et **(NT = 1)** et **(NA > 2)** et **(PT =** /) **et (PA=** positive) alors **Etat = Maîtrisée**
Sinon si **(Score > 60)** et **(NT <= 3)** et **(NA > 1)** et **(PT =** positive) **et (PA=** positive) alors **Etat = acquise** avec **IC = fort**
Sinon si **(Score > 60)** et **(NT <= 3)** et **(NA > 1)** et **(PT =** aléatoire) **ou (PA=** aléatoire) alors **Etat = acquise** avec **IC = moyen**
Sinon si **(Score > 60)** et **(NT <= 3)** et **(NA >1)** et **(PT =** négative) **ou (PA=** négative) alors **Etat = acquise** avec **IC = faible**
Sinon si **(Score > 40)** et **(NT > =1)** et **(NA > =1)** et **(PT =** positive) **ou (PA=** positive) alors **Etat = en cours d'acquisition** avec **IC =fort**
Sinon si **(Score > 40)** et **(NT > =1)** et **(NA > =1)** et **(PT =** aléatoire) **ou (PA=** aléatoire) alors **Etat = en cours d'acquisition** avec **IC = moyen**
Sinon si **(Score > 40)** et **(NT > =1)** et **(NA >=1)** et **(PT =** négative) **ou (PA=** négative) alors **Etat = en cours d'acquisition** avec **IC = faible**

2. LA DÉTECTION DES APPRENANTS EN DIFFICULTÉ

Dans notre proposition, nous avons compris l'importance d'évaluer non seulement les compétences acquises par les apprenants, mais aussi les défis auxquels ils sont confrontés, pour les aider à les surmonter. Pour ce faire, nous avons mis en place une fonction de détection des difficultés rencontrées par les apprenants.

Dans notre système de simulation, nous avons défini deux types de difficultés :

- **La faible performance.** Elle est observable à travers un état de compétence majoritairement « non acquis ». Par exemple, lorsqu'un apprenant n'a acquis qu'une seule compétence sur trois après avoir participé à des activités qui les mettent en pratique.
- **L'incohérence cognitive.** Elle se manifeste par une incohérence dans le comportement de l'apprenant dans notre simulation. Par exemple, lorsque nous examinons **N** comme la dernière activité réalisée par un apprenant, et **N-1** comme l'activité précédente par rapport à **N**, ces deux activités mettent en œuvre la même compétence.
- **Si** l'état (N) est « non acquise avec indice de certitude fort » ou « en cours d'acquisition avec indice de certitude faible » **ET Si** l'état (N-1) est « acquise avec indice de certitude fort », **alors** l'apprenant est considéré en difficulté de type incohérence cognitive.

La fonction de détection des difficultés prend en entrée les scores de l'apprenant, les activités réalisées, l'état de ses compétences, et produit une liste des apprenants en difficulté, en précisant le type de difficulté pour chacun. Cette fonction nous permet d'identifier rapidement les apprenants qui ont besoin d'aide, et d'adapter notre approche pédagogique pour répondre à leurs besoins.

III. ÉVALUATION DE « SCIADO TRACKER »

Afin d'évaluer notre module de suivi, nous avons réalisé une première expérimentation au sein de notre entreprise (Sciado Partenaires) avec six formateurs internes. Ces formateurs sont habitués à assurer des formations à distance à l'aide de nos simulateurs. Ils ont testé le module pour donner leur avis sur son utilité. Pour recueillir leurs commentaires, nous avons utilisé le questionnaire d'évaluation d'utilisabilité (SUS).

Les retours des formateurs ont été globalement positifs, avec un niveau élevé de satisfaction et un taux de 77 %. Les formateurs ont trouvé que ce mo-

dule leur permet de visualiser en temps réel la progression des apprenants, et de mesurer l'acquisition de leurs compétences. De plus, ce module offre la possibilité d'intervenir de manière efficace pour aider les apprenants en difficulté. En outre, il permet de mieux comprendre et visualiser le rythme d'apprentissage de chaque apprenant, et de situer les difficultés de ce dernier afin de l'aider à progresser.

Finalement, nous avons également reçu certaines propositions pour améliorer ce module. À titre d'exemple, un expert nous a suggéré d'ajouter une représentation graphique, pour illustrer le positionnement de la progression d'un apprenant par rapport à la progression globale de la classe des apprenants. En termes d'ergonomie, les formateurs ont aussi souligné que certaines améliorations peuvent être apportées pour rendre ce module plus facile à utiliser.

CONCLUSION

Les simulateurs offrent un environnement très prometteur pour l'apprentissage. Ils présentent de nombreux avantages, notamment la possibilité de simuler des situations complexes et d'acquérir de l'expérience en pratique. Cependant, pour une utilisation optimale, il est nécessaire de travailler sur la personnalisation de l'apprentissage, en fonction des besoins individuels des apprenants. Au sein de notre entreprise, nous proposons des simulateurs pédagogiques qui permettent la montée en compétences de l'apprenant dans le domaine de la gestion d'entreprise et du management. Avec l'intégration du module de suivi « Sciado Tracker », nous avons renforcé le processus de montée en compétences, donnant ainsi la possibilité à chaque apprenant de progresser selon son rythme.

Notre expérimentation a montré que le module de suivi peut être très utile pour évaluer la progression des apprenants, et améliorer leur expérience d'apprentissage. À l'avenir, nous comptons continuer à améliorer ce module en ajoutant des outils de prédiction de comportement de l'apprenant. Les objectifs sont d'optimiser le processus d'apprentissage, et de l'intégrer à d'autres simulateurs, pour mesurer son impact sur la progression des apprenants dans différents domaines. Nous sommes convaincus que la combinaison de simulateurs pédagogiques et de personnalisation de l'apprentissage peut offrir un potentiel considérable pour les apprenants, les aidant à atteindre leurs objectifs de formation plus rapidement et avec plus de succès.

Bibliographie

S. L. Brunton, J. L. Proctor, J. N. Kutz (2016). « Discovering governing equations from data by sparse identification of nonlinear dynamical systems ». *Proceedings of the national academy of sciences, 113(15). 3932-3937.*

A.Gruler, J. D. Armas Adrián, A. A. Juan, D. Goldsman (2019). « Modelling human network behaviour using simulation and optimization tools: the need for hybridization ». *SORT: statistics and operations research transactions,* 43(2). 0193-222.

D. Martin, B. McEvoy (2003). « Business simulations: a balanced approach to tourism education ». *International Journal of Contemporary Hospitality Management,* 15(6). 336-339.

Lara Kathleen Smetana and Randy L. Bell (2012). « Computer simulations to support science instruction and learning: A critical review of the literature ». *International Journal of Science Education.* 34(9): 1337-1370.

J. Smed (2002). *Production Planning in Printed Circuit Board Assembly.* PhD thesis, Turku, Finland:University of Turku.1-71.

P. Spagnoletti, E. D'Atri, A. D'Atri (2013). « Managing Decision-Making in Supply Chain and Value Networks : The Beer Game Evolution ». *International Journal of Electronic Commerce Studies.* 4(1), 63-78.

A. D'Atri, P. Spagnoletti, A. Banzato, C. Bonelli (2009). *Supply Chain and Virtual Enterprises: the Beer Game evolution.*

U. Thorsteinsson (1995). *The LOGI-Game-a Dynamic Modular Logistic Game.* Simulation games and learning in production management, 175-184.

A. Siddiqui, M. Khan, S. Akhtar (2008). « Supply chain simulator: A scenario-based educational tool to enhance student learning ». *Computers & Education,* 51(1), 252-261.

E. W. Chan, C. E. Fan, M. W. Lewis, K. King, P. Dreyer, C. Nelson (2009). *The RSS-POD Supply Chain Management Game.* Rand Health.

A. Abela (2009). Chart suggestions-a thought starter. Extreme Presentation.

L. O. Wilson (2016). *Anderson and Krathwohl–Bloom's taxonomy revised.* Understanding the New Version of Bloom's Taxonomy.

J. Brooke (1986). *System usability scale (SUS): a quick-and-dirty method of system evaluation user information. Reading, UK: Digital equipment co ltd,* 43, 1-7.

ISBN 978-2-9591124-1-6

© Robert Delaunay – *Rythme 2* (gouache sur papier) – 1938 — CCØ Paris Musées / Musée d'Art Moderne de la Ville de Paris