Ecole doctorale : ED 476

Neurosciences et Cognition

Institut de psychologie

Laboratoire d'Etude des Mécanismes Cognitifs

Discipline: Psychologie cognitive

Etude de l'impact de contenus pédagogiques multimédia interactifs sur l'attention et l'engagement des apprenants

Minne Youri

Réalisée sous la direction du Pr. George Andrew MICHAEL et de la Pr. Élise LAVOUÉ

Table des matières

Table des matières 2
Table des figures 6
Remerciements 7
Introduction générale 8
Chapitre 1 : Cadre conceptuel de l'apprentissage multimédia 12
A. Théorie cognitive de l'apprentissage multimédia 13
1. La théorie du double codage 13
2. La théorie de la charge cognitive 14
3. Processus cognitifs 15
4. Principes 17
5. Limites de cette théorie 18
B. Approches complémentaires 19
1. Théorie cognitive et affective de l'apprentissage multimédia 19
2. Le design émotionnel 20
3. Le rôle de l'anthropomorphisme 21
4. Théorie de l'agentivité sociale 23
5. Agents pédagogiques 24
C. Engagement 25
1. Cadre théorique 25
2. Motivation, engagement et performance 28
D. L'attention dans l'Apprentissage Multimédia 31
1. Inhibition 32
1.1. Cadre théorique 32
1.2. Hétérogénéité de l'inhibition 33
2. Gestion des Ressources Attentionnelles 34
2.1. Cadre théorique 34
2.2. Hétérogénéité de la gestion des ressources 35
2.3. Lien entre ressources attentionnelles et inhibition 35
3. Déploiement des ressources attentionnelles 37
4. L'attention dans l'apprentissage multimédia 40

E. Points d’ombre et objectifs de recherche 41
Chapitre 2. Modèle Intégratif de l’engagement et de l’attention 44
1. Introduction 46
2. Multimedia learning theories 47
2.1. Cognitive theory of multimedia learning 47
2.2. Cognitive and affective theory of multimedia learning 50
3. Learners’ engagement 51
3.1. Situational engagement 52
3.2. Student engagement and its components 53
3.3. Engagement in multimedia learning 54
4. Attention in multimedia learning 56
4.1. Deployment of attentional resources 56
4.2. Salience and multimedia learning 57
4.3. Does attention reflect engagement? 59
5. An integrative model of engagement and attention in multimedia learning 61
5.1. Antecedents of learners’ situational engagement 61
5.1.1. Individual characteristics 62
5.1.2. Task characteristics and multimedia support 63
5.1.3. Environment 64
5.2. Link between individual characteristics and engagement 64
5.3. Engagement and deployment of attentional resources 65
5.4. Outcomes 67
6. Conclusion 68
7. References 69
Transition et objectif de recherche 79
Chapitre 3 : Impact de l'Anthropomorphisme dans l’Apprentissage Multimédia 81
Abstract 82
1. Methodology 85
2. Statistical Analyses 89
3. Results 89
4. Discussion 95
5. References 100
Transition 102

Chapitre 4 : Le rôle des capacités attentionnelles sur l'impact des agents pédagogiques et l'engagement 103
Abstract 105
1. Introduction 106
2. Methodology 109
3. Results 115
4. Discussion 119
5. References 125
Analyses complémentaires 131
Transition 134
Chapitre 5 : Le rôle de l'interactivité dans l'apprentissage multimédia 136
Abstract 138
1. Introduction 139
2. Methodology 146
3. Results 153
4. Discussion 155
5. References 158
Discussion générale 161
A. Synthèse 161
B. Implications théoriques 162
1. L'engagement 162
2. Motivation 168
3. L'attention 170
4. Attention et engagement 171
5. Performances d'apprentissage 173
C. Implications pratiques 176
Conclusion 178
Bibliographie 180

Table des figures

Chapitre 1

Figure 1 : Illustration de la théorie du double codage.. 14
Figure 2 : Illustration des trois processus cognitifs mis en jeu lors d'un apprentissage multimédia.... 17
Figure 3 : Illustration du mécanisme d'activation social tirée de la théorie de l'agentivité sociale (Mayer, 2005).. 24
Figure 4 : Schéma du modèle MAM.. 39

Chapitre 2

Figure 1: Integrative model of engagement and attention in multimedia learning............................. 62

Chapitre 3

Figure 1: Learning materials. Interfaces of the e-learning environment: on the left the main interface to access the modules, on the right a screenshot of the interface to navigate in the video...................... 87
Figure 2: Learning materials. Screenshots of two videos: the neutral version on the left and the rich version on the right 87
Figure 3; Path plot mediation of (1) External Regulation (2) UES and (3) mean d' 96

Chapitre 4

Figure 1: Learning materials. Interfaces of the e-learning environment: on the left the main interface to access the modules, on the right a screenshot of the interface to navigate in the video. 110
Figure 2: Learning materials. Screenshots of the three videos: the absent version on the left, the weak version on the middle and the strong version on the right. 111
Figure 3: Interaction between ID and memorization, and RD and memorization. The higher the index, the less the subject has good associated abilities... 119

Chapitre 5

Figure 1: Theory of social agency from Mayer (2012).. 141
Figure 2: Learning materials. Interfaces of the e-learning environment: on the left the main interface to access the modules, on the right a screenshot of the interface to navigate in the video................... 148
Figure 3: Learning materials. On the left Video of the control APA condition and on the right video of the interactive APA condition... 148

Remerciements

Je tiens à exprimer ma profonde gratitude envers toutes les personnes qui m'ont accompagné et soutenu tout au long de ce travail de thèse.

En premier lieu, je remercie chaleureusement mon directeur de thèse, Georges Michaël, pour sa guidance et ses enseignements précieux. Ses conseils avisés et sa rigueur scientifique ont été essentiels pour la réalisation de ce travail.

Je souhaite également exprimer toute ma reconnaissance à ma co-directrice de thèse, Élise Lavoué. Son soutien constant et son encouragement tout au long de mon parcours doctoral ont été d'une aide inestimable.

Un grand merci à Clarence Thiery, ma responsable entreprise, à qui je dois énormément. Son soutien et ses conseils ont grandement contribué à la réussite de cette thèse.

Je remercie également mes collègues doctorants, Enzo, Rémi et Anthony, pour leur camaraderie et leur soutien tout au long de ces années. Nos discussions et collaborations ont enrichi mon travail et rendu ce parcours plus agréable.

Je n'oublie pas mes collègues de Sydo, pour leur soutien et leurs encouragements, et plus particulièrement Sylvain Tillon, pour sa précieuse collaboration et son amitié.

Enfin, je tiens à remercier toutes les personnes, proches, amis et famille, qui m'ont soutenu, de près ou de loin, dans cette aventure doctorale.

Introduction générale

Apprendre, acquérir, accumuler et intégrer de nouvelles connaissances est un processus essentiel à l'adaptation humaine. Historiquement, l'apprentissage se produisait de manière naturelle et souvent informelle, à travers l'expérience quotidienne, l'observation, l'imitation et la transmission des savoirs au sein des communautés (Gray, 2016). Ce type d'apprentissage est caractérisé par son aspect direct et pratique, où les connaissances sont acquises de manière fluide au sein du contexte dans lequel elles sont directement applicables. Avec le développement de notre société, la complexification des savoirs et l'émergence d'institutions dédiées à l'éducation, l'apprentissage est devenu plus formel et explicite. Cette formalisation a entraîné la mise en place de structures éducatives où l'enseignement est dispensé selon des programmes prédéfinis. Dans le cadre de l'apprentissage explicite, l'apprenant est souvent perçu comme passif, recevant des connaissances de la part d'un enseignant ou d'une source externe. Cette perception s'ancre dans des modèles pédagogiques où l'enseignant est le détenteur et le transmetteur du savoir, et l'apprenant, le récepteur de ce savoir. Ce modèle est critiqué par des penseurs comme Peter Gray (2016), qui prônent un retour vers des formes d'apprentissage plus naturelles et engageantes. Cependant, comme le soulignent Geary (2008), l'apprentissage de connaissances abstraites, telles que celles enseignées aujourd'hui, ne suit pas toujours un processus naturel, nécessitant des approches pédagogiques spécifiques pour surmonter ces défis. C'est dans ce contexte de réflexion sur l'efficacité et l'évolution des méthodes d'enseignement que la digitalisation est apparue comme pouvant supporter de nouvelles formes d'apprentissage.

Dans le contexte actuel marqué par une numérisation croissante des secteurs économiques, allant du secteur bancaire à l'administration publique, le domaine de l'éducation connaît également une transformation significative. Cette transformation est particulièrement évidente en Europe, où l'on observe une consolidation de politiques éducatives favorisant l'utilisation et le développement de l'éducation numérique. En France, l'impact de la pandémie de COVID-19 a accéléré cette transition numérique, modifiant le paysage de l'apprentissage et soulignant le rôle croissant de la filière EdTech, dont le chiffre d'affaires était estimé à 1,3 milliard d'euros en 2021. Cette orientation politique a conduit à une augmentation de la diversité des supports pédagogiques disponibles et a encouragé les enseignants à adopter des cadres de compétences axés sur l'exploitation de ces ressources numériques pour innover et améliorer les pratiques d'enseignement et de formation.

En conséquence, on observe une intégration croissante de matériel numérique dans les ressources éducatives, telles que les manuels scolaires. À titre d'exemple, le nombre de tableaux numériques interactifs est passé de 2 pour 1000 élèves dans les écoles élémentaires en 2009 à 17 pour 1000 élèves en 2019. Dans les collèges, il est passé de 3 pour 1000 élèves à 17,7 pour 1000 élèves au cours de la même période. Dans le même temps, on observe une forte augmentation du pourcentage d'établissements et d'écoles dont le projet fait référence aux technologies de l'information et de la communication (DEPP, 2021).

La numérisation de l'éducation présente plusieurs avantages, notamment une meilleure accessibilité, nécessitant uniquement l'usage d'un ordinateur portable, ainsi qu'une certaine flexibilité permettant aux apprenants de progresser à leur propre rythme et de reconsulter le contenu pédagogique en fonction de leurs besoins individuels (Zimmerman, 2012, Lee & Choi, 2011). De plus, le support numérique offre la possibilité d'adapter l'apprentissage aux besoins de l'apprenants, par exemple en leur proposant une gamification adaptée (Lavoué, 2021).

Toutefois, cette transition numérique s'accompagne également de défis, notamment la spécificité de l'apprentissage multimédia asynchrone. Ce type d'apprentissage comprend des modules d'enseignement en ligne que l'apprenant peut suivre en différé. En d'autres termes, les étudiants ont à leur disposition différentes ressources pédagogiques (fichiers audio, vidéos, QCM, contenus interactifs…) qu'ils peuvent consulter librement. L'apprenant et le formateur, ou l'étudiant et le professeur, ne sont donc pas impliqués au même moment dans le processus pédagogique. La connexion n'étant pas simultanée, les échanges se font nécessairement par mail ou, par exemple, sur un forum de discussion dédié. L'apprentissage se fait donc en autonomie, ce qui représente un changement de paradigme par rapport aux environnements éducatifs traditionnels, principalement parce qu'il retire de l'équation la présence physique d'un enseignant.

Pourtant, dans un cadre traditionnel, les enseignants jouent un rôle crucial non seulement dans la transmission de connaissances mais aussi dans le suivi et la régulation de l'attention des apprenants. Ils sont capables de détecter les signes de décrochage attentionnel chez les élèves et d'intervenir en conséquence, que ce soit par la modification de leur stratégie pédagogique, l'engagement direct avec l'élève, ou l'introduction d'activités pour raviver l'intérêt de la classe (Michael G.A. & Minne Y, 2024). Or, ces défis se trouvent exacerbés dans l’apprentissage multimédia asynchrone du fait de la présence de distractions potentielles, telles que les notifications constantes des réseaux sociaux, des e-mails etc.

La capacité à maintenir l'attention et à se concentrer sur des tâches d'apprentissage spécifiques sur de longues périodes est une compétence cruciale pour réussir. La concentration soutenue sur une tâche unique, sans se laisser distraire par des éléments externes, est essentielle pour l'assimilation profonde des connaissances. Les interruptions peuvent avoir un impact significatif sur la performance dans des tâches nécessitant une attention soutenue (Brazzolotto & Michael, 2019). La complexité et le timing des interruptions influencent leur effet perturbateur (Brazzolotto & Michael, 2019, 2020). De plus, les interruptions perçues comme déplaisantes semblent avoir un impact plus important sur la performance (Brazzolotto & Michael, 2020). Même de brèves interruptions peuvent réduire significativement la qualité de la performance sur des tâches nécessitant une attention soutenue (Foroughi et al., 2015). Les environnements d'apprentissage multimédia présentent des défis supplémentaires pour le maintien de l'attention. Ces environnements sont souvent cognitivement plus chargés en raison de la présence de nombreuses informations simultanées, telles que du texte, des images, des vidéos et des hyperliens (Mayer, 2014). Cette richesse d'informations peut surcharger la mémoire de travail de l'apprenant et rendre plus difficile la concentration sur les éléments essentiels (Sweller, 2020). La problématique de la motivation et de l’engagement des apprenants s’ajoute à cette question attentionnelle. Certaines études observent de faibles taux de rétention et d'achèvement (Muljana & Luo, 2019 ; Radovan, 2019 ; Xavier & Meneses, 2020) avec des taux d'abandon plus élevés que ceux de l'apprentissage en face à face (Gaytan, 2015 ; Radovan, 2019). Plusieurs facteurs peuvent contribuer au faible engagement dans l'apprentissage multimédia asynchrone. L'isolement perçu est l'un des principaux défis, car l'absence de contact direct avec les enseignants et les pairs peut diminuer le sentiment d'appartenance à une communauté d'apprentissage. En outre, la surcharge d'informations résultant de la multitude de ressources disponibles peut submerger les apprenants, rendant difficile la concentration et la sélection des informations pertinentes.

Dans l'apprentissage multimédia asynchrone, les problématiques de l'attention et de l'engagement des apprenants se trouvent alors interconnectées (Folk et al. 2009). Un apprenant engagé pourrait ne pas allouer les ressources attentionnelles nécessaires à son apprentissage. De plus, un apprenant avec de faibles capacités attentionnelles pourrait être surchargé d'informations, et de ce fait se décourager et se désengager de la tâche (Sweller, 2020). De fait, il est crucial de développer des contenus, formats et outils pédagogiques adaptés afin de maximiser l'attention et l'engagement des apprenants, tout en évitant leur dispersion. L'objectif principal de cette thèse est de proposer et valider un modèle intégratif unifiant ces différents aspects. Dans le premier chapitre, nous présenterons le cadre théorique dans lequel s'inscrit cette thèse, nous permettant d'appréhender l'apprentissage multimédia, l'attention et l'engagement. Le deuxième chapitre sera consacré à la présentation de notre modèle intégratif de l'engagement et de l'attention dans l'apprentissage multimédia. Ce modèle, élaboré à partir d'une revue de la littérature, vise à clarifier les relations entre ces deux concepts clés et à proposer un cadre explicatif de leur rôle dans l'efficacité de l'apprentissage multimédia. Nous détaillerons les différentes composantes du modèle, leurs interactions et les hypothèses qui en découlent quant aux conditions d'un engagement et d'un déploiement attentionnel optimaux.

Les chapitres suivants seront consacrés à la validation empirique de ce modèle par le biais de trois études expérimentales. Le chapitre 3 se penchera sur l'effet de l'anthropomorphisme sur l'engagement, la motivation et les performances des apprenants. Le chapitre 4 examinera plus en détail le rôle des agents pédagogiques en variant leur présence et leur niveau d'expressivité, tout en analysant leur interaction avec les capacités attentionnelles des apprenants. Enfin, le chapitre 5 abordera l'apport des technologies d'intelligence artificielle dans le développement d'agents pédagogiques interactifs, en évaluant leur influence sur l'engagement et les performances d'apprentissage, ainsi que leur relation avec les capacités attentionnelles.

Dans la discussion générale, nous synthétiserons les principaux résultats de nos études et les analyserons à travers le prisme de notre modèle intégratif et de la littérature existante. Cette synthèse nous permettra de mettre en lumière les déterminants clés de l'engagement et du déploiement de l'attention des apprenants dans le contexte spécifique de l'apprentissage multimédia asynchrone. En nous appuyant sur ces résultats, nous proposerons des implications pratiques pour la conception de dispositifs pédagogiques adaptés, visant à optimiser l'engagement et l'attention des apprenants dans les environnements d'apprentissage multimédia. Nous identifierons également les limites de nos travaux et les perspectives de recherche qu'ils ouvrent, en particulier concernant l'exploitation des avancées récentes en matière d'intelligence artificielle et de personnalisation des environnements d'apprentissage en fonction des caractéristiques individuelles des apprenants.

Chapitre 1 : Cadre conceptuel de l’apprentissage multimédia

Ce chapitre est consacré à la présentation du cadre théorique de l'apprentissage multimédia, qui constitue le contexte général de notre recherche. Nous nous intéresserons particulièrement à deux concepts clés : l'attention et l'engagement, et à leur rôle dans l'efficacité de l'apprentissage multimédia.

Nous commencerons par un examen approfondi de la Théorie Cognitive de l'Apprentissage Multimédia développée par Richard Mayer (2005), qui est prédominante dans ce champ théorique. Nous mettrons en exergue ses principes fondamentaux et ses implications pour la conception de matériel pédagogique multimédia. Nous élargirons ensuite notre perspective en intégrant les aspects affectifs de l'apprentissage, tels que proposés par Moreno et Mayer (2007). Cette extension est cruciale pour comprendre comment les états émotionnels et motivationnels des apprenants influencent leur engagement et leur traitement cognitif dans un environnement multimédia. Nous explorerons également la théorie de l'agentivité sociale de Mayer (2005), qui examine comment la dimension sociale de l'apprentissage multimédia peut augmenter le déploiement des ressources cognitives.

Dans ce contexte, nous aborderons deux éléments spécifiques de l'apprentissage multimédia : l'anthropomorphisme et l'agent pédagogique. Ces éléments ont été étudiés pour leur impact potentiel sur l'efficacité de l'apprentissage multimédia, notamment à travers leur influence sur l'engagement et l'attention des apprenants (Schroeder et al., 2013). Nous nous pencherons sur le concept d'engagement, qui est un élément central pour comprendre l'impact des principes de design multimédia sur le déploiement des ressources cognitives (Kirschner et al. 2011). L'engagement sera considéré non pas comme un simple intérêt passager, mais comme une immersion active dans le processus d'apprentissage. Nous soulignerons le rôle de l'engagement dans l'amélioration des performances d'apprentissage et son lien avec la motivation (Astleitner et al. 2004).

Enfin, nous traiterons de l'attention, à la fois en termes de capacités et de déploiement, ainsi que de sa spécificité dans l'apprentissage multimédia. Nous examinerons comment les capacités attentionnelles limitées des apprenants influencent leur traitement des informations multimédias, et comment le déploiement efficace de l'attention peut être soutenu par des principes de design pédagogique fondés sur des preuves (Mayer, 2014). Nous explorerons les défis spécifiques de la gestion de l'attention dans des environnements d'apprentissage multimédia, tels que la nécessité de guider l'attention vers les informations pertinentes tout en évitant la surcharge cognitive.

Pour conclure, nous introduirons notre question de recherche sur le lien entre attention et engagement dans l'apprentissage multimédia. Nous proposerons un modèle intégratif qui vise à comprendre comment ces deux dimensions interagissent pour favoriser un apprentissage multimédia profond et durable. Ce modèle s'appuiera sur l'ensemble des cadres théoriques discutés dans le chapitre, offrant ainsi une vision holistique qui capture la complexité de l'apprentissage multimédia dans les environnements éducatifs contemporains.

A. Théorie cognitive de l’apprentissage multimédia

La Théorie Cognitive de l'Apprentissage Multimédia s'articule autour de modèles et de théories préexistants en psychologie cognitive. Trois théories sous-tendent ce modèle, la théorie du double codage de Paivio (1986), la théorie de la charge cognitive de Sweller (1998) et l’approche de Baddeley (1986) de la mémoire de travail.

1. La théorie du double codage

Développée par Allan Paivio (1971), la théorie du double codage propose que l'information soit traitée à travers des canaux verbaux et visuels distincts. Cette dualité de codage permet à une information d'être représentée sous deux formes différentes, améliorant ainsi la probabilité de son rappel. Par exemple, le concept de "chat" peut être encodé à la fois sous forme verbale ("chat") et sous forme visuelle (image d'un chat). Des travaux de neuropsychologie, notamment sur des patients "split-brain"[1] menées dans les années 50-60 (Gazzaniga, 2005), ont montré que ce double codage est un processus cognitif automatique. Ce processus d’eencodage d’un stimulus de deux manières différentes augmente la probabilité de

[1] Syndrome résultant de la déconnexion des deux hémisphères (Gazzaniga, 2005)

se souvenir de ce stimulus, comparativement à un stimulus encodé de façon unimodale. De plus, une image accompagnée d'une information verbale pertinente est mieux mémorisée (Anderson & Bower, 1973). Dans le contexte de l'Apprentissage Multimédia, la théorie du double codage de Paivio offre donc un cadre pour comprendre comment les informations multimédias sont traitées efficacement. Les apprenants traitent les informations visuelles et verbales dans deux canaux séparés, ce qui optimise le traitement cognitif et la rétention de l'information. L'intérêt du multimédia réside donc en partie sur sa capacité à proposer une information de manière multimodale via des processus cognitifs distincts. Cependant, ces processus peuvent être surchargés cognitivement, on parle alors de charge cognitive.

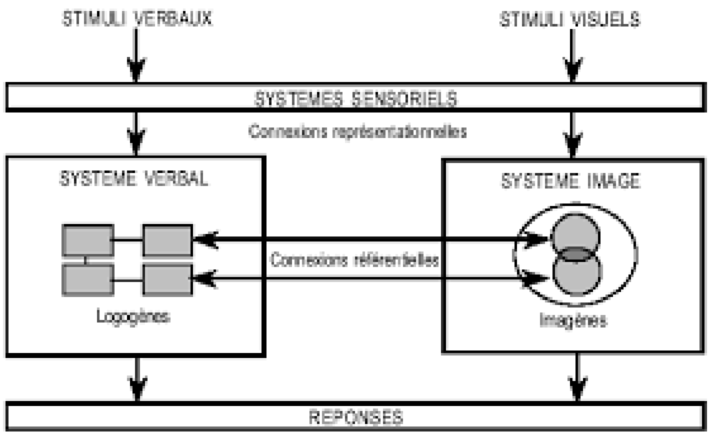

Figure 1 : Illustration de la théorie du double codage

2. La théorie de la charge cognitive

Cette théorie repose sur l'hypothèse que la capacité de la mémoire de travail est limitée, et que cette limitation joue un rôle crucial dans le processus d'apprentissage. Sweller et ses collègues (1994) ont mis en avant que pour optimiser l'apprentissage, il est essentiel de considérer la quantité d'information et les processus cognitifs sollicités simultanément, afin de ne pas surcharger la mémoire de travail. Le processus de traitement de l'information consiste en la création d'un schéma mental en mémoire de travail. Ce schéma mental est une représentation cognitive simplifiée et organisée des informations pertinentes pour une tâche donnée (Chi, Glaser, & Rees, 1982). La construction de ces schémas mentaux permet de réduire la charge cognitive en mémoire de travail en regroupant les informations en unités significatives et en éliminant les détails non pertinents. Une fois construit, le schéma mental peut être automatisé et stocké en mémoire à long terme, libérant ainsi des ressources en mémoire de travail pour le traitement de nouvelles informations. Cependant, si la quantité d'informations à traiter dépasse la capacité de la mémoire de travail, la construction du schéma mental sera entravée, ce qui nuira à l'apprentissage (Sweller et al., 1998). On parle alors de charge cognitive comme étant le coût que représente la tâche. La théorie distingue trois types principaux de charge cognitive : intrinsèque, extrinsèque et utile.

La charge cognitive intrinsèque est déterminée par le nombre de concepts à manipuler simultanément, leurs interactions et leurs complexités. Le niveau d'expertise de l'apprenant va alléger la charge intrinsèque d'un dispositif. Il dépend (1) du nombre d'éléments à intégrer (2) de leurs interactions et (3) de leur complexité.

Le deuxième type de charge dans une situation d'apprentissage serait la charge cognitive extrinsèque. La charge extrinsèque désigne les ressources cognitives allouées à la prise en compte de tous les éléments qui ne sont pas directement nécessaires vis-à-vis de l'objectif pédagogique. Elle comprend tous les éléments venant s'ajouter à la tâche comme les distractions, des explications trop complexes ou un niveau de détails trop élevé.

Le reste des ressources cognitives sera attribué à la charge utile. Cette charge permet d'ancrer des schémas mentaux de mémoire de travail en les confrontant avec les connaissances antérieures stockées en mémoire à long terme. Un apprentissage significatif se produit lorsque des instructions appropriées sont appliquées pour diminuer la charge cognitive extrinsèque et pour libérer des ressources cognitives dédiées au traitement intrinsèque et pertinent, contribuant ainsi à un apprentissage amélioré. Un apprentissage n'est effectif que si l'apprenant a suffisamment de ressources pour construire un schéma. Or, si la tâche devient particulièrement « coûteuse » pour la mémoire de travail en raison d'une charge cognitive intrinsèque ou d'une charge cognitive extrinsèque trop importante, l'apprenant peut manquer de ressources cognitives nécessaires pour la construction d'un schéma mental.

3. Processus cognitifs

La Théorie Cognitive de l'Apprentissage Multimédia identifie trois processus cognitifs principaux qui interviennent dans l'apprentissage multimédia : la sélection, l'organisation et l'intégration (Mayer, 2005). La sélection des modalités fait référence à la première étape de l'apprentissage, où l'apprenant se concentre sur la perception des éléments pertinents du matériel présenté, que ce soit dans le canal auditif ou visuel (Mayer, 2005). Cette étape de sélection permet à l'apprenant de filtrer les informations non essentielles, mobilisant ainsi la mémoire de travail visuospatiale et verbale de manière efficace (Baddeley, 2003). Ensuite, l'organisation se produit lorsque l'apprenant organise mentalement l'information sélectionnée en modèles logiques et en représentations visuelles, un processus qui repose largement sur le composant de la mémoire de travail dédié à la manipulation et à l'organisation des informations (Baddeley, 2003). Ce processus permet de transformer des informations éparses en un ensemble cohérent, facilitant la compréhension et la mémorisation. Le dernier processus, l'intégration, se produit lorsque l'apprenant relie ces informations organisées à ses connaissances antérieures, permettant la construction d'un schéma cognitif enrichi (Mayer, 2009). Cette étape est essentielle pour une compréhension approfondie et durable, mobilisant à la fois la mémoire de travail et la mémoire à long terme pour l'intégration des nouvelles informations (Baddeley, 2003).

Le modèle de Construction-Intégration de Kintsch (1988) offre un cadre complémentaire pour comprendre ce processus. Selon ce modèle, la compréhension se déroule en deux phases distinctes mais interdépendantes. La première phase, la construction, implique la formation d'une représentation mentale du contenu basée sur les informations explicites présentes dans le matériel. Durant cette phase, le lecteur ou l'apprenant établit une "base de texte" en extrayant les propositions et en les reliant entre elles (Kintsch, 1998). La seconde phase, l'intégration, consiste à intégrer cette représentation aux connaissances antérieures de l'individu pour former une compréhension globale cohérente, appelée "modèle de situation" (Kintsch, 1998). Ce processus d'intégration permet non seulement de donner du sens aux informations nouvellement acquises, mais aussi d'enrichir et de modifier les connaissances préexistantes.

Le modèle de Kintsch établit également une distinction importante entre la mémorisation et la compréhension. La mémorisation correspond à la capacité de retenir et de restituer les informations explicites présentes dans le matériel, c'est-à-dire la "base de texte". En revanche, la compréhension implique la construction d'un "modèle de situation" cohérent, qui intègre les informations nouvelles aux connaissances antérieures (Kintsch, 1994). Ainsi, il est possible de mémoriser un texte sans nécessairement le comprendre en profondeur, tandis qu'une compréhension véritable nécessite l'intégration des informations dans un réseau de connaissances préexistant.

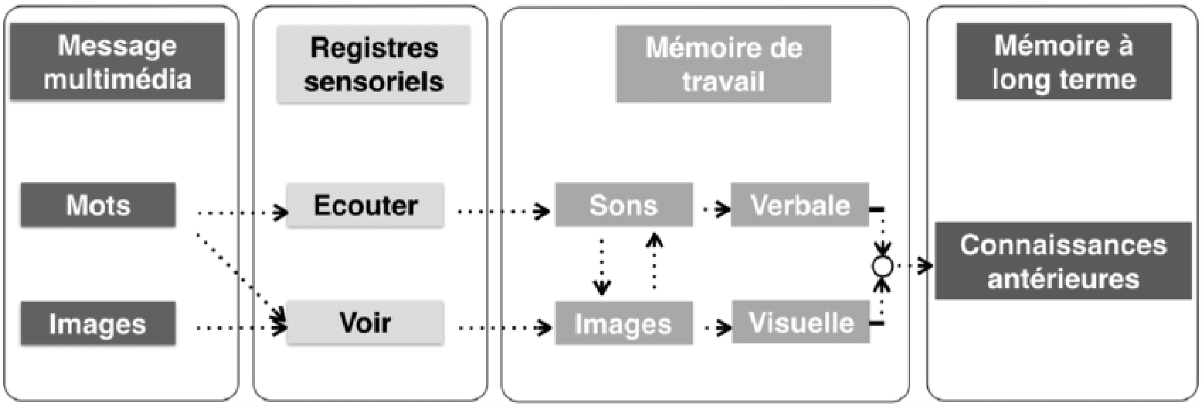

Figure 2 : Illustration des trois processus cognitifs mis en jeu lors d'un apprentissage multimédia

4. Principes

A partir de ce cadre théorique, de nombreuses recherches empiriques ont étudié des principes de conception pédagogique ayant pour objectif de favoriser la rétention et la compréhension du dispositif (pour une revue, voir Mayer, 2021, 3ème édition). Un ensemble de 14 principes de design sont identifiés, classés en trois catégories en fonction de leur influence sur les processus cognitifs. Ces principes peuvent être regroupés selon le type de charge cognitive qu'ils visent à influencer : réduire la charge extrinsèque, atténuer la charge intrinsèque, ou augmenter la charge utile.

Pour atténuer la charge cognitive extrinsèque, cinq principes sont mis en avant : Cohérence, Signalisation, Redondance, Contiguïté Spatiale, et Contiguïté Temporelle. Le principe de cohérence préconise l'élimination de tout contenu distrayant ou non essentiel par rapport aux objectifs d'apprentissage (Mayer & Jackson, 2005). La signalisation recommande d'intégrer des marques visuelles pour guider l'attention des apprenants vers les informations cruciales (Van Gog, 2014), tandis que le principe de redondance suggère d'éviter la surcharge cognitive en limitant la présentation simultanée de textes et de narrations redondants (Sweller, 2005). La contiguïté spatiale et temporelle, respectivement, soulignent l'importance de la proximité physique entre les mots et les images, et de la synchronisation entre la narration et les animations pour faciliter l'apprentissage (Mayer & Fiorella, 2014 ; Mayer, 2009).

Trois principes ont été proposés afin de gérer la charge cognitive intrinsèque : Segmentation, Pré-entraînement et Modalité. Le principe de segmentation stipule que les étudiants peuvent mieux apprendre lorsque le matériel d'apprentissage est segmenté en plus petits modules (Mayer & Pilegard, 2014). Selon le principe de pré-entrainement, l'apprentissage peut être amélioré si les concepts clés et les principales caractéristiques sont fournis avant l'apprentissage (Mayer, 2009). Les apprenants peuvent avoir besoin de temps pour construire mentalement un modèle causal dans le contenu multimédia, surtout lorsqu'il est complexe. Le principe de modalité indique qu'il est préférable de fournir l'information sous forme de narration plutôt que d'un texte à l'écran car deux canaux (visuel et auditif) sont utilisés lorsque les mots sont servis sous forme de narration (Moreno & Mayer, 1999).

Trois autres principes aident les apprenants à maximiser la charge cognitive utile, à savoir les principes de Personnalisation, de Voix et d'Image. Le principe de personnalisation indique qu'avoir un style plus conversationnel améliore l'apprentissage par rapport à un style formel (Mayer et al., 2004). Ainsi, les concepteurs pédagogiques devraient éviter d'utiliser autant que possible le langage académique. Il est affirmé dans le principe de voix que « les gens apprennent mieux lorsque la narration est prononcée avec une voix humaine plutôt qu'avec une voix de machine » (Mayer, 2009, p. 242). Enfin, le principe d'image stipule que l'ajout de photos des intervenants lors de la présentation du matériel d'apprentissage ne signifie pas que les résultats d'apprentissage sont améliorés. Il est préférable d'utiliser des animations et visuels pertinents au lieu d'afficher une tête parlante de l'instructeur.

5. Limites de cette théorie

La théorie cognitive de l'apprentissage multimédia de Mayer présente des limitations significatives qui méritent d'être examinées. La première concerne sa capacité de généralisation, notamment sur ses principes qui ne sont pas effectifs dans tous les cas. Par exemple, dans le cas de l'apprentissage d'une langue, Munassar et ses collègues (2010) ont mis en évidence que les principes de modalité et de redondance de Mayer n'étaient pas nécessairement applicables, car la redondance d'une même information via le canal visuel et auditif favorise l'apprentissage d'un mot dans son orthographe et sa prononciation. Une deuxième limite concerne l'aspect très cognitif de cette théorie, ne prenant pas en compte les aspects affectifs et motivationnels de l'apprentissage (Brom, 2018).

L'extension de cette approche par la Théorie Cognitivo-Affective de l'Apprentissage avec les Médias de Moreno (2006) permet donc d'introduire une dimension émotionnelle et motivationnelle dans la conception de l'apprentissage multimédia. Cette théorie avance que les facteurs émotionnels jouent un rôle médiateur dans le processus d'apprentissage multimédia, influençant la disposition de l'apprenant à déployer un effort mental nécessaire pour apprendre (Park et al., 2015). Elle prend également en considération la manière dont la métacognition de l'apprenant, incluant ses styles et capacités cognitifs, médie le processus d'apprentissage multimédia (Moreno & Durán, 2004 ; Park et al., 2015), suggérant que la mémoire de travail de l'apprenant, sa motivation (i.e., la charge pertinente) et son style cognitif sont influencés par son état émotionnel.

B. Approches complémentaires

1. Théorie cognitive et affective de l'apprentissage multimédia

L'attention, les fonctions mnésiques et le contrôle exécutif sont des fonctions cognitives nécessaires pour un apprentissage efficace. Or, ces processus sont étroitement liés aux émotions. En effet, les émotions influencent particulièrement l'attention, notamment en modulant la sélectivité de l'attention ainsi qu'en motivant l'action et le comportement lors d'un apprentissage (Tyng et al., 2017). La théorie cognitive et affective de l'apprentissage multimédia (CATLM) développée par Moreno (2005), constitue une extension de la Théorie Cognitive de l'Apprentissage Multimédia de Mayer. L'innovation principale de la CATLM réside dans son incorporation d'éléments affectifs et émotionnels. Dans le cadre des théories de l'apprentissage, l'affect fait référence aux états émotionnels et aux dispositions des apprenants, qui peuvent influencer leur engagement, leur motivation et leur traitement cognitif de l'information (Park, Plass & Brünken, 2014).

La CATLM met en lumière la façon dont les états affectifs des apprenants peuvent influencer leur engagement et leur traitement cognitif de l'information. Selon cette théorie, la motivation est considérée comme un facteur affectif qui influence l'engagement cognitif, c'est-à-dire l'investissement et l'effort mental fourni par l'apprenant lors du processus d'apprentissage (Moreno, 2006). Des recherches ont montré que l'engagement émotionnel des apprenants peut avoir un impact significatif sur leur capacité à traiter et à retenir les informations (Um, Plass, Hayward & Homer, 2012 ; Plass, Heidig, Hayward, Homer & Um, 2014).

De plus, la CATLM souligne l'importance de la médiation métacognitive dans l'apprentissage. Les processus métacognitifs, tels que l'auto-régulation, impliquent une interaction entre des dimensions cognitives et émotionnelles/affectives (Efklides, 2011 ; Zimmerman & Schunk, 2011). Les apprenants qui sont conscients de leurs propres processus de pensée et qui régulent leurs émotions sont plus à même de s'engager efficacement dans l'apprentissage (Pintrich, 2000 ; Ben-Eliyahu & Linnenbrink-Garcia, 2013).

Enfin, la CATLM prend en compte les différences individuelles des apprenants, telles que leurs connaissances antérieures et leurs caractéristiques cognitives, qui peuvent influencer l'efficacité de l'apprentissage multimédia (Moreno, 2006). Des études ont montré que ces variations individuelles interagissent avec la conception des supports multimédias pour impacter les résultats d'apprentissage (Lusk & Atkinson, 2007 ; Höffler, 2010).

En intégrant des aspects affectifs, motivationnels et métacognitifs, la théorie de Moreno offre une compréhension plus complète de l'apprentissage multimédia. Elle souligne l'interaction complexe entre le traitement de l'information, régulé par des processus cognitifs et métacognitifs, et les états émotionnels et motivationnels des apprenants. De cette théorie émerge la question de la conception de supports multimédias émotionnellement stimulants, tout en considérant les limites des capacités cognitives des apprenants.

2. Le design émotionnel

Dans la conception d'environnements d'apprentissage multimédia, la manière dont l'information est représentée visuellement dans les supports d'apprentissage joue un rôle crucial sur les émotions des apprenants, lesquelles influencent à leur tour le processus d'apprentissage (Um et al.2012). La notion que les stimuli esthétiquement plaisants peuvent positivement impacter l'apprentissage est généralement résumée par « ce qui est beau est bon » (Dion, Berscheid et Walster, 1972). La vision est, à bien des égards, l'outil d'apprentissage le plus puissant dont disposent les humains. La recherche a montré que diverses conceptions esthétiques peuvent évoquer des émotions spécifiques, impactant ainsi les performances et les processus cognitifs des apprenants (Harp & Mayer, 1997 ; Mayer & Moreno, 1998 ; Szabo & Kanuka, 1998 ; Wolfson et Case, 2000). Il a également été découvert que la conception d'éléments multimédias, tels que le design visuel, la mise en page, la couleur, et le son, génère des perceptions positives de l'apprentissage chez les utilisateurs (Tractinsky et al., 2000). Un exemple particulièrement illustratif du design émotionnel est le principe de l'anthropomorphisme, où des éléments du contenu multimédia prennent des traits ou des caractéristiques humaines.

3. Le rôle de l'anthropomorphisme

Les recherches dans le domaine du design émotionnel ont démontré que l'utilisation d'éléments anthropomorphiques et de couleurs agréables dans les supports multimédias d'apprentissage peut influencer positivement les états affectifs et motivationnels des apprenants, ainsi que leurs performances d'apprentissage (Mayer & Estrella, 2014 ; Plass et al., 2014 ; Um et al., 2012). Les éléments anthropomorphiques sont des représentations visuelles qui imitent des caractéristiques humaines, comme des visages souriants ou des personnages animés (Brom, Stárková & D'Mello, 2018). Les couleurs agréables, quant à elles, sont généralement des couleurs vives et chaleureuses (Um et al., 2012).

Une méta-analyse conduite par Brom et al. (2018) sur un ensemble de 33 études expérimentales a montré que l'utilisation d'éléments anthropomorphiques, souvent combinée à des couleurs agréables, avait un impact positif sur les performances d'apprentissage. Les résultats ont révélé des effets positifs et significatifs sur la rétention des informations (Mayer & Estrella, 2014), la compréhension (Plass et al., 2014) et le transfert des connaissances dans d'autres situations (Um et al., 2012).

Au-delà des performances, l'anthropomorphisme influencerait positivement les états motivationnels des apprenants. La méta-analyse de Brom et al. (2018) a mis en évidence une augmentation de la motivation intrinsèque (Um et al., 2012), ainsi que des effets plus modérés sur l'appréciation et le plaisir d'apprendre (Plass et al., 2014). Les éléments anthropomorphiques permettraient également de réduire la perception de difficulté de la tâche (Plass et al., 2014).

Plusieurs hypothèses ont été avancées pour expliquer ces effets bénéfiques. L'une d'elles est que les visages et les personnages souriants génèrent des émotions positives qui, selon la théorie cognitive-affective de l'apprentissage multimédia (CATLM ; Moreno, 2005), favorisent l'engagement cognitif des apprenants. Une autre explication est que les éléments anthropomorphiques agissent comme des signaux visuels qui guident l'attention des apprenants vers les informations pertinentes (Schneider et al., 2018). Enfin, en rendant le matériel plus attractif et mémorable, l'anthropomorphisme pourrait simplement augmenter l'intérêt et la motivation à apprendre (Um et al., 2012).

Cependant, la méta-analyse de Brom et al. (2018) n'a pas mis en évidence d'effet significatif de l'anthropomorphisme sur les émotions positives ressenties par les apprenants. Les auteurs ont observé un effet faible et marginalement significatif sur l'affect positif ($d+ = 0,113$).

Ces résultats suggèrent que l'impact de l'anthropomorphisme sur les états émotionnels des apprenants pourrait être moins robuste que son influence sur les performances d'apprentissage et la motivation intrinsèque. Plusieurs interprétations sont envisageables pour expliquer cette absence d'effet significatif sur les émotions positives. Il est possible que l'anthropomorphisme n'influence pas réellement les émotions des apprenants, mais agisse plutôt sur leur motivation intrinsèque et leur engagement cognitif. Une autre explication serait que les mesures auto-rapportées utilisées dans les études pour évaluer les émotions ne soient pas suffisamment sensibles pour détecter des changements subtils dans les états affectifs. Enfin, on ne peut exclure que l'anthropomorphisme génère bien des émotions positives, mais que celles-ci soient trop fugaces pour être captées par les mesures post-apprentissage. Quoi qu'il en soit, ces résultats semblent indiquer que l'impact bénéfique de l'anthropomorphisme sur l'apprentissage serait davantage lié à ses effets sur la motivation intrinsèque et le guidage attentionnel des apprenants, plutôt qu'à une influence directe sur leurs émotions positives. Des recherches supplémentaires intégrant des mesures plus fines des processus émotionnels et motivationnels en cours d'apprentissage permettraient de mieux comprendre les mécanismes par lesquels l'anthropomorphisme favorise les apprentissages.

Une interprétation alternative serait que l'aspect "humain" des éléments anthropomorphiques déclencherait une forme d'agentivité sociale chez les apprenants. Selon la théorie de l'agentivité sociale (Mayer, 2014), la présence d'un agent social dans un environnement d'apprentissage, même sous une forme schématique ou symbolique, activerait chez l'apprenant un sentiment de présence sociale et un mode de traitement social de l'information. Cette activation des schémas sociaux pourrait alors favoriser l'engagement, la motivation et, in fine, les performances d'apprentissage, sans nécessairement passer par une augmentation des émotions positives.

4. Théorie de l’agentivité sociale

L'apprentissage multimédia asynchrone est profondément solitaire, or, il a été montré que les interactions entre apprenants (discussions, collaborations, entraide...) et avec l'enseignant (feedback, soutien...), ainsi que le sentiment de connexion interpersonnelle et d'appartenance à un groupe ayant des objectifs communs sont des facteurs clés d'engagement et de rétention des apprenants (Rovai, 2002). Bien que les utilisateurs puissent interagir avec le contenu à leur propre rythme, l’absence de ces éléments peut souvent mener à un sentiment d'isolement qui affecte la motivation et le processus d'apprentissage.

La théorie de l'agentivité sociale dans les environnements d'apprentissage multimédia, développée par Mayer et al. (2003) et Moreno et al. (2001), propose que ces environnements puissent être conçus pour encourager les apprenants à opérer sous l'hypothèse que leur relation avec l'ordinateur est sociale, suivant les conventions de communication humaine décrites par Reeves et Nass (1996). Cette théorie suggère que l'utilisation de signaux sociaux verbaux et visuels dans les environnements informatiques peut favoriser le développement d'un partenariat, en amenant les apprenants à percevoir leur interaction avec l'ordinateur comme similaire à une conversation humaine. Par exemple, Nass et ses collègues (Moon & Nass, 1996 ; Nass, Moon, Fogg, Reeves, & Dreyer, 1995) ont montré que les individus réagissent aux signaux sociaux verbaux présents dans une interaction textuelle avec un ordinateur comme si ces signaux provenaient d'un humain. Percevoir l'ordinateur comme un partenaire social encourage l'apprenant à s'engager dans un processus de compréhension qui augmente la probabilité d'un transfert positif (Mayer et al., 2003). Selon Ginns et al. (2013), modifier les matériels d'apprentissage pour adopter un style plus informel et conversationnel augmente la probabilité de cette réponse d'activation sociale, motivant davantage l'apprenant.

Dans ce contexte, il est supposé que les indices sociaux puissent amorcer l'activation des réponses sociales chez les apprenants, notamment l'engagement à donner du sens à ce que l'orateur dit, comme le souligne Mayer (2005). Une réponse sociale devrait conduire à une motivation accrue, souvent désignée par l'intérêt à s'engager dans un traitement cognitif actif et, par conséquent, à un traitement plus profond lors de l'apprentissage. Plusieurs études ont démontré que la présence d'agents pédagogiques dans les environnements d'apprentissage multimédia pouvait améliorer les performances des apprenants, en particulier lorsque ces agents présentaient des caractéristiques sociales réalistes (Schroeder et al., 2013 ; Wang et al., 2018).

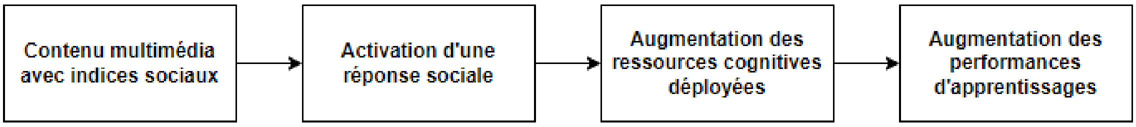

Figure 3 : Illustration du mécanisme d'activation social tirée de la théorie de l'agentivité sociale (Mayer, 2005)

A partir de ce cadre théorique, la question de la présence de l'agent pédagogique, comme un instructeur virtuel à l'écran, accompagnant l'apprenant se justifie théoriquement étant donné son apport social et ses effets sur l'engagement de l'apprenant.

5. Agents pédagogiques

Un des principes de l'apprentissage multimédia reposant sur cette théorie de l'agentivité est la présence d'agents pédagogiques, définis comme des personnages virtuels qui servent de guides pédagogiques ou de facilitateurs dans l'environnement d'apprentissage. Les agents pédagogiques sont conçus pour fournir un soutien supplémentaire, améliorer l'engagement des apprenants et promouvoir une acquisition efficace des connaissances (Schroeder et al., 2013).

Les agents pédagogiques peuvent avoir un impact significatif sur l'expérience d'apprentissage. L'agent maintient l'engagement des élèves (Dunsworth et Atkinson, 2007 ; Kim et Baylor, 2006 ; Woo, 2008) et confèrent de la crédibilité aux activités de l'environnement d'apprentissage (Towns, Voerman, Callaway, & Lester, 1998). L'utilisation des agents pédagogiques peut également améliorer les résultats non cognitifs des élèves, tels que la motivation et l'engagement (Chin, Hong, Huang, Shen et Lin, 2016).

Cependant, les recherches dans ce domaine révèlent des résultats variés, parfois contradictoires, concernant l'efficacité des agents pédagogiques (Castro-Alonso, 2021). Ces différences peuvent être attribuées à la diversité des conceptions et des contextes d'utilisation des agents (Davis, 2018 ; Schroeder, Adesope et Gilbert, 2013). Il devient alors évident que le design de l'agent pédagogique joue un rôle déterminant dans l'optimisation de son impact sur l'apprentissage. Une attention particulière doit donc être accordée à la manière dont ces agents sont conçus et intégrés dans l'environnement d'apprentissage. L'un des axes de recherche dans ce domaine concerne la "force" de l'agent, terme se référant au degré d'animation et d'expressivité des agents pédagogiques (Beege,2023). Il est suggéré que des agents dotés d'une grande expressivité et capables d'imiter les comportements humains, tels que les expressions faciales ou les gestes, peuvent créer une connexion plus forte avec les apprenants (Abdullah et al., 2017). Par exemple, Mayer et DaPra (2012) ont montré que les étudiants apprenaient mieux à partir d'une animation narrée lorsque le narrateur était représenté par un agent pédagogique au comportement non-verbal naturel (expressions faciales, gestes) plutôt que par une simple voix. De même, une méta-analyse de Schroeder et Adesope (2014) a révélé que les agents pédagogiques ayant une apparence humaine réaliste et exprimant des émotions positives favorisaient davantage l'apprentissage que les agents au design plus minimaliste ou neutre.

Pour conclure sur l'impact du design de ces agents pédagogiques, il est également à noter que la simple présence, l'animation et la complexité des agents peuvent parfois agir comme un facteur de distraction, détournant potentiellement l'attention des apprenants du contenu pédagogique (Schroeder et Adesope, 2013), compromettant ainsi les performances d'apprentissage (Park, 2015 ; Ayres & Sweller, 2005). L'impact des agents peuvent également s'avérer délétères en introduisant une charge cognitive supplémentaire, dans laquelle les apprenants doivent partager leur attention entre l'agent et le contenu (Sweller, 1988). Les agents pédagogiques efficaces devraient diriger l'attention vers un contenu pertinent sans introduire d'éléments superflus et gênants qui nécessiteraient des processus supplémentaires, tels que l'inhibition, pour contrôler les interférences potentielles. Le besoin d'inhibition peut épuiser les ressources attentionnelles disponibles (Michael et al., 2006), réduisant ainsi l'efficacité de l'apprentissage. L'agent pédagogique doit donc répondre à un double enjeu : engager les apprenants par le biais de sa « force », mais également de guider leurs attentions par sa conception et ses gestes.

C. Engagement

1. Cadre théorique

L'engagement des apprenants peut être défini comme l'effort, l'intérêt, la persévérance dans la tâche (Molinari et al., 2016). L'approche principale de l'engagement est une approche par composantes. L'engagement est un concept composé de trois dimensions complémentaires et en interaction : affective, comportementale et cognitive (Fredricks et al., 2004 ; Linnenbrink et Pintrich, 2003).

L'engagement cognitif est lié au déploiement de stratégies d'apprentissage (cognitives, autorégulées ou liées à la gestion des ressources) et à l'effort mental déployé lors de la réalisation d'une tâche (Pintrich, 1999). Il peut être envisagé sous un angle quantitatif (quantité de ressources allouées à la tâche, intensité des efforts) et qualitatif (degré de sophistication des stratégies d'apprentissage, adéquation des efforts). Ainsi, un apprenant peut investir beaucoup de ressources dans la mise en œuvre d'une stratégie peu sophistiquée ou une même quantité de ressources dans une stratégie plus sophistiquée. La notion d'engagement cognitif renvoie, ainsi, aux fonctions attentionnelles nécessaires à la réalisation d'une tâche (Fredricks, 2004).

L'engagement affectif inclut l'intérêt, les émotions et les valeurs perçues par les apprenants pendant les activités d'apprentissage (Huang et al., 2022). Il peut se référer aux émotions positives ou négatives (par exemple, l'intérêt, la tristesse ou l'anxiété) que les apprenants ressentent à l'égard de l'école, de leurs enseignants et leurs pairs, ou encore du contenu à apprendre voire de l'acte même d'apprendre.

Enfin, l'engagement comportemental se fonde sur l'idée de participation et d'indicateurs observables de cette participation (Fredricks et al., 2004). La notion de participation renvoie elle-même à différents niveaux : le respect des règles de la classe et l'absence de comportements dérangeants ; l'effort, la concentration, l'attention et la participation aux activités en classe ; la participation à des activités extra-curriculaires. Dans le cas d'un apprentissage en ligne, ces différents niveaux de participation peuvent être analysés à travers l'étude des comportements observables des apprenants au sein de l'environnement numérique. Lavoué et al. (2021) proposent ainsi d'étudier les relations entre la motivation des apprenants et leurs comportements engagés dans un environnement d'apprentissage gamifié. Ils s'appuient notamment sur les travaux de Fincham et al. (2018) qui définissent l'engagement comportemental comme la quantité d'efforts observables qu'un apprenant consacre au processus d'apprentissage, et de Motz et al. (2019) qui soulignent l'importance de la fréquence et de la constance des interactions pour caractériser l'engagement en ligne.

Plus récemment, Reschly et al. (2012) ont proposé un modèle qui représente l'engagement des apprenants à la fois comme un processus et un résultat d'une activité d'apprentissage. Ils ont identifié une quatrième dimension qui correspondrait à l'engagement académique. Selon leur modèle, l'engagement affectif et l'engagement cognitif représentent les perceptions des étudiants, qui se manifestent alors par un engagement académique et comportemental. L'engagement académique se distingue de l'engagement comportemental par le fait qu'il se concentre spécifiquement sur les indicateurs liés à la performance et à la réussite des apprenants dans les activités d'apprentissage. Les indicateurs utilisés pour mesurer l'engagement académique dans les environnements d'apprentissage en ligne sont assez courants : le temps consacré aux activités de cours (par exemple, consulter des pages, répondre à des questionnaires et à des devoirs), la participation aux cours (ou le nombre de connexions), la précision et le taux de réussite des quiz et des devoirs. En revanche, l'engagement comportemental est davantage lié à la participation et à l'implication de l'apprenant dans les activités d'apprentissage, sans nécessairement tenir compte de la performance ou de la réussite. Il s'agit donc de deux dimensions complémentaires mais distinctes de l'engagement des apprenants. Bien que les définitions exactes des quatre dimensions puissent varier légèrement d'un article à l'autre, elles sont cohérentes avec d'autres modèles de l'engagement (Joksimovic et al., 2018 ; Fincham et al., 2019).

Une autre approche de l'engagement se rapprochant plus des cadres théoriques de psychologie cognitive définit l'engagement situationnel. L'engagement situationnel dans les environnements d'apprentissage est décrit comme un état temporaire d'immersion intense et de concentration sur une activité spécifique (Shernoff, 2013). Ce type d'engagement joue un rôle dans l'amélioration de la compréhension, l'augmentation de la rétention de l'information et la stimulation de la motivation à continuer d'apprendre (Shernoff, 2013). Ce niveau d'engagement est influencé par plusieurs facteurs, notamment l'intérêt pour le sujet et la pertinence perçue de la tâche (Wigfield & Eccles, 2000).

Par ailleurs, le concept de flow, élaboré par Mihaly Csikszentmihalyi (1990), décrit un état de concentration totale où l'individu est tellement absorbé par une activité qu'il en perd la notion du temps. L'atteinte de cet état de flow dans un contexte éducatif peut conduire à une expérience d'apprentissage extrêmement enrichissante, marquée par un plaisir intrinsèque et une efficacité accrue dans l'acquisition des connaissances (Csikszentmihalyi, 1990). Dans les environnements d'apprentissage, le flow peut être facilité par des tâches qui maintiennent un équilibre entre les défis présentés et les compétences de l'apprenant, des objectifs clairement définis, une rétroaction immédiate et une immersion profonde dans la tâche (Shernoff et al., 2013). Bien que l'engagement situationnel et le flow soient des concepts distincts, ils sont étroitement liés. Un haut niveau d'engagement situationnel peut être un précurseur à l'expérience du flow (Shernoff, 2013). Les environnements d'apprentissage qui favorisent cet engagement peuvent ainsi créer des conditions propices à l'expérience du flow. Inversement, vivre l'expérience du flow peut renforcer l'engagement situationnel, rendant l'apprentissage plus gratifiant et motivant (Csikszentmihalyi & Rathunde, 1993).

2. Motivation, engagement et performance

La relation entre motivation et engagement suscite un vif intérêt dans le domaine de l'éducation, plusieurs études soulignant l'association étroite entre ces deux concepts (par exemple, Bempechat & Shernoff, 2012 ; Bingham & Okagaki, 2012) et leur lien avec le bien-être des apprenants (Brault-Labbé & Dubé, 2010 ; Burton, Lydon, D'Alessandro, & Koestner, 2006 ; Reis, Sheldon, Gable, Roscoe, & Ryan, 2000). L'engagement est souvent perçu comme la concrétisation de la motivation, se manifestant par les pensées, actions, et émotions initiées ou ressenties par l'individu en raison de sa motivation (Martin, Ginns et al., 2017). Ainsi, l'engagement agit en tant que reflet de la motivation, avec l'idée que la motivation est à l'origine de l'action, tandis que l'engagement représente l'acte même d'apprendre (Cleary & Zimmerman, 2001). Cette interaction crée une boucle de rétroaction où l'expérience de l'engagement peut renforcer ou modifier les aspects de la motivation, y compris l'auto-efficacité, et influencer la volonté de poursuivre ou de renouveler l'engagement dans des tâches similaires (Sharek, 2016). Bien que la relation entre motivation et engagement soit largement reconnue, la nature spécifique de leur interdépendance reste un sujet d'étude important, avec une littérature suggérant que la motivation est un prérequis essentiel à l'engagement, mais pas suffisant en elle-même (Appleton, Christenson, & Furlong, 2008). Mayer (2014) a souligné le rôle de la motivation dans la promotion d'un engagement générateur, qui à son tour améliore l'apprentissage et la performance. De même, les travaux de Reeve et Lee (2014) ont démontré que les changements dans l'engagement des étudiants entraînent des modifications dans leur motivation de classe, ce qui affecte positivement les résultats académiques.

Si plusieurs recherches ont porté sur l'association entre la motivation ou l'engagement et différentes variables liées à l'expérience d'apprentissage, peu d'études empiriques documentent à ce jour les liens spécifiques qui existent entre les deux concepts (Fredricks et al., 2004). Selon plusieurs auteurs, la motivation constituerait un élément essentiel, mais non suffisant à l'engagement (Appleton, Christenson, & Furlong, 2008), Reeve (2012) va en ce sens en concevant la motivation comme un processus biologique non observable qui deviendrait tangible par l'entremise d'un comportement d'engagement. Parmi les quelques chercheurs ayant tenté d'établir des liens empiriques entre les deux construits chez diverses populations étudiantes, on retrouve notamment les travaux de Skinner, Kindermann et Furrer (2009). Ces derniers ont examiné les relations entre la motivation et l'engagement auprès d'élèves du primaire et du secondaire. Leurs résultats indiquent que la motivation intrinsèque et le sentiment d'auto-efficacité sont positivement associés à l'engagement comportemental et émotionnel des élèves. De plus, ils soulignent que l'engagement agit comme médiateur dans la relation entre la motivation et la réussite scolaire. D'autres études, menées auprès d'étudiants universitaires, corroborent ces résultats. Par exemple, une recherche de Rothes et al. (2022) a montré que la motivation autodéterminée et les objectifs personnels de réussite jouent un rôle crucial dans l'engagement des étudiants. Ils ont démontré que les objectifs de maîtrise médiatisent la relation entre la motivation autonome et les résultats éducatifs, tandis que les objectifs de performance-évitement médiatisent la relation entre la régulation introjectée, la régulation externe et l'engagement comportemental et émotionnel des étudiants Ces résultats suggèrent que lorsque les étudiants poursuivent leurs études par intérêt et parce qu'ils en perçoivent la valeur, ils sont plus susceptibles de s'engager activement dans leurs apprentissages.

Cependant, une recherche de Cheon et Reeve (2015) a montré que l'engagement peut fluctuer au fil du temps, indépendamment du niveau de motivation des élèves. Les auteurs ont suivi des élèves du secondaire pendant une année scolaire et ont constaté que leur engagement avait tendance à diminuer au cours de l'année, et ce, même si leur motivation demeurait stable. Ces résultats indiquent que d'autres facteurs, tels que le soutien de l'enseignant ou le climat de classe, peuvent influencer l'engagement au-delà de la motivation individuelle. La question de la relation entre l'engagement, la motivation et l'intérêt suscite un vif débat dans le domaine de l'apprentissage. Bien que ces concepts soient souvent considérés comme étroitement liés, il est crucial de s'interroger sur la nature de leurs interactions et sur la possibilité qu'un apprenant soit engagé dans une tâche sans être nécessairement motivé par le contenu pédagogique en lui-même.

Hidi et Renninger (2006 ; 2011) abordent cette problématique sous l'angle de l'intérêt, un concept distinct mais lié à la motivation. Selon ces auteurs, ainsi que Krapp (2011), l'intérêt est un état psychologique caractérisé par une tendance à focaliser son attention sur un contenu spécifique, résultant d'une interaction dynamique entre l'individu et son environnement. Le développement de l'intérêt serait initié par un processus de "déclenchement" lorsqu'un contenu parvient à capter l'attention de l'apprenant (Dewey, 1913 ; Hidi et Baird, 1986).

Deux formes principales d'intérêt sont distinguées : l'intérêt situationnel, suscité par les caractéristiques de l'activité (nouveauté, créativité, défi, thème, possibilité de choix), et l'intérêt individuel, qui renvoie aux préférences personnelles pour un sujet donné et influence l'attention, l'effort, la persévérance et l'acquisition de connaissances (Hidi & Renninger, 2006 ; Pintrich, Marx & Boyle, 1993 ; Schiefele, 1991).

L'intérêt situationnel, en particulier, pourrait être considéré comme un facteur d'engagement situationnel par sa capacité à capter l'attention de l'apprenant. Un stimulus à la fois nouveau et pertinent peut accroître cet intérêt, ainsi que l'attention et les processus cognitifs associés, tels que la mémorisation et la compréhension (Darst & Pangrazi, 2001). Cependant, la relation entre l'activation physiologique et psychologique provoquée par ce stimulus et l'attention n'est pas linéaire mais curvilinéaire, une activation excessive pouvant entraîner une diminution de l'attention.

La question centrale qui émerge alors est de savoir si cet intérêt situationnel influence directement l'engagement sans nécessairement impacter la motivation de l'apprenant, étant donné que l'intérêt et la motivation sont des concepts distincts bien que liés. Plusieurs études suggèrent que l'intérêt situationnel peut effectivement favoriser l'engagement sans modifier la motivation intrinsèque de l'individu (Hidi & Renninger, 2006 ; Rotgans & Schmidt, 2011), car il est souvent de courte durée et dépend de stimuli externes, tandis que la motivation intrinsèque est une disposition plus stable ancrée dans les préférences personnelles.

Cependant, si l'intérêt situationnel est maintenu et renforcé au fil du temps, il peut potentiellement évoluer vers un intérêt individuel plus durable et ainsi influencer positivement la motivation intrinsèque (Hidi & Renninger, 2006). Cette transition dépend de facteurs tels que les opportunités d'engagement répétées, le soutien de l'environnement d'apprentissage et les expériences de maîtrise (Renninger & Su, 2012).

En somme, il est possible qu'un apprenant soit engagé et intéressé par une tâche sans être motivé intrinsèquement par le contenu, l'intérêt situationnel agissant principalement sur l'engagement à court terme. Cette analyse souligne la complexité des relations entre intérêt, engagement et motivation, et la nécessité de considérer ces concepts de manière nuancée dans les recherches sur l'apprentissage, en tenant compte du fait que la motivation et l'engagement, bien qu'essentiels, ne sont pas toujours directement liés.

D. L’attention dans l'Apprentissage Multimédia

L'attention, définie comme la capacité à se concentrer sur une tâche ou un stimulus spécifique tout en ignorant les distractions (James, 1890), joue un rôle prépondérant dans la compréhension des mécanismes d'apprentissage multimédia. Elle mobilise tant les capacités intrinsèques de l'individu que les stratégies pédagogiques mises en œuvre pour optimiser l'engagement et la rétention de l'information. L'attention peut être considérée comme un processus cognitif complexe qui implique la sélection, la focalisation et le maintien de la concentration sur les informations pertinentes (Posner & Petersen, 1990). Dans le contexte de l'apprentissage multimédia, l'attention est cruciale pour filtrer et traiter efficacement les multiples sources d'information présentées simultanément, telles que le texte, les images et les sons (Mayer, 2009).

Cette partie de notre manuscrit se propose d'explorer les dimensions complexes et multifacettes de l'attention, envisagée à la fois comme une capacité cognitive et comme un processus dynamique déployé au sein des environnements d'apprentissage multimédia. Dans un premier temps, nous aborderons les fondements théoriques qui sous-tendent la notion d'attention dans le cadre de l'apprentissage multimédia. Nous nous intéresserons particulièrement aux capacités attentionnelles telles que l'inhibition et la gestion des ressources cognitives, éléments clés permettant de comprendre comment les apprenants parviennent à filtrer les informations pertinentes et à les intégrer efficacement (Mayer, 2009 ; Sweller, Ayres, & Kalyuga, 2011). L'inhibition attentionnelle, qui permet de supprimer les informations non pertinentes ou distrayantes, est essentielle pour maintenir la concentration sur les éléments clés du contenu pédagogique (Diamond, 2013). De même, la gestion efficace des ressources cognitives, qui implique l'allocation de l'attention en fonction de la demande de la tâche, est primordiale pour éviter une surcharge cognitive et optimiser l'apprentissage (Sweller et al., 2011).

Ensuite, la présentation du modèle Master Activation Map (MAM) offrira un cadre explicatif quant au déploiement de l'attention, et plus particulièrement appliqué à l’apprentissage multimédia. La relation entre l'engagement et l'attention sera également traitée, mettant en lumière les synergies entre l'implication active de l'apprenant et sa capacité à maintenir une attention soutenue.

Cette analyse sera étayée par une revue de la littérature existante, ainsi que par la proposition d'un modèle intégratif visant à expliciter les liens entre ces concepts fondamentaux dans l'efficience de l'apprentissage multimédia.

1. Inhibition

1.1. Cadre théorique

L'inhibition est un processus cognitif permettant de supprimer ou de moduler des informations, des réponses, des pensées ou des émotions non pertinentes ou indésirables (Michael et al., 2006). Ce mécanisme joue un rôle central dans le contrôle attentionnel et le traitement de l'information, en permettant de résister à l'interférence et de maintenir un traitement efficace des stimuli pertinents (Kok, 1999).

Le modèle cognitif proposé par Michael et al. (2006) conçoit l'inhibition comme un processus actif qui implique la suppression des informations non pertinentes. Selon ce modèle, l'inhibition est étroitement liée à la mémoire de travail, car elle nécessite le maintien des objectifs et des règles en mémoire pour guider le comportement. Les auteurs soulignent également que l'inhibition est un processus flexible et adaptable, qui peut être modulé en fonction des demandes de la tâche et du contexte. Cette flexibilité est soutenue par des données neurophysiologiques montrant que l'activité des régions cérébrales impliquées dans l'inhibition, telles que le cortex préfrontal et le cortex cingulaire antérieur, peut être modulée en fonction des exigences de la tâche (Michael et al., 2006).

Kok (1999) propose une analyse détaillée des différents niveaux auxquels l'inhibition peut opérer dans le traitement de l'information. Selon cet auteur, l'inhibition peut intervenir au niveau de la sélection des informations pertinentes, en supprimant le traitement des stimuli non pertinents. Ce processus est particulièrement important dans les situations où de multiples stimuli sont en compétition pour les ressources attentionnelles limitées. L'inhibition peut également opérer au niveau de la suppression des informations non pertinentes qui ont déjà été traitées, évitant ainsi qu'elles n'interfèrent avec le traitement en cours. Enfin, l'inhibition joue un rôle crucial dans le contrôle des réponses, en permettant de supprimer les réponses inappropriées. Ces différents niveaux d'inhibition sont soutenus par des données comportementales et neurophysiologiques, montrant par exemple une activation accrue des régions préfrontales lors de la suppression des réponses dans des tâches de type Go/No-Go (Kok, 1999).

Les patterns développementaux de l'inhibition active ont été explorés en détail par Michael et al. (2013). Ces auteurs ont mené une méta-analyse des études portant sur le développement de l'inhibition, en se focalisant sur les tâches de type Go/No-Go, Stroop et Stop-Signal. Les résultats montrent une amélioration marquée des capacités d'inhibition au cours de l'enfance et de l'adolescence, avec des gains particulièrement importants entre 6 et 12 ans. Ces améliorations sont associées à des changements dans l'activité des réseaux neuronaux impliqués dans le contrôle cognitif, notamment une augmentation de l'activation du cortex préfrontal et une diminution de l'activation des régions sous-corticales telles que le striatum (Michael et al., 2013). Ces résultats suggèrent que le développement de l'inhibition reflète une maturation progressive des circuits neuronaux sous-tendant le contrôle cognitif, avec un shift progressif vers un contrôle top-down médié par les régions préfrontales.

1.2. Hétérogénéité de l'inhibition

Les variations individuelles dans l'efficacité de l'inhibition, comme le soulignent Harnishfeger et Bjorklund (1994), contribuent à des différences individuelles dans le traitement cognitif. Harnishfeger et Bjorklund (1994) ont mis en lumière l'importance des variations individuelles dans l'efficacité de l'inhibition. Ces variations influencent la capacité d'une personne à contrôler ses réponses aux stimuli externes et internes. Des facteurs tels que l'âge, le développement neurologique, et certaines conditions psychologiques ou neurologiques peuvent affecter cette capacité.

Dans le cadre des cas typiques de variation de l'inhibition, les recherches de Yang et al. (2022) ont mis en lumière un affaiblissement notable de l'inhibition avec l'âge. Cette tendance est corroborée par les travaux de Campbell, Lustig et Hasher (2020), qui mettent en évidence des différences significatives dans le contrôle inhibiteur liées à l'âge, marquées par une capacité variable à supprimer des stimuli, des pensées et des actions non pertinentes. En outre, l'influence du niveau d'éducation sur l'efficacité de l'inhibition est également notable (Aron, 2007). Les individus ayant atteint un niveau d'éducation supérieur tendent à présenter de meilleures capacités d'inhibition, probablement en raison du développement cognitif accru et de l'exercice mental régulier associé à des études avancées.

Parallèlement, au sein de la population générale, on observe une large variabilité des capacités d'inhibition, même en dehors des considérations liées à l'âge ou au niveau d'éducation.

Cette diversité peut être attribuée à des facteurs génétiques, à des influences environnementales variées, et même à l'état psychologique et émotionnel des individus, soulignant ainsi la complexité et la nature multifactorielle de l'inhibition. Ces divers aspects de l'hétérogénéité de l'inhibition mettent en évidence l'importance de stratégies individualisées dans des domaines tels que l'éducation et la santé mentale, et soulignent la nécessité d'une compréhension approfondie des capacités d'inhibition pour une application efficace dans ces contextes.

2. Gestion des Ressources Attentionnelles

2.1. Cadre théorique

Selon le modèle de Khaneman (1973), l'attention humaine est conceptualisée comme un ensemble de ressources finies qui doivent être allouées judicieusement pour optimiser le traitement cognitif. Ce modèle se fonde sur l'idée que les capacités attentionnelles humaines sont limitées et ne peuvent être pleinement investies dans plusieurs tâches simultanément. Cette approche marque un départ significatif des théories antérieures, telles que celle de Broadbent (1958), qui envisageaient l'attention principalement comme un mécanisme de filtrage.

Dans le modèle de Kahneman, la distribution des ressources attentionnelles se fait en fonction des exigences de la tâche. Les tâches qui sont plus complexes ou qui requièrent un niveau d'engagement plus élevé consommeront une plus grande part de ces ressources. Cette distribution n'est pas statique mais dynamique, s'adaptant continuellement aux changements dans l'environnement et aux priorités de l'individu. L'efficacité avec laquelle une tâche est exécutée dépend directement de la quantité de ressources attentionnelles qui lui sont allouées. Une tâche qui reçoit une grande part de ces ressources sera traitée plus efficacement et profondément. Inversement, des tâches traitées avec des ressources limitées peuvent ne pas être exécutées avec la même qualité ou peuvent même être négligées.

Kahneman aborde également la question de l'attention partagée. Lorsqu'un individu est confronté à plusieurs tâches, les ressources attentionnelles doivent être partagées entre ces différentes exigences. Cela peut conduire à une diminution de l'efficacité du traitement pour chaque tâche individuelle, car les ressources ne sont pas infinies et ne peuvent être pleinement dédiées à plusieurs processus simultanément. La gestion des ressources attentionnelles, d'autre part, se rapporte à la manière dont nous allouons notre attention limitée aux multiples stimuli et tâches qui exigent notre concentration.

2.2 Hétérogénéité de la gestion des ressources

L'attention fluctue naturellement, fréquemment et de manière irrégulière et imprévisible en termes d'intensité. Cette intensité varie tout au long de la journée avec de longues périodes où l'attention peut être maintenue à un haut niveau et des périodes "creuses" pendant lesquelles maintenir son attention est difficile. Incrustées dans ces longues périodes de fluctuation, nous trouvons également des variations sur de bien plus petites périodes de temps. Nous pouvons ainsi imaginer de longues périodes d'attention hautement performantes à l'intérieur desquelles il y a de petits moments de relâchement et d'éclipse. Si une information critique survient pendant ces périodes de relâchement, son traitement ne sera pas optimal (Michael et al., 2015). La seconde dimension concerne l'objet sur lequel porte l'attention. Celle-ci peut être portée soit sur des événements extérieurs, qu'il s'agisse de la tâche à accomplir, soit sur des événements internes, comme par exemple ses propres pensées. L'intensité et l'objet de l'attention interagissent.

Parfois, les fluctuations d'intensité coïncident avec la fluctuation de l'objet de l'attention : nous sommes moins attentifs sur le contenu d'un enseignement car l'intensité est présentement faible et l'attention s'est tournée profondément vers nos propres pensées, et inversement. Cette structure complexe fait en sorte à ce que contrôler délibérément son attention dans tout contexte et de manière permanente soit difficile à accomplir, et la conséquence n'est autre que la diminution de la performance (Couffe & Michael, 2017) que cela soit dans un contexte d'activité professionnelle (Brazzolotto & Michael, 2019) ou des activités d'apprentissage (Farley et al., 2013). Il y a cependant des marges d'optimisation qui sont définies à la fois par les capacités individuelles et par le fonctionnement naturel de l'attention (Stevens & Bavelier, 2012). Nous pouvons ainsi améliorer les activités d'apprentissage en apprenant à orienter et à maintenir l'attention, toujours dans les limites fixées par les différences individuelles.

2.3. Lien entre ressources attentionnelles et inhibition

L'inhibition et la gestion des ressources attentionnelles sont deux processus cognitifs étroitement liés, qui interagissent de manière complexe pour permettre un traitement efficace de l'information et une adaptation optimale aux demandes de l'environnement. Comprendre la nature de cette interaction est essentiel pour appréhender le fonctionnement de l'attention et son rôle dans l'apprentissage, en particulier dans le contexte des environnements multimédias où les apprenants sont confrontés à de multiples sources d'information concurrentes.

Comme le souligne Farrell (2016), les processus d'inhibition sont cruciaux pour un fonctionnement cognitif efficace dans une variété de tâches intellectuelles, notamment la mémoire de travail et le raisonnement. En effet, l'inhibition permet de supprimer les informations non pertinentes ou distractrices, libérant ainsi des ressources attentionnelles qui peuvent être allouées aux stimuli et aux tâches pertinentes. Cette interaction entre l'inhibition et la gestion des ressources est particulièrement importante pour comprendre comment nous maintenons notre concentration sur des tâches spécifiques tout en inhibant les distractions.

Selon le modèle de Kahneman (1973), les ressources attentionnelles sont limitées et doivent être allouées de manière sélective en fonction des demandes de la tâche et des priorités de l'individu. Dans ce contexte, l'inhibition joue un rôle clé en permettant de filtrer les informations non pertinentes et de prioriser le traitement des stimuli les plus importants. Ainsi, un individu disposant de bonnes capacités d'inhibition sera en mesure de mieux gérer ses ressources attentionnelles, en les allouant de manière optimale aux tâches et aux informations pertinentes.

Inversement, des déficits d'inhibition peuvent conduire à une mauvaise gestion des ressources attentionnelles, car l'individu aura des difficultés à supprimer les distractions et à maintenir sa concentration sur les éléments clés. C'est notamment le cas dans certains troubles neurodéveloppementaux comme le TDAH, où les difficultés d'inhibition sont associées à une tendance à se laisser distraire par des stimuli non pertinents et à une difficulté à maintenir une attention soutenue (Barkley, 1997).

Au niveau neurophysiologique, l'interaction entre l'inhibition et la gestion des ressources attentionnelles est sous-tendue par des réseaux neuronaux largement distribués, impliquant notamment le cortex préfrontal et les aires pariétales (Banich et al., 2000). Ces régions cérébrales sont impliquées à la fois dans les processus d'inhibition et dans le contrôle attentionnel, suggérant une forte intrication entre ces deux fonctions au niveau neural.

La propagation de l'inhibition, selon Neumann et al. (1996), se fait de manière analogue à l'activation dans la mémoire de travail. Cette analogie suggère que l'inhibition et l'activation sont deux processus qui fonctionnent en tandem, chacun maximisant ses fonctions dans le cadre de leurs limites respectives de ressources. Ainsi, une inhibition efficace permet de libérer des ressources attentionnelles qui peuvent être allouées à l'activation et au maintien des informations pertinentes en mémoire de travail.

Dans le contexte de l'apprentissage multimédia, la relation entre l'inhibition et la gestion des ressources attentionnelles est particulièrement cruciale. En effet, les apprenants sont souvent confrontés à de multiples sources d'information (texte, images, sons) qui peuvent entrer en compétition pour les ressources attentionnelles limitées. Une bonne capacité d'inhibition permettra à l'apprenant de filtrer les informations non pertinentes, de résister à la distraction et de focaliser ses ressources sur les éléments clés du message pédagogique. À l'inverse, des difficultés d'inhibition pourront conduire à une surcharge cognitive et à une dispersion de l'attention, nuisant à la compréhension et à la rétention des informations.

En somme, l'inhibition et la gestion des ressources attentionnelles sont deux processus cognitifs étroitement liés, qui interagissent de manière complexe pour permettre un traitement efficace de l'information. Comprendre la nature de cette interaction est essentiel pour optimiser les apprentissages, en particulier dans le contexte des environnements multimédias où les apprenants sont confrontés à de multiples sources d'information concurrentes. En favorisant le développement des capacités d'inhibition et en concevant des supports pédagogiques qui tiennent compte des limites des ressources attentionnelles, il est possible de créer des conditions propices à un apprentissage efficace et durable.

3. Déploiement des ressources attentionnelles

La question du déploiement des ressources attentionnelles est au cœur des recherches sur l'attention visuelle. Comment notre système cognitif parvient-il à allouer ses ressources limitées pour traiter efficacement le flux d'informations visuelles ? Quels mécanismes sous-tendent la sélection des éléments pertinents et le filtrage des distracteurs ? Le modèle MAM (Master Activation Map), proposé par Michael et al. (2006, 2013, 2014, 2020), offre un cadre théorique pour aborder ces questions.

L'originalité du MAM réside dans sa conception de l'attention comme un processus dynamique et multifacette, impliquant l'interaction de différents sous-systèmes et étapes de traitement. Selon ce modèle, le déploiement attentionnel émerge de la compétition entre des signaux de saillance bottom-up, reflétant les propriétés des stimuli, et des influences top-down, guidées par les objectifs et les intentions de l'observateur.

Un apport clé du MAM est de proposer une architecture cognitive détaillée pour rendre compte de cette dynamique. L'information visuelle serait d'abord traitée en parallèle par deux voies distinctes, spécialisées respectivement dans les dimensions spatiales et non spatiales (couleur, forme). La voie non spatiale jouerait un rôle critique en calculant les différences entre les éléments du champ visuel pour chaque dimension perceptive. Ces différences seraient ensuite représentées sous forme de signaux sur une carte de saillance, mettant en évidence les éléments les plus distinctifs.

Cette carte de saillance constituerait une première étape dans le déploiement attentionnel, en orientant les ressources vers les zones les plus informatives du champ visuel. Cependant, la saillance seule ne suffit pas à guider efficacement l'attention. C'est ici qu'intervient la carte MAM, une carte cognitive intégrative qui combine les signaux de saillance bottom-up avec des influences top-down générées par une carte de pertinence. Cette dernière coderait les zones du champ visuel les plus pertinentes en fonction des buts de l'observateur.

L'interaction entre ces différentes cartes au sein du MAM rendrait compte de la dynamique complexe du déploiement attentionnel. La compétition entre les signaux de saillance et de pertinence déterminerait quelle zone du champ visuel va recevoir prioritairement les ressources attentionnelles. Le signal le plus fort sur la carte MAM guiderait l'orientation de l'attention et des mouvements oculaires vers la localisation correspondante.

Un autre apport du MAM est de souligner le rôle clé de l'inhibition dans la gestion des ressources attentionnelles. Là où de nombreux modèles ont tendance à considérer l'inhibition comme un processus secondaire ou réactif, le MAM lui confère un statut de premier plan. L'inhibition y est conçue comme un processus actif et indépendant, au même titre que l'orientation attentionnelle. Selon le modèle, l'inhibition, guidée par des influences top-down, modulerait en continu la compétition entre les signaux au sein de la carte MAM. Son rôle serait de favoriser les signaux pertinents par rapport aux distracteurs, permettant ainsi un déploiement optimisé des ressources attentionnelles. Lorsque plusieurs stimuli se disputent l'attention, l'inhibition interviendrait pour réduire l'impact des stimuli non pertinents, permettant à l'individu de focaliser ses ressources sur les éléments les plus importants au regard de ses objectifs.

Ce mécanisme inhibiteur serait crucial pour maintenir une allocation efficiente des ressources dans des environnements complexes et riches en distracteurs potentiels, comme les interfaces d'apprentissage multimédia. En filtrant en continu les informations non pertinentes, l'inhibition préviendrait la surcharge cognitive et permettrait un engagement optimal des ressources attentionnelles au service des buts d'apprentissage.

Le MAM offre ainsi une grille de lecture pertinente pour analyser les défis attentionnels posés par les environnements d'apprentissage multimédia. Il suggère que l'efficacité de ces environnements dépend de leur capacité à guider de manière optimale le déploiement des ressources attentionnelles de l'apprenant, en combinant judicieusement des éléments de guidage bottom-up (saillance perceptive) et top-down (consignes, objectifs).

.

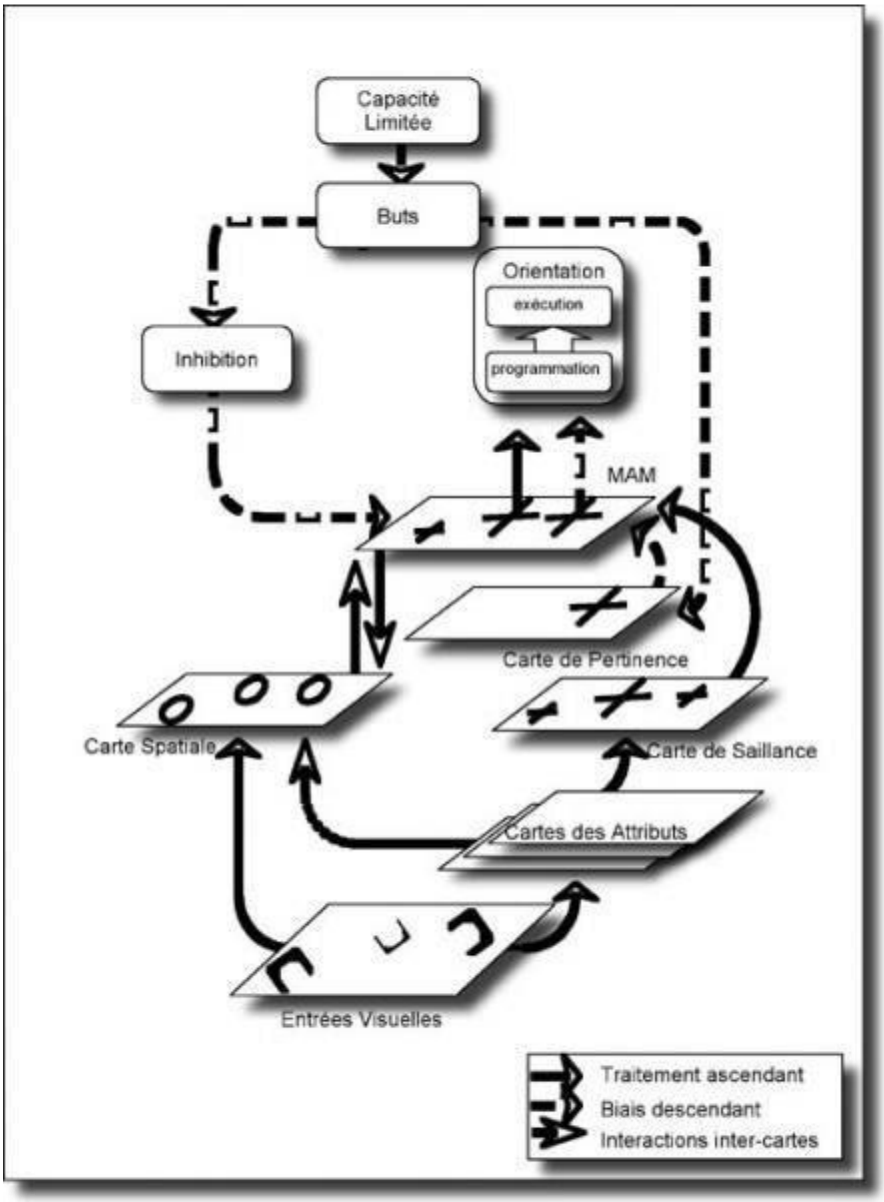

Figure 4 : Schéma du modèle MAM

4. L'attention dans l'apprentissage multimédia

L'attention permet de se centrer sur les informations importantes et d'éliminer ou de diminuer l'impact d'informations secondaires ou sans lien avec les activités en cours. Les bruits et voix de fond dans une classe, les illustrations et animations présentés sur les diapositives d'un support électronique, les fenêtres pop-up intempestives sont autant de signaux qui peuvent interrompre la tâche en cours et détourner l'attention pour une plus ou moins longue période (Couffe & Michael, 2017 ; Michael G.A. & Minne Y. ;2024). Cependant, rester attentif n'est pas chose aisée. L'Humain est doté de systèmes naturellement dédiés à l'interruption des activités en réponse à des sollicitations externes (Michael et al., 2001 ; 2014 ; 2015) dans le sens où la prise en compte des changements de l'environnement immédiat permettrait une meilleure appréhension de la situation afin de réagir de façon appropriée et d'adapter son comportement en conséquence (Couffe & Michael, 2017 ; Michael & Couffe, 2018). Mais ce comportement normal peut être pénalisant pour les activités en cours selon la nature de l'information qui est à l'origine de l'interruption.

Dans la cadre d'un apprentissage multimédia, l'abondance de stimuli visuels et auditifs peut représenter un défi en termes d'attention (Sweller, 2011). Les apprenants peuvent être confrontés à la nécessité de traiter simultanément plusieurs formes d'informations, ce qui peut les rendre plus vulnérables aux distractions. Par exemple, dans le cas de surcharge sensorielle, lorsqu'il y a trop d'informations à traiter en même temps, il devient difficile pour l'apprenant de se concentrer sur l'essentiel. Les éléments visuels et auditifs peuvent se concurrencer, ce qui peut entraîner une perte d'attention. Ce phénomène est généralement connu sous le terme d'attention divisée. La présence de multiples médias peut inciter les apprenants à diviser leur attention entre différentes sources d'information, ce qui peut nuire à leur compréhension globale.

Les distractions, qu'elles proviennent de l'environnement physique ou numérique sont nombreuses. Les recherches ont montré que les individus ayant une capacité d'inhibition plus élevée sont plus susceptibles de réussir dans les tâches d'apprentissage multimédia (Gazzaley & Nobre, 2012). La capacité de gestion des ressources, quant à elle, se réfère à la capacité d'allouer efficacement les ressources cognitives disponibles entre les différentes tâches et processus impliqués dans l'apprentissage multimédia (Paas, Tuovinen, Tabbers, & Van Gerven, 2016). Cela inclut la répartition de l'attention entre les modalités visuelles et auditives, ainsi que la gestion des ressources de travail et de mémoire à long terme. Les apprenants doivent être capables de coordonner ces ressources pour traiter et intégrer les informations multimédias de manière efficace et sans surcharge cognitive (Sweller, Ayres, & Kalyuga, 2011). Les recherches ont montré que les individus ayant une capacité de gestion des ressources plus élevée sont plus susceptibles de réussir dans les tâches d'apprentissage multimédia (Paas et al., 2003). Cela peut être réalisé en optimisant la présentation des informations, en utilisant des techniques de segmentation et de pré-entraînement, et en adaptant la difficulté des tâches aux compétences et aux connaissances des apprenants (Mayer, 2005).

E. Points d'ombre et objectifs de recherche

Au cours de cette introduction, nous avons présenté le cadre théorique de l'apprentissage multimédia, en mettant l'accent sur la théorie cognitive de l'apprentissage multimédia de Mayer (2005) et ses extensions intégrant les aspects émotionnels et sociaux. Nous avons également examiné les concepts d'engagement et d'attention, ainsi que leur rôle dans l'efficacité de l'apprentissage multimédia.

Malgré les avancées significatives dans ce domaine, plusieurs zones d'ombre subsistent. Tout d'abord, la nature précise des relations entre la motivation, l'engagement et l'intérêt dans le contexte de l'apprentissage multimédia reste à clarifier. Bien que ces concepts soient étroitement liés, il est possible qu'un apprenant soit engagé dans une tâche sans être nécessairement motivé par le contenu pédagogique en lui-même. L'intérêt situationnel, suscité par les caractéristiques de l'activité, pourrait influencer directement l'engagement sans modifier la motivation intrinsèque de l'individu. Des recherches supplémentaires sont nécessaires pour mieux comprendre ces interactions complexes.

Ensuite, le rôle spécifique de l'attention dans l'apprentissage multimédia mérite d'être approfondi. Si l'importance de l'attention a été soulignée, ses interactions avec d'autres processus cognitifs comme l'inhibition et la gestion des ressources attentionnelles restent à clarifier. De plus, l'impact des éléments de design émotionnel et social, tels que l'anthropomorphisme et les agents pédagogiques, sur le déploiement attentionnel des apprenants nécessite des investigations supplémentaires.

Enfin, la question de l'hétérogénéité interindividuelle dans les capacités d'inhibition et de gestion des ressources attentionnelles est un point crucial à considérer. Les variations liées à l'âge, au niveau d'éducation ou à certaines conditions psychologiques et neurologiques peuvent influencer significativement l'efficacité de l'apprentissage multimédia. Il est essentiel de mieux comprendre ces différences individuelles pour concevoir des environnements d'apprentissage adaptés et inclusifs.

Face à ces zones d'ombre, notre objectif principal est de proposer et de valider un modèle intégratif unifiant les différents aspects de l'engagement et de l'attention dans l'apprentissage multimédia. Ce modèle visera à clarifier les relations entre ces concepts clés, en tenant compte de leur interaction avec la motivation, l'intérêt situationnel et les capacités d'inhibition et de gestion des ressources attentionnelles.

Pour atteindre cet objectif, nous procéderons en deux étapes. Dans un premier temps, nous réaliserons une revue systématique de la littérature existante sur ces thématiques, afin d'identifier les points de convergence et les lacunes dans les connaissances actuelles. Cette analyse approfondie nous permettra de dégager les composantes essentielles de notre modèle intégratif et de formuler des hypothèses quant aux relations entre ces composantes.

Dans un second temps, nous mettrons à l'épreuve ce modèle à travers une série de trois études expérimentales complémentaires :

- La première étude examinera l'impact de l'anthropomorphisme sur l'engagement, la motivation et les performances d'apprentissage des apprenants. En manipulant le niveau d'anthropomorphisme d'un environnement d'apprentissage multimédia et en mesurant l'engagement situationnel, cette étude permettra de clarifier les liens entre design émotionnel, états motivationnels, engagement et performances.

- La deuxième étude explorera plus spécifiquement le rôle des agents pédagogiques, en faisant varier leur présence et leur niveau d'expressivité. En analysant l'interaction entre ces caractéristiques de design et les capacités attentionnelles des apprenants, cette étude offrira un éclairage sur les conditions d'efficacité des agents pédagogiques sur l'engagement et les performances d'apprentissage en fonction des profils attentionnels.

- Enfin, la troisième étude s'intéressera à l'apport des technologies d'intelligence artificielle pour développer des agents pédagogiques interactifs et adaptatifs. En comparant leur influence sur l'engagement et les performances d'apprentissage par rapport à des agents traditionnels, cette étude questionnera le potentiel de l'IA pour optimiser l'expérience d'apprentissage multimédia.

À travers ces trois études, nous explorerons donc l'impact de différents éléments de design émotionnel et social sur l'engagement, la motivation et le déploiement attentionnel des apprenants, tout en prenant en compte le rôle des différences individuelles dans les capacités d'inhibition et de gestion des ressources attentionnelles.

Les résultats de ces travaux permettront de valider et d'affiner notre modèle intégratif, offrant ainsi une compréhension plus fine des mécanismes sous-jacents à un apprentissage multimédia efficace. Ils fourniront également des pistes concrètes pour optimiser la conception des environnements d'apprentissage, en prenant en compte les spécificités attentionnelles et motivationnelles des apprenants. À terme, notre ambition est de contribuer à l'émergence de stratégies pédagogiques innovantes, adaptées aux défis cognitifs et émotionnels de l'apprentissage multimédia.

Chapitre 2. Modèle Intégratif de l'engagement et de l'attention

Integrating Attention and Engagement in Multimedia Learning: a Narrative Review

Y. Minne [a; b; c], G. A. Michael [a], C. Thiery [b], E. Lavoué [c]

[a] *Unité de Recherche Étude des Mécanismes Cognitifs, Université de Lyon, Université Lyon 2, Lyon, Lyon, France*

[b] *Sydo, 69001 Lyon, France*

[c] *Iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, Université de Lyon, Université Lyon 3, Lyon, France*

Abstract:

The literature on multimedia learning is growing with the increasingly systematic use of digital technologies in learning. One of the predominant theories in this framework is the cognitive theory of multimedia learning developed by Mayer (2005), which describes learner's cognition during a learning task. Although it focuses on learners' cognition, it does not integrate or clearly define the concepts of engagement and attention, nor the relationships between them. Therefore, in this paper, we present a review of the literature in the field of multimedia learning and cognitive psychology to identify how the notions of engagement and attention are defined and used. We then propose to integrate these two concepts into a new model of engagement and attention in multimedia learning. We believe that this contribution will make it possible to better align the pedagogical principles that derive from the multimedia learning theory by considering learners' characteristics, multimedia and task features, together with the environment, as antecedents for learners' engagement and deployment of attentional resources.

Keywords: multimedia learning, attention, engagement, design, attentional abilities.

1. Introduction

Learners' enrollment in online courses continues to grow (Jordan, 2015a). The COVID-19 pandemic accelerated this phenomenon with lockdown on large numbers of students. Indeed, the only way to avoid classrooms, colleges, and universities was to continue online indefinitely. Previously, online learning, distance learning, and correspondence courses were traditionally considered non-formal curricula, but it now appears that if these conditions continue steadily, they will gradually overtake the formal education structure. Therefore, in this context, further studies on distance multimedia learning are needed. A well-recognized theory of multimedia learning was proposed by Mayer and Moreno (2003) to study learners' cognition when using a multimedia device.

The present article relies on this theory to address two issues that arise in multimedia learning. First, one of the major problems is the high dropout rate (Jordan, 2014; Patterson & McFadden, 2009; Rice, 2006; Roblyer, 2006). One solution to this problem is to increase learners' engagement. A second issue relates to learners' attention, which varies over long and short periods (Robertson et al. 1997; Guilford, 1927; Mackworth, 1948). Learners do not pay attention to all parts of a multimedia lesson (Mayer & Moreno, 2003; Moreno & Mayer, 2002; Moreno & Mayer, 2000). Therefore, our review aims to identify those factors that allow learners' attention to focus on multimedia content to promote learning, and to remain engaged in the learning activity.

In this article, we review some aspects of attention and engagement, taking Mayer's (2003) cognitive theory of multimedia learning as a theoretical basis. We first present some cognitive theories about multimedia learning to identify the shortcomings regarding the concepts of engagement and attention. We then provide a detailed analysis of the literature on learners' engagement, especially in multimedia learning, followed by an analysis of the attentional processes thought to be involved in this same context. Finally, we propose an integrative model of engagement and attention in multimedia learning. This model allows us to clarify the role of these closely related cognitive processes in multimedia learning. We also believe that it is useful for designers in creating engaging content that encourages deployment of attentional resources.

2. Multimedia learning theories

2.1. Cognitive theory of multimedia learning

Multimedia learning is defined as "an integrative approach to the study of human cognition, education, and technology that encompasses the use of multiple representational formats (e.g., verbal, graphical, pictorial, auditory, and kinesthetic) to present information and to promote learning" (Mayer, 2009, p. 10). Theoretical frameworks on multimedia learning have been proposed in HCI (Human-Computer Interaction), psychology, educational science domains, and so on. In line with our objective to integrate the concepts of attention and engagement in multimedia learning, we focus on the cognitive theory of multimedia learning (Mayer, 2005), in contrast with other theories centered on technology, the goal of which is to study the various existing technologies in the framework of learning and their impact. The cognitive theory of multimedia learning (Mayer, 2005) seems to stand out from the rest (Li et al., 2019). This learner-centered theory focuses on learners' cognition during learning to study how to adapt multimedia to promote human cognition and, based on this, to propose a set of recommendations for pedagogical designs. One of the main postulates of this theory is that we learn better when we have images and texts rather than text alone. This theory is based on three assumptions, detailed below: (i) the principle of active processing (Mayer, 2009; Wittrock, 1989), (ii) dual channel (Paivio, 1971), and (iii) limited processing capacity (Cherry, 1953; Broadbent, 1958) and the cognitive load theory (Sweller, 1988).

The first assumption of Mayer's theory is the principle of active processing, which can be defined as the set of operations performed by the learner on the input to produce a representation that is more useful for learning than the input itself (Mayer, 2014). Meaningful learning takes place when learners engage in cognitive processes that are appropriate for the task (Mayer, 2010). The processes involved include attentional selection of information of the instructional material presented, mental organization into a coherent structure of the selected information, and integration of the information with prior knowledge in long-term memory (Mayer, 2005).

The second assumption is the dual channel. This asserts that human cognition has two different channels for processing–verbal auditory information and pictorial visual information. It thus generates two separate representations of the same information so that a given stimulus can be processed as a word and as an image, which can be used for later recall (Pavio, 1987). The ability to simultaneously encode a stimulus in two different ways increases the probability of remembering that stimulus, compared to encoding in only one way. However, the individual's ability to process information is limited. Learners can only process a few pieces of information at a time in each of these channels (Paivio, 1971).

This leads to the third assumption of this theory, namely that human beings have a limited processing capacity. This capacity is often referred to as attentional resources, which is a set of limited mental resources used for processing information (Cherry, 1953). Theoretically, attentional resources are limited, which implies that it is impossible to process all the available information at the same time (Broadbent, 1958, Kahneman, 1973). The cognitive load theory (CLT) is a theory of human learning and memory that proposes that cognitive capacity is limited, and that learning can be facilitated if instructional materials are designed to reduce the cognitive load (Sweller, 1988). Within this context, the term of cognitive load is evoked and defined as the total quantity of resources dedicated to the processing of learning material. Therefore, this would result in three types of cognitive load: intrinsic cognitive load, extrinsic load, and germane load.

Intrinsic load represents the resources allocated to manage the complexity of the material to be learned. It is determined by the number of concepts to be handled simultaneously, their interactions, and their complexity. The level of expertise and the strategy used by the learner will lighten the intrinsic load of a device. *Extrinsic load* corresponds to those resources attributed to the processing of information not relevant to learning. It includes all the useless items added to the task such as distractions, overly complex explanations, or too high a level of detail. These elements, when considered, will use resources, thereby impacting the resources available for intrinsic load. The remainder of the resources available are directly useful for learning and are called *germane load*. This load is dedicated to the integration of acquired information in long-term memory. This processing corresponds to the construction of verbal and pictorial models of information, to their confrontation with knowledge in long-term memory, and finally to their storage in long-term memory. The resources allocated to this acquisition are assumed to be the subtraction of the total resources minus the resources used in intrinsic and extrinsic loads.

The relationship between the three types of cognitive load is thought to be additive: a device designed with overly high intrinsic and extrinsic loads may result in less available germane load resources. The interest for an instructional designer is therefore to create a device that reduces the extrinsic load of the device to make learning more effective. The goal of an instructional designer is to try to reduce the extrinsic load of the task and of the learning context so that the learner allocates resources solely to the task. With this objective in mind, various researchers have proposed different principles of instructional design based on empirical research. The idea behind these principles is to reduce the extrinsic and intrinsic cognitive load (for a review of these principles, see Mayer, 2014). induced by elements not necessary for the task, in other words, to avoid distribution of attentional resources on salient, but irrelevant elements. The presence of decorative images and attractive details are elements that can increase the extrinsic cognitive load of the task. Eliminating these elements in the learning task corresponds to the principle of coherence as fewer elements as possible. To avoid this dispersion of attentional resources toward irrelevant elements, it is also judicious to emphasize those elements that are useful for the pedagogical goal, such as the use of bold text and images. This corresponds to the principle of signaling, which makes it possible to highlight the most relevant elements in the task, and thus, to capture learners' attention in the information selection process (Mayer & Fiorella, 2014).

However, even with these principles, according to Miller (1996), learners will not always allocate correctly the necessary resources to complete the task but rather allocate their resources to obtain the level of understanding they desire. This desired level of understanding may therefore diverge from the pedagogical objective required and may naturally vary among learners. The challenge is therefore to be able to engage learners so that they mobilize the resources necessary to understand the content. A learner may have resources available in the germane load but may not activate them due to a lack of engagement in the task. Thus, the tools are proposed with the aim of ensuring learners make the effort to allocate more resources to the task by increasing their engagement. Yet, the notion of engagement is not addressed in this cognitive theory of multimedia learning. Furthermore, the emotional dimension of learning defined as an interaction between the emotional states of the learner and the content being learned (Wang & Walberg, 2004) is not addressed either. Consequently, different theoretical frameworks of multimedia learning have emerged in order to integrate these dimensions. One line of thought on the subject is the cognitive and affective theory of multimedia learning introduced by Moreno (2008), which is an extension of Mayer's theory by adding the affective and emotional dimensions in multimedia learning.

2.2. Cognitive and affective theory of multimedia learning

The Cognitive and affective theory of multimedia learning (Moreno, 2006) focuses on emotions as a trigger for the effort in germane load in the framework of multimedia learning. One fact revealed by Moreno (2006) is that the cognitive theory of multimedia learning proposed by Mayer (2005) does not consider affective factors such as the emotions felt by the individual during learning, or conative factors, i.e., learners’ motivation in the task. To this end, Moreno (2008) proposed the cognitive and affective theory of learning with media (CATLM). This theory is mainly based on Mayer's theory (2005), but it adds the affective dimension of multimedia learning and the motivational aspect. For Moreno (2007, p. 310), the “motivational factors mediate learning by increasing or decreasing cognitive engagement”, whereas “affective factors mediate learning by increased emotional engagement or reducing anxiety towards the learning task”. Affective engagement is increased when learners feel positive emotions toward the task, and when they feel supported and valued by their peers and educators (Pekrun & Linnebrink-Garcia, 2012). The interest in mobilizing learners’ affective and motivational resources is to promote the effective deployment of resources in the task. The main idea behind this theory is that a pleasant, attractive stimulus will impact learners' emotions and, indirectly, their learning, by increasing their cognitive engagement. Yet, Moreno's model is not complete since it does not explain how an emotional stimulus impacts engagement or what exactly cognitive engagement is. We therefore propose to extend this theory for a deeper understanding of the concept of engagement.

3. Learners’ engagement

As mentioned earlier, one of the problems with multimedia learning is the high dropout rate (Jordan, 2014; Patterson & McFadden, 2009; Rice, 2006; Roblyer, 2006), which is reported to be 10 to 20% higher in an online course than in a traditional classroom. The issue of learners’ engagement with a multimedia device is therefore key in e-learning design. Engagement is the "*holy grail of learning*" (Sinatra, Heddy & Lombardi, 2015) and is thought to characterize learners' active investment in a learning situation (Christenson, Reschly & Wylie, 2012). Many links have been established between engagement and academic success (Frederick, Blumenfeld, & Paris 2004, Heddy & Sinatra, 2013; Tytler & Osborne, 2012). Learners' engagement can be related to persistence and use of effective learning strategies (Krapp, 2000), as well as the ability to regulate themselves, and to make efforts to learn (Lee, Lee & bong, 2014).

Although engagement is a necessary condition for effective learning, there is no general definition but rather a set of operationalizations according to the associated field of study: student engagement, situational engagement, academic engagement, school engagement, and learner engagement (Reschly & Christenson, 2012). Sinatra, Heddy, and Lombardi (2015) uncovered two types of engagement corresponding to two levels of granularity. *Macro-level engagement* refers to learners' engagement over a long period of time within the larger school context (teacher, institution, peers) and with classroom activities. *Micro-level engagement*, also called situational engagement, is described by Lawson (2017) as a state of experience. It focuses mainly on learners' engagement in a task for a limited period.

3.1. Situational engagement

Situational engagement refers more to learners' experience at a specific time (minutes to hours) during the activity than to a process that varies over a long time (Lawson, 2017). Situational engagement is an optimal learning experience considered to be a "heightened, simultaneous experience of concentration, interest, and enjoyment in the task" (Shernoff, 2013, p. 12). It corresponds to a high level of challenge, interest, and competence.

This type of engagement is mainly derived from positive psychology theories and in particular, the flow theory (Csikszentmihalyi, 1990). A flow experience is characterized by a high level of cognitive and affective engagement (Csikszentmihalyi, 1990) during which "*nothing else seems to matter"* (Csikszentmihalyi, 1990, p. 4). The experience is described as autotelic in the sense that the individual performs the activity for the pleasure of doing it. When learners are in a flow state, they have a high level of concentration, lose the notion of time and self-awareness, and have a feeling of control over the activity and over themselves. For an activity to provide such a state of immersion, it must be sufficiently difficult to stimulate, but not so difficult as to discourage the individual. The objectives of the activity must be understandable and clear, and learners should get some feedback from the activity on their successes and failures, thus allowing them to adjust their behavior (Voelkl & Ellis, 2002).

3.2. Student engagement and its components

This second type of engagement corresponds to the macro-level and has been studied primarily in educational literature from kindergarten to higher education. Engagement is viewed as a multidimensional process that varies over time and is sometimes defined as a mediator between the learner's context (family, peers, schools) and learning outcomes (Rechly & Christenson, 2012). The main approach to engagement (Frederick et al., 2004) considers it to be a malleable and multidimensional process, in which three complementary dimensions interact: behavioral, affective, and cognitive.

Behavioral engagement refers to all observable indicators of student behavior in the classroom and to the social dynamics of an individual within an activity. It has been shown that the more effort and persistence students put into tasks, the more likely they are to learn (Pintrich & Schunk, 1996).

Affective engagement is the set of positive and negative emotions felt by the learner toward teachers and classmates. It also refers to the learners' sense of belonging to the field of study and the institution (Frederick et al., 2004). It has been shown that the more positive the emotions that learners feel toward the environment (institution, social environment, and learning task), the more likely they will return to the task and actively participate (Pekrun & Linnebrink-Garcia, 2012; Heddy & Sinatra, 2013)

The third component is cognitive engagement, corresponding to the set of cognitive resources that the learner will deploy when performing a task. It can be considered from a quantitative (quantity of resources allocated to the task, intensity of effort) and qualitative (degree of sophistication of learning strategies, adequacy of effort) perspectives. A high level of cognitive engagement is associated with better learning performance and higher academic achievement (Pintrich & Schrauben, 1992; Weinstein & Mayer, 1986).

3.3. Engagement in multimedia learning

These two types of engagement (student engagement and situational engagement) are particularly interesting and raise two issues. The first relates to the design of a multimedia device to ensure that learners returning to the experience do not give up over time. The second issue relates to how learners can be engaged to promote the deployment of cognitive resources for proper understanding of the content in real time, since learners generally do not provide all the resources necessary to complete the task in order to learn (Miller, 2015).

Different models focus on the factors impacting learner engagement with the aim of fostering learning. Bouvier, Lavoué and Sehaba (2014) proposed a model for engaged behaviors in the context of gaming, in which learners' individual characteristics and expectations toward the task play a significant role in their engagement. Learners may get involved if their expectations of the task are met. This issue of expectations can be linked to learners' initial motivation and interest in the content itself and how the learning contents are perceived. Finally, the media-specific factors, such as immersion and involvement, can be considered. If all these factors converge, learners will engage in the task.

Carroll et al. (2021) proposed an integrative model grouping the different approaches to engagement and added a third category of factors: the environment. This model is essentially based on the continuum introduced by Sinatra et al. (2015) between the micro and macro levels of engagement, while adding the flow state as a third type of engagement (not described here as situational engagement, but as an additional level of engagement). According to this model, these three levels are in continuous interaction: learners in a flow state have the potential to establish a micro-level state of engagement, and the same is true for the micro-level and macro-level states of engagement. In this hierarchical model, engagement is influenced mainly by individual factors such as motivation, cognitive abilities, interest, personality, or anxiety, and task factors such as the presence of challenge, clear goals, and meaningful content. Finally, environmental factors are considered such as autonomy, safety, and support. Autonomy is mainly based on the possibility for learners to make choices in the completion of the task. When an individual is allowed to make choices, this leads to self-determination and encourages participation and cognitive engagement (Deci, Vallerand, Pelletier, & Ryan, 1991; Rotgans & Schmidt, 2011). Psychological safety occurs when an individual feels secure and able to take risks without fear of negative consequences. This type of environment allows people to engage fully in tasks and activities, and leads to better performance (Kanfer & Ackerman, 1989).

To conclude this section, we note that the three categories of factors listed above impact the processes of learner engagement, while one of the major challenges in multimedia learning is to create an engaging device that captures learners' attention. In the next section, we therefore focus on deployment of learners' attentional resources in the task. In fact, the way learners select the different stimuli is not clearly described in Mayer's theory (2005). We integrate attentional theories to identify the elements that tend to be selected by learners during the learning processes, and position attention in relation to engagement.

4. Attention in multimedia learning

4.1. Deployment of attentional resources

The presence of multiple distracting information, unexpected interrupting tasks, and app pop-up windows may compromise the learning process (Brazzolotto & Michael, 2022). The issue is therefore to capture learners' attention in order to engage them sufficiently to resist distraction. Indeed, diverted attention is naturally mentioned by teachers as one of the major causes of difficulties in class (Cicekci & Sadik, 2019). The additional difficulty in asynchronous multimedia learning is that teachers are not able to detect loss of attention and cannot guide learners to focus on the material to be learned. There are also individual differences in the way attention is controlled (Ruff & Rothbart, 2001), thus defining various degrees of abilities in learning. For example, the ability to maintain focus on content (i.e., sustained attention) has been shown to vary with motivation and anxiety (Unsworth & Miller, 2021) and decline with aging (Davies & Parasuraman, 1982). The same applies to selective attention and the ability to inhibit irrelevant and distracting information (Craik & McDowd, 1987). One study found that individuals who were better at ignoring irrelevant information were better able to learn from multimedia presentations that contained both relevant and irrelevant information (Mayer, 2005). These findings suggest that individuals with better attentional abilities may be able to learn better from multimedia presentations. Learners' ability to inhibit irrelevant information and distractors seems to play a major role due to the inherent characteristics of multimedia learning.

During multimedia learning, learners deploy attention on the display that contains videos, text, and sounds. This may take place either involuntarily (i.e., captured by the elements present on the screen) or voluntarily while searching for specific information (Posner, 1980). Voluntary orientation of attention is primarily guided by learners' goals, previous knowledge, and experiences (Michael et al., 2006). In the multimedia context, this may correspond to a given pedagogical goal, to the interest and knowledge that the learner may have in the content, and to the elements that constitute the multimedia support. It is interesting to note that the educational goal envisaged by the designer can differ from the learner's objective. Therefore, deployment of learners' attention may not comply with what was envisaged by the designer, with the result that learning may be less efficient.

4.2. Salience and multimedia learning

The notion of involuntary deployment refers to non-conscious processes that guide attention toward a salient item of multimedia support (Michael et al., 2006). Many studies have shown that stimuli that have attention-capturing power are those that differ from the rest on a perceptual dimension (Treisman & Gormican, 1988; Duncan & Humphreys, 1988), therefore defining a local contrast of either color, orientation, or motion, and those stimuli appearing suddenly or undergoing changes in luminance, as well as infrequent or novel sounds (e.g., Turatto, Galfano, Gardini, & Mascetti, 2004; Oonk & Abrams, 1998; Abrams & Christ, 2003). Such characteristics facilitate the search for relevant items on any multimedia support (Jamet et al., 2008). In multimedia learning, this is referred to as the signaling principle, which, in terms of cognitive processing, involves attention cueing that can use selection, organization, and integration cues. Selection cues guide learners' attention to relevant information in the support (Crooks et al., 2012). Organization cues help learners to organize the elements of the support in a way that facilitates material processing and improves retention (Crooks et al., 2012). Integration cues help learners to integrate the elements of the support into a coherent whole (de Koning et al., 2009). In terms of perceptual processing, unique colors (Anderson & Schooler, 1991) or moving objects (Duncan, 1984) seem to be effective in capturing learners' attention. Two features that influence the perceptibility of visual representations include visual contrast and dynamic contrast, which seem to direct learners' attention to the target (de Koning et al., 2009). Auditory salience also plays a role and can be defined as the degree to which an auditory stimulus is noticed by an individual. It is determined by many factors including the loudness, timbre, pitch, location, and duration of the sound (Kaya & Elhilali, 2017).

However, a salient element is not always a relevant one (Michael et al., 2006), while the multitude of signals present may divert attention away from relevant aspects of the support. This may lead to difficulties in selecting the relevant information, thereby constituting an obstacle to learning the content (Mayer, 2005). The objective of a multimedia designer is therefore to highlight, i.e., confer salience to the relevant elements of the content in connection with the educational goals laid down to decrease the distracting role of salient but irrelevant elements, since integrating salient and relevant aspects of stimuli may increase performance (Michael et al., 2006). Distractors in multimedia learning are a more complex problem to tackle than in classical learning. In a face-to-face classroom context, the teacher can frame and dismiss potentially distracting elements (Harris, 1998). However, multimedia learning differs from face-to-face classroom learning in that learners' context cannot usually be controlled by the teacher. The effect of the number of distractors can vary greatly from one learner to another depending on the framework they have set themselves.

If more than one relevant item is highlighted, then attention has to be split. However, in such cases, performance drops (Tarmizi & Sweller, 1988). Multiple sources of information should be presented in such a way that learners do not need to split their attention between them (Mayer, 2009). There has been a great deal of research on split attention. For example, Moreno and Mayer (1998) found that learners presented with both auditory and visual information recalled less information than those who were only presented with visual information. The split attention effect is more pronounced when learners are asked to perform a task that is difficult or unfamiliar (e.g., solving a math problem).

4.3. Does attention reflect engagement?

In existing models of engagement, attention is closely linked to the behavioral, cognitive, and affective dimensions of engagement. Carroll et al. (2021) considered cognitive engagement to be learners' concentration in the task, absorption and loss of self-awareness. The notion of absorption is particularly interesting as it represents a heightened level of attention and immersion in the learning activity, strongly resembling the state of flow (Csikszentmihalyi, 1990) which is characterized by complete absorption in an activity, to the point where one loses track of time and self-consciousness. Empirical studies have shown that flow experiences are associated with high levels of engagement, motivation, and enjoyment (Shernoff et al., 2003; Csikszentmihalyi et al., 2005). Thus, attentional absorption can be seen as a key indicator of cognitive engagement in multimedia learning.

Moreover, attention is also tied to the behavioral dimension of engagement. Observable measures such as reading time, eye movements, and fixation duration can provide valuable insights into learners' attentional deployment and, by extension, their level of engagement (Miller, 2015). For instance, longer fixation times on specific elements of a multimedia lesson may signify greater interest and engagement with the content. Eye-tracking studies have shown that engaged learners tend to exhibit more focused gaze patterns, spending more time on relevant information and less on peripheral or irrelevant details (Mayer, 2010; van Gog & Scheiter, 2010). However, it is important to note that these behavioral measures are indirect and should be interpreted with caution, as they can be influenced by factors beyond engagement, such as prior knowledge or perceptual salience.

The relationship between attention and engagement may also vary depending on the characteristics of the multimedia learning task. Interactive simulations and educational games, for example, may be more conducive to attentional absorption and flow than passive video lectures, due to their immersive and participatory nature (Annetta, 2010; D'Mello et al., 2012). The level of challenge, feedback, and control offered by the task can also impact engagement. Tasks that strike a balance between challenge and skill, provide clear goals and immediate feedback, and allow for a sense of control are more likely to promote flow and engagement (Csikszentmihalyi, 1990; Shernoff et al., 2003).

Individual differences in attentional abilities should also be considered when examining the link between attention and engagement in multimedia learning. Learners vary in their capacity to sustain attention, resist distraction, and allocate attentional resources efficiently (Unsworth & McMillan, 2014; Unsworth et al., 2015). These variations may impact their ability to fully engage with multimedia content, particularly in the presence of competing stimuli or high cognitive load. Designing multimedia environments that accommodate different attentional profiles, such as by providing options for pacing, segmentation, or content selection, may help to optimize engagement for a wider range of learners (Mayer & Pilegard, 2014).

Recognizing attention as reflecting engagement has important implications for the cognitive theory of multimedia learning (Mayer, 2014). By integrating attentional processes into the model, we can better understand how learners allocate cognitive resources during multimedia learning and how this allocation is influenced by factors such as task design, individual differences, and the learning environment. This expanded framework can guide instructional designers in creating multimedia experiences that not only manage cognitive load but also actively promote learner engagement through attentional absorption and flow. By optimizing attention and engagement, we may ultimately enhance learning outcomes and foster more meaningful, enjoyable, and effective multimedia learning experiences.

In light of these considerations, we propose an integrative model that explicitly links attention and engagement in the context of multimedia learning (Figure 1).

5. An integrative model of engagement and attention in multimedia learning

We propose to integrate attention and engagement in the cognitive theory of multimedia learning by distinguishing learners' antecedents, deployment of attentional resources resulting from this engagement, and its outcomes (see Fig.1).

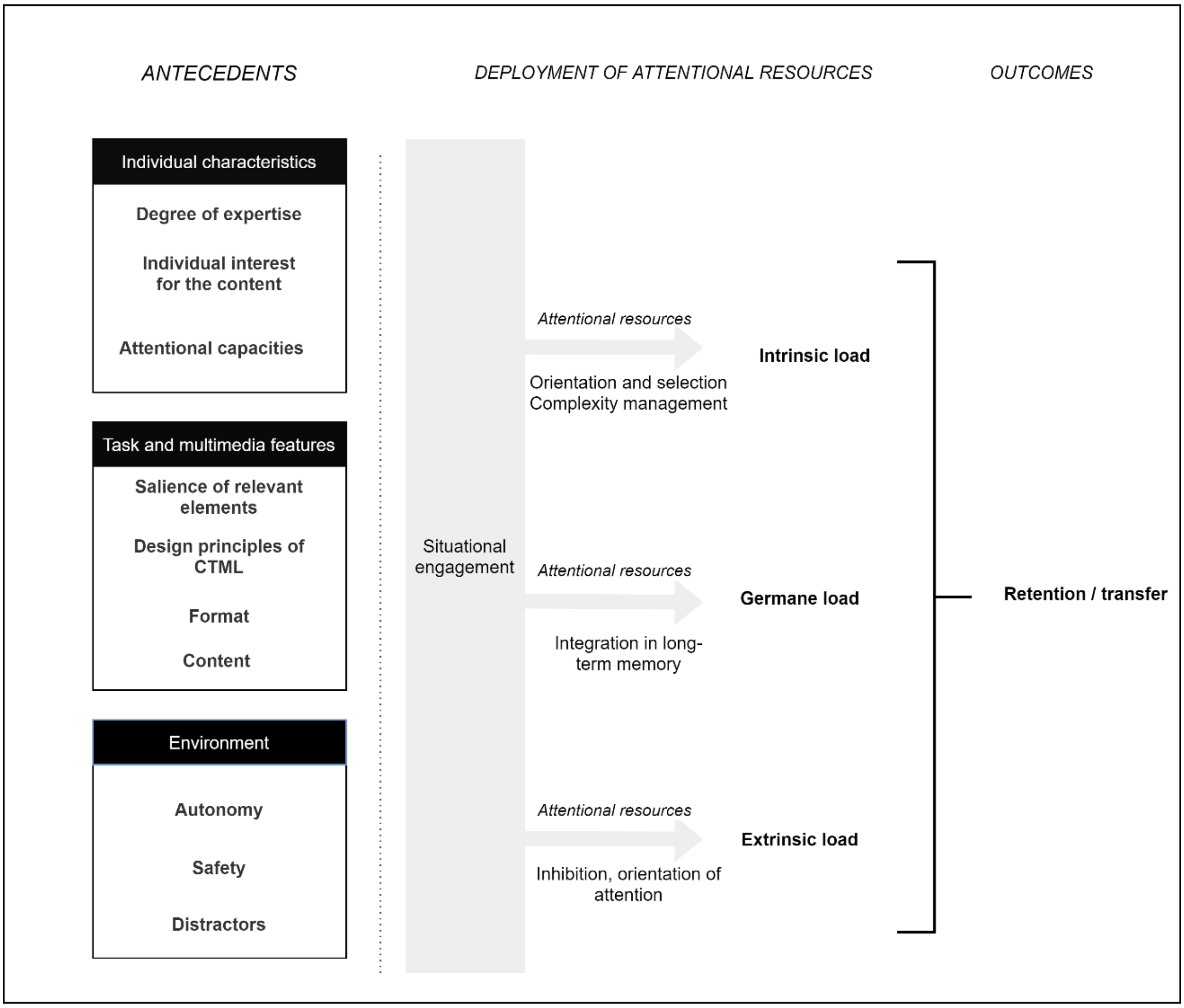

Fig1. Integrative model of engagement and attention in multimedia learning

5.1. Antecedents of learners' situational engagement

As proposed by Caroll et al. (2021), we distinguish individual characteristics, task and multimedia features, as well as the environment, as factors of learners' situational engagement. We highlight specific characteristics in multimedia learning, in relation to the deployment of attentional resources.

5.1.1. Individual characteristics

Degree of expertise. Many inter-individual factors influencing learning have been suggested. The main factor is learners' prior knowledge in the domain (Kalyuga, 2013). The degree of expertise of learners on the content directly influences the cognitive load, as mentioned in the cognitive theory of multimedia learning. Indeed, the less familiar learners are with content (vocabulary, etc.), the more the design needs to be rich to guide them in their learning. On the contrary, experts focus more on relevant information, which accounts for more effective learning compared to novices (Canham & Hegarty, 2010; Jarodzka et al., 2010).

Individual interest in the content. Learners have an individual interest in the content-to-learn, which, independently from the task itself, impacts their engagement (Renninger & Hidi, 2019). Any content that is personally relevant to the learner is more likely to capture and maintain their interest than content that is not.

Attentional abilities. Our model emphasizes that learners possess varying attentional abilities, which significantly influence their understanding of content (Ruff & Rothbart, 2001). Attentional abilities, defined as cognitive skills that enable selective focus, sustained attention, resistance to distraction, and efficient mental resource allocation, do not directly impact situational engagement but affect the amount and deployment of attentional resources a learner can utilize (Cowan, 2017; Posner & Rothbart, 2007). These abilities differ among learners due to factors like age, cognitive development, and individual differences in executive functioning (Diamond, 2013; Rothbart & Posner, 2015).

For instance, consider two types of learners: those who are highly engaged in the content but have low attentional abilities, deploying all their resources during learning, and those who have high attentional abilities but are not strongly engaged. Which of these learners will be more successful? Can engagement compensate for low attentional abilities? This question underscores the need for future research to clarify the relationship between engagement and attention.

In multimedia learning, attentional abilities interact with task characteristics, determining the learner's experienced difficulty. Factors such as task complexity, interactivity, presentation format, and pacing can influence cognitive load and attentional demands (Mayer, 2014; Sweller et al., 2011). For example, a highly complex multimedia lesson with multiple information sources and rapid presentation may challenge learners with lower attentional capacities, as they may struggle to focus on relevant details and process information efficiently.

5.1.2. Task characteristics and multimedia support

We identify the following characteristics of tasks and multimedia support as playing a key role in learners' engagement.

Salience of relevant elements and design principle of the cognitive theory of multimedia learning. As shown above, these characteristics have an impact on learners' ability to engage in and focus on the task. There is a strong relationship between salience and engagement: when something is salient, it is more likely to capture attention and engage us (i.e., be important to us). For example, if we see a sign that is bright and flashing, it is more likely to catch our attention than a sign that is dull and motionless.

Format and content of the task. According to the flow theory (Nakamura & Csikszentmihalyi, 2014), the task should be clear, coherent, comprehensive, and consistent with the learning objectives. The level of difficulty of the task can also have an impact on learners' ability to engage in the task.

5.1.3. Environment

The characteristics of the environment presented in this model mainly rely on the those proposed by Caroll et al. (2021).

Autonomy. Autonomy refers to the possibility for learners to have choices in the task. It has been shown that this increases engagement in the task by giving learners control over the situation (Fredricks et al., 2004). In multimedia learning environments, autonomy can be achieved by providing learners with a sense of control and ownership over their learning experience. By feeling in control, learners feel more comfortable taking risks, exploring new ideas, and asking questions (e.g., to be engaged).

Safety and support. Safety and support, both referring to the learning environment, are crucial to create a psychologically safe environment for learning. One way is to provide learners with feedback on their progress that is positive and constructive. (Johnson & Priest, 2014).

Distractors. Distractors can be more present in multimedia learning due to the inherent nature of the support. Distractors can divert attention away from the important aspects of the support and make learning less effective (Wang & Adesope, 2014).

5.2. Link between individual characteristics and engagement

The relationship between task difficulty and engagement is a key aspect of our model, drawing upon the flow theory of engagement (Csikszentmihalyi, 1990; Shernoff et al., 2003). According to this theory, optimal engagement occurs when there is a balance between the challenge posed by the task and the individual's skills. When a task is too easy, learners may become bored and disengaged; when it is too difficult, they may experience frustration and anxiety, leading to disengagement. In the context of multimedia learning, attentional abilities can be viewed as a critical skill set that determines whether a learner perceives a task as appropriately challenging or overwhelming.

Thus, our model proposes that the alignment between task difficulty and attentional abilities plays a significant role in shaping engagement outcomes. Learners with strong attentional skills may be better equipped to handle complex, fast-paced multimedia content, maintaining engagement even in the face of high cognitive load. In contrast, learners with less developed attentional abilities may struggle to remain engaged when the attentional demands of the task exceed their capacities. This interaction between attentional skills and task characteristics has important implications for multimedia design, suggesting that designers should consider the attentional profiles of their target learners and strive to create content that provides an optimal level of challenge.

Beyond task difficulty, our model also acknowledges the role of other learner characteristics in shaping engagement, such as prior knowledge, motivation, and interest (Tobias, 1994; Renninger & Hidi, 2016). These factors can influence the allocation of attentional resources and the depth of cognitive processing during multimedia learning. For instance, learners with high intrinsic motivation and interest in the topic may be more likely to sustain attention and persist in the face of challenging content, even if their attentional abilities are not exceptionally high (Hidi & Renninger, 2006; Schiefele, 2009). Conversely, learners with low motivation may fail to engage deeply with the material, even if they possess strong attentional skills.

5.3. Engagement and deployment of attentional resources

Our model proposes that the level of engagement significantly influences the deployment of attentional resources during multimedia learning. According to Miller (2015), the initial orientation of attention towards a specific element of the learning material represents the first level of engagement, reflecting the allocation of cognitive resources. As learners become more engaged, driven by factors such as motivation, interest, and the learning context, they tend to deploy additional attentional resources to process the information more deeply. This heightened engagement is characterized by increased cognitive engagement, which refers to the quantity and quality of allocated attentional resources (see Section 2.2). The deployment of attentional resources is closely tied to the management of cognitive load during multimedia learning, as described by Cognitive Load Theory (Sweller, 1988). The theory distinguishes between three types of cognitive load: intrinsic, extrinsic, and germane load, each of which has a distinct relationship with engagement:

Intrinsic Load. Intrinsic load refers to the inherent complexity of the material being learned, which determines the amount of information that must be processed to complete the task (Sweller, 1988). Engaged learners are better able to focus their attention on relevant information, despite the complexity. However, as the difficulty of the task increases, more cognitive resources are required. High intrinsic load necessitates a greater degree of selective attention to manage the information effectively. When intrinsic load is too high, even highly engaged learners may struggle to process information efficiently, leading to potential declines in performance.

Extrinsic Load. Extrinsic load is generated by external distractions or irrelevant elements within the learning environment (Sweller, 1988). High engagement helps learners allocate attentional resources more effectively, allowing them to resist distractions and focus on the task. However, if the extrinsic load is excessive, even engaged learners might find their attentional resources diverted, leading to reduced task performance. Effective engagement minimizes the impact of extrinsic load by enhancing the learner's ability to concentrate on relevant information and inhibit distractions.

Germane Load. Germane load refers to the cognitive effort required to process, store, and integrate new information into long-term memory (Sweller, 1988). Engagement plays a crucial role in managing germane load by motivating learners to invest the necessary cognitive resources. When learners are highly engaged, they are more likely to allocate sufficient attentional resources to understand and retain new information. Conversely, if engagement is low, learners may lack the motivation to process information deeply, reducing the effectiveness of learning. Germane load is thus optimized when learners are both engaged and able to manage their cognitive resources effectively.

5.4. Outcomes

Transfer on multimedia learning. Transfer is a generalization from the test to the real world. When we learn something, we need to be able to apply it to new situations. There are several different ways to measure transfer on learning. One way is to simply ask the learner how they are doing with the new material. Another way is to provide the learner with a test on the new material and measure their performance. The most important factor in transfer is how well the learner understands the new material. If the learner does not understand the material, they will not be able to apply it to new situations, whereas if the learner understands the material, they will be able to apply it to new situations. The best way to ensure that learners understand the material is to provide them with practice opportunities. The more practice learners have with the material, the better they will be able to understand it and apply it to new situations.

Retention on multimedia learning. Retention is a measure of how well learners remember the material. Retention can be measured in several ways, for instance by providing the learner with a test on the material. The test can be carried out immediately after the learning session or after a certain set period of time. The test can be either objective (e.g., multiple choice) or subjective (e.g., essay)

These two outcomes are the classic measures of learning performance in multimedia learning.

6. Conclusion

In this paper, we have proposed a model that integrates situational engagement and attention in the cognitive theory of multimedia learning with the goal of better understanding the role of engagement on the deployment of attentional resources. For this purpose, we have pointed out the different limits of the theoretical framework of multimedia learning developed by Mayer (2005) and, particularly, the lack of precise relationships between engagement and attention.

We have shown that attention and engagement are linked by the fact that attention reflects learners' engagement, at least partly. We have highlighted the importance of learners' attentional abilities in the effectiveness of learning and proposed to consider these abilities more systematically in existing models. In our integrative model, we have also added other factors that can interfere with engagement and subsequent deployment of attentional resources. With this in mind, we have identified three categories of factors that influence engagement (the environment, individual characteristics, and multimedia characteristics). In the second part, we have clarified the link between deployment of attentional resources and cognitive load. Finally, we have presented the impact on learning outcomes.

This review opens up new questions. Do attentional abilities play a role in the effectiveness of the principles of pedagogical design described in the multimedia learning theory? What is the role of engagement in these principles? Can an engaged learner with low attentional abilities learn better than an unengaged learner with high attentional abilities? One way to address these questions would be to assess participants attention abilities before presenting them with material to learn and manipulate the characteristics of the support in a way to influence the degree of engagement. This model also has potential implications for educational practices. For example, it could help instructional designers create more effective multimedia content by considering the role of attention and engagement. Additionally, this model could help educators identify learners who could benefit from additional support in attentional skills when working with multimedia content.

7. References

Abrams, R. A., & Christ, S. E. (2003). Motion Onset Captures Attention. *Psychological Science*, *14*(5), 427-432. https://doi.org/10.1111/1467-9280.01458

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the Environment in Memory. *Psychological Science*, *2*(6), 396-408. https://doi.org/10.1111/j.1467-9280.1991.tb00174.x

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*, 369-386. https://doi.org/10.1002/pits.20303

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *The psychology of learning and motivation: II* (p. xi, 249-xi, 249). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Beck, L. A. (1992). Csikszentmihalyi, Mihaly. (1990). Flow: The Psychology of Optimal Experience. *Journal of Leisure Research*, *24*(1), 93-94. https://doi.org/10.1080/00222216.1992.11969876

Blackorby, J., & Cameto, R. (s. d.). *8. Changes in the School Engagement and Academic Performance of Students with Disabilities*. 24.

Blumenfeld, P., & Paris, A. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research - REV EDUC RES*, *74*, 59-109. https://doi.org/10.3102/00346543074001059

Bouvier, P., Lavoué, E., & Sehaba, K. (2014). Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming. *Simulation and Gaming*, *45*(4-5), 491-507. https://doi.org/10.1177/1046878114553571

Brault-Labbé, A., & Dubé, L. (2008). Engagement, surengagement et sous-engagement académiques au collégial : Pour mieux comprendre le bien-être des étudiants. *Revue des sciences de l'éducation*, *34*(3), 729751. https://doi.org/10.7202/029516ar

Brazzolotto, P., Duran, G., & Michael, G. A. (2022). How do we handle interruptions? Investigating the processes underlying the resumption of interrupted tasks. *Psychologie Française*, *67*(1), 77-95.

Broadbent, D. E. (1958). *Perception and communication.* Pergamon Press. https://doi.org/10.1037/10037-000

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction*, *20*, 155-166. https://doi.org/10.1016/j.learninstruc.2009.02.014

Carroll, M., Lindsey, S., Chaparro, M., & Winslow, B. (2021). An applied model of learner engagement and strategies for increasing learner engagement in the modern educational environment. *Interactive Learning Environments*, *29*(5), 757-771. https://doi.org/10.1080/10494820.2019.1636083

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975-979. https://doi.org/10.1121/1.1907229

Christenson, S. L., Reschly, A. L., & Wylie, C. (Éds.). (2012). *Handbook of Research on Student Engagement*. Springer US. https://doi.org/10.1007/978-1-4614-2018-7

Cicekci, M., & Sadik, F. (2019). Teachers' and Students' Opinions About Students' Attention Problems During the Lesson. *Journal of Education and Learning*, *8*, 15. https://doi.org/10.5539/jel.v8n6p15

Clark, J. M., & Paivio, A. (1987). A Dual Coding Perspective on Encoding Processes. In M. A. McDaniel & M. Pressley (Éds.), *Imagery and Related Mnemonic Processes: Theories, Individual Differences, and Applications* (p. 5-33). Springer. https://doi.org/10.1007/978-1-4612-4676-3_1

Craik, F., & Mcdowd, J. (1987). Age Differences in Recall and Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 474-479. https://doi.org/10.1037/0278-7393.13.3.474

Crooks, S. M., Cheon, J., Inan, F., Ari, F., & Flores, R. (2012). Modality and cueing in multimedia learning: Examining cognitive and perceptual explanations for the modality effect. *Computers in Human Behavior*, *28*(3), 10631071. https://doi.org/10.1016/j.chb.2012.01.010

Csikszentmihalyi, M. (2000). *FLOW: The Psychology of Optimal Experience*. 6.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. Academic Press.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a Framework for Attention Cueing in Instructional Animations: Guidelines for Research and Design. *Educational Psychology Review*, *21*(2), 113-140. https://doi.org/10.1007/s10648-009-9098-7

Dickinson, L. (1995). Autonomy and motivation a literature review. *System*, *23*(2), 165-174. https://doi.org/10.1016/0346-251X(95)00005-5

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology. General*, *113*(4), 501-517. https://doi.org/10.1037//0096-3445.113.4.501

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433-458. https://doi.org/10.1037/0033-295X.96.3.433

Folker, S., Folker@uni-Bielefeld, S., De, H., Ritter, & Sichelschmidt, L. (2005). Processing and integrating multimodal material: The influence of color-coding. *Proceedings of the 27th annual conference of the cognitive science society*.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, *74*(1), 59-109. https://doi.org/10.3102/00346543074001059

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Friedman, N. P., & Miyake, A. (2004). The Relations Among Inhibition and Interference Control Functions: A Latent-Variable Analysis. *Journal of Experimental Psychology: General*, *133*, 101-135. https://doi.org/10.1037/0096-3445.133.1.101

Guilford, J. P. (1927). « Fluctuations of Attention » with Weak Visual Stimuli. *The American Journal of Psychology*, *38*(4), 534-583. https://doi.org/10.2307/1414394

Halverson, L. R., & Graham, C. R. (2019). Learner Engagement in Blended Learning Environments: A Conceptual Framework. *Online Learning*, *23*(2), Art. 2. https://doi.org/10.24059/olj.v23i2.1481

Harms, C. M., Niederhauser, D. S., Davis, N. E., Roblyer, M. D., & Gilbert, S. B. (s. d.). *Journal of Communication/La Revue Electronique de Communication, 16(1& 2), 2006*. 20.

Harp, S., & Mayer, R. (1998). How Seductive Details Do Their Damage: A Theory of Cognitive Interest in Science Learning. *Journal of Educational Psychology*, *90*, 414-434. https://doi.org/10.1037/0022-0663.90.3.414

Harris, A. (1998). Effective Teaching: A review of the literature. *School Leadership & Management*, *18*(2), 169-183. https://doi.org/10.1080/13632439869628

Heddy, B. C., & Sinatra, G. M. (2013). Transforming Misconceptions: Using Transformative Experience to Promote Positive Affect and Conceptual Change in Students Learning About Biological Evolution: TRANSFORMING MISCONCEPTIONS. *Science Education*, *97*(5), 723-744. https://doi.org/10.1002/sce.21072

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*(2), 135-145. https://doi.org/10.1016/j.learninstruc.2007.01.011

Janosz, M. (2012). Part IV Commentary: Outcomes of Engagement and Engagement as an Outcome: Some Consensus, Divergences, and Unanswered Questions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Éds.), *Handbook of Research on Student Engagement* (p. 695-703). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_33

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, *20*(2), 146-154. https://doi.org/10.1016/j.learninstruc.2009.02.019

Johnson, C. I., & Priest, H. A. (2014). The Feedback Principle in Multimedia Learning. In R. E. Mayer (Éd.), *The Cambridge Handbook of Multimedia Learning* (2e éd., p. 449-463). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.023

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, *15*(1). https://doi.org/10.19173/irrodl.v15i1.1651

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329-354. https://doi.org/10.1037/0033-295X.87.4.329

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kalantzis, M., & Cope, W. (2020). Afwheter the COVID-19 crisis: Why higher education may (and perhaps should) never be the same. *ACCESS: Contemporary Issues in Education*, *40*. https://doi.org/10.46786/ac20.9496

Kalyuga, S. (2013). Effects of Learner Prior Knowledge and Working Memory Limitations on Multimedia

Learning. *Procedia - Social and Behavioral Sciences*, *83*, 25-29. https://doi.org/10.1016/j.sbspro.2013.06.005

Kaya, E., & Elhilali, M. (2017). Modelling auditory attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*, 20160101. https://doi.org/10.1098/rstb.2016.0101

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge; New York, NY, USA: Cambridge University Press. http://archive.org/details/comprehensionpar0000kint

Krapp, A., & Prenzel, M. (2011). Research on Interest in Science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27-50. https://doi.org/10.1080/09500693.2010.518645

Lawson, M. A. (2017). Commentary: Bridging student engagement research and practice. *School Psychology International*, *38*(3), 221-239. https://doi.org/10.1177/0143034317708010

Lee, W., Lee, M.-J., & Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology*, *39*, 86-99. https://doi.org/10.1016/j.cedpsych.2014.02.002

Li, J., Antonenko, P. D., & Wang, J. (2019). Trends and issues in multimedia learning research in 1996–2016: A bibliometric analysis. *Educational Research Review*, *28*, 100282. https://doi.org/10.1016/j.edurev.2019.100282

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*(1), 6-21. https://doi.org/10.1080/17470214808416738

Mayer, R. E. (2003). Elements of a Science of E-Learning. *Journal of Educational Computing Research*, *29*(3), 297-313. https://doi.org/10.2190/YJLG-09F9-XKAX-753D

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In *The Cambridge handbook of multimedia learning* (p. 31-48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E. (2009). *Multimedia Learning* (2e éd.). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2010). Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning. In R. Moreno (Éd.), *Cognitive load theory*. (p. 131-152). Cambridge University Press. https://doi.org/10.1017/CBO9780511844744.009

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning, 2nd ed.* (p. 43-71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E. (2014b). *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of educational psychology*, *90*(2), 312.

Mayer, R. E., & Moreno, R. (2001). Animation as an Aid to Multimedia Learning. *Educational Psychology Review*, 13.

Mayer, R. E., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 43-52. https://doi.org/10.1207/S15326985EP3801_6

Mayer, R., & Fiorella, L. (2014). *Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles* (p. 279-315). https://doi.org/10.1017/CBO9781139547369.015

Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology*, *80*, 514523. https://doi.org/10.1037/0022-0663.80.4.514

Michael, G. A., Garcia, S., Fernandez, D., Sellal, F., & Boucart, M. (2006). The ventral premotor cortex (vPM) and resistance to interference. *Behavioral Neuroscience*, *120*(2), 447462. https://doi.org/10.1037/0735-7044.120.2.447

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97. https://doi.org/10.1037/h0043158

Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in Academic Work: The Role of Learning Goals, Future Consequences, Pleasing Others, and Perceived Ability. *Contemporary Educational Psychology*, *21*(4), 388-422. https://doi.org/10.1006/ceps.1996.0028

Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, *22*(3), 149-158. https://doi.org/10.1111/j.1365-2729.2006.00170.x

Moreno, R., & Mayer, R. (2000). *A learner-centered approach to multimedia explanations: Deriving instructional design principles from cognitive theory*. *2*.

Nakamura, J., & Csikszentmihalyi, M. (2014). The Concept of Flow. In M. Csikszentmihalyi (Ed.), *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (pp. 239–263). Springer Netherlands. https://doi.org/10.1007/978-94-017-9088-8_16

Oonk, H. M., & Abrams, R. A. (1998). New perceptual objects that capture attention produce inhibition of return. *Psychonomic Bulletin & Review*, *5*, 510-515. https://doi.org/10.3758/BF03208830

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1993). The Efficiency of Instructional Conditions: An Approach to Combine Mental Effort and Performance Measures. *Human Factors*, *35*(4), 737-743. https://doi.org/10.1177/001872089303500412

Paivio, A. (1971). *Imagery and verbal processes.* (p. xi, 596). Holt, Rinehart & Winston.

Patterson, B., & McFadden, C. (2009). Attrition in Online and Campus Degree Programs. *Online Journal of Distance Learning Administration*, *12*(2).

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In *Handbook of research on student engagement* (p. 259-282). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-2018-7_12

Peters, M. A., Rizvi, F., McCulloch, G., Gibbs, P., Gorur, R., Hong, M., Hwang, Y., Zipin, L., Brennan, M., Robertson, S., Quay, J., Malbon, J., Taglietti, D., Barnett, R., Chengbing, W., McLaren, P., Apple, R., Papastephanou, M., Burbules, N., Misiaszek, L. (2022). Reimagining the new pedagogical possibilities for universities post-Covid-19: An EPAT Collective Project. *Educational Philosophy and Theory*, *54*(6), 717-760. https://doi.org/10.1080/00131857.2020.1777655

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In *Student perceptions in the classroom* (p. 149-183). Lawrence Erlbaum Associates, Inc.

Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in Education: Theory, Research, and Applications*. Merrill.

Posner, M. (1980). Orienting of Attention. *The Quarterly journal of experimental psychology*, *32*, 3-25. https://doi.org/10.1080/00335558008248231

Reeve, J. (2012). A Self-determination Theory Perspective on Student Engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Éds.), *Handbook of Research on Student Engagement* (p. 149172). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_7

Renninger, K. A., & Hidi, S. E. (2019). Interest Development and Learning. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge Handbook of Motivation and Learning* (pp. 265–290). Cambridge University Press. https://doi.org/10.1017/9781316823279.013

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In *Handbook of research on student engagement.* (p. 3-19). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-2018-7_1

Rice, K. L. (2006). A Comprehensive Look at Distance Education in the K–12 Context. *Journal of Research on Technology in Education*, *38*(4), 425-448. https://doi.org/10.1080/15391523.2006.10782468

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). `Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747758. https://doi.org/10.1016/S0028-3932(97)00015-8

Ruff, H. A., & Rothbart, M. K. (2001). Individual Differences in Attention. In H. A. Ruff & M. K. Rothbart (Eds.), *Attention in Early Development: Themes and Variations* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195136326.003.0010

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008a). *Motivation in education: Theory, research, and applications* (3rd ed). Pearson/Merrill Prentice Hall.

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008b). *Motivation in education: Theory, research, and applications* (3rd ed). Pearson/Merrill Prentice Hall.

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement.* (p. xvii, 368). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-7089-2

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist*, *50*(1), 1-13. https://doi.org/10.1080/00461520.2014.1002924

Sousa, D. A., & Tomlinson, C. A. (2011). *Differentiation and the Brain: How Neuroscience Supports the Learner-Friendly Classroom*. Solution Tree Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285. https://doi.org/10.1016/0364-0213(88)90023-7

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, *80*, 424-436. https://doi.org/10.1037/0022-0663.80.4.424

Treisman, A., & Gormican, S. (s. d.). *Feature Analysis in Early Vision: Evidence from Search Asymmetries*. 34.

Turatto, M., Galfano, G., Gardini, S., & Mascetti, G. (2004). Stimulus-Driven Attentional Capture: An Empirical Comparison of Display-Size and Distance Methods. *The Quarterly journal of experimental psychology. A Human experimental psychology*, *57*, 297-324. https://doi.org/10.1080/02724980343000242

Tytler, R., & Osborne, J. (2012). Student Attitudes and Aspirations Towards Science. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Éds.), *Second International Handbook of Science Education* (p. 597625). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9041-7_41

Unsworth, N., & Miller, A. L. (2021). Individual Differences in the Intensity and Consistency of Attention. *Current Directions in Psychological Science*, *30*(5), 391–400. https://doi.org/10.1177/09637214211030266

Voelkl, J., & Ellis, G. (2002). *Optimal experience in daily life: Examination of the predictors, dimensions and outcome*. Leisure Research Symposium, Ashburn, VA: National Recreation and Park Association.

Wang, B., & Theeuwes, J. (2018). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*, *80*(4), 860-870. https://doi.org/10.3758/s13414-018-1493-z

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies in M, wittrock (ED) handbook of research on teaching pp (315-327). *New Yourk, Macillan*.

Wiebe, E., & Sharek, D. (2016). ELearning. In H. O'Brien & P. Cairns (Éds.), *Why Engagement Matters: Cross-Disciplinary Perspectives of User Engagement in Digital Media* (p. 53-79). Springer International Publishing. https://doi.org/10.1007/978-3-319-27446-1_3

Wittrock, M. C. (1986). *Handbook of research on teaching* (3rd ed). Macmillan; Collier Macmillan.

Wittrock, M. C. (1989). Generative Processes of Comprehension. *Educational Psychologist*, *24*(4), 345-376.

https://doi.org/10.1207/s15326985ep2404_2

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building academic success on social and emotional learning: What does the research say?* (P. x, 244). Teachers College Press.

Transition et objectif de recherche

L'objectif de ce premier article théorique était de construire un modèle intégratif posant l'engagement comme facteur central modulant le déploiement des ressources attentionnelles durant l'apprentissage multimédia. Nous avons également considéré l'attention en tant que capacité cognitive, comme un prédicteur de la capacité d'un apprenant à s'engager pleinement dans une tâche d'apprentissage via la régulation de ses ressources attentionnelles. Cependant, pour soumettre ce modèle théorique à une validation empirique solide, il est essentiel d'élaborer un protocole de mesure rigoureux permettant d'évaluer avec une grande finesse les différentes dimensions de l'engagement situationnel et de la motivation, ainsi que leurs relations. En effet, l'engagement et la motivation sont des construits étroitement liés, mais distincts (Appleton et al., 2008 ; Reeve, 2012). Comprendre comment ils interagissent dans le contexte de l'apprentissage multimédia nécessite une approche méthodologique intégrée.

L'évaluation de l'engagement et de la motivation dans leur complexité représente un défi méthodologique de taille, requérant la collecte et l'analyse conjointe de données multimodales, à la fois subjectives et objectives (Bouvier et al., 2014). Les mesures subjectives, comme les questionnaires, permettent d'accéder aux perceptions et états psychologiques auto-rapportés par les apprenants. Parallèlement, l'analyse des traces d'interaction des apprenants avec l'environnement d'apprentissage offre des indices comportementaux objectifs de leur engagement (Henrie et al., 2015).

Afin de pouvoir tester empiriquement les prédictions de notre modèle intégratif, nous avons donc mis en place une première étude expérimentale visant, entre autres, à valider un protocole de mesure multimodale de l'engagement, de la motivation et des performances d'apprentissage. Ce protocole combinant diverses méthodes subjectives et objectives a été appliqué à l'évaluation d'un principe de la théorie de l'apprentissage multimédia : le principe d'anthropomorphisme. Bien que l'impact positif de l'anthropomorphisme (par exemple à l'aide d'agents pédagogiques animés) sur les performances d'apprentissage ait été démontré (Brom et al., 2018), peu d'études à ce jour se sont attachées à examiner son influence sur les états motivationnels et l'engagement réel des apprenants.

Notre expérience a donc fait varier le niveau d'anthropomorphisme d'un environnement d'apprentissage multimédia, tout en mesurant, grâce à notre protocole, l'engagement, la motivation, la charge cognitive et les performances d'apprentissage. Cela nous a permis de poursuivre un triple objectif : (i) Valider la sensibilité et la fiabilité de notre protocole de mesure de l'engagement et de la motivation pour détecter leurs différentes dimensions, (ii) tester les effets de l'anthropomorphisme sur l'engagement situationnel et la motivation des apprenants, au-delà des performances d'apprentissage et (iii) clarifier les relations entre l'engagement, la motivation et les performances d'apprentissage multimédia.

Chapitre 3 : Impact de l'Anthropomorphisme dans l’Apprentissage Multimédia

Enhancing engagement, motivation, and learning outcomes in e-learning: The role of anthropomorphism

Y. Minne[1,2,3], E. Lavoué[3] C. Thiery[2] and G. A. Michael[1]

[1]Unité de Recherche Étude des Mécanismes Cognitifs, Université de Lyon, Université Lyon 2, Lyon, France
[2] Sydo, 69001 Lyon, France
[3] Univ Lyon, Université Jean Moulin Lyon 3, iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, F-69008

Abstract

This study aimed at examining the influence of rich multimedia support, characterized by anthropomorphism, on learner engagement, motivation, cognitive load, and learning outcomes in an e-learning context. Two versions of a multimedia support were compared: rich (with anthropomorphism) and neutral (without anthropomorphism). Participants who used the rich version reported higher levels of engagement, more autonomous and self-determined motivation, and better learning outcomes than those exposed to the neutral version. The study also revealed that engagement mediated the impact of external motivation on learning outcomes. However, we did not find any mediation between the support version, motivation, engagement, and learning outcomes. No significant difference was found in cognitive load between the two groups, suggesting that the rich multimedia support used did not excessively increase cognitive load. The findings underscore the potential benefits of integrating rich multimedia elements in e-learning environments to enhance learner engagement, motivation, and learning outcomes. Future research is recommended to further explore the impact of anthropomorphism on different dimensions of learners' motivation and engagement in multimedia learning through more detailed scales.

Keywords: Engagement, motivation, distance education and online learning, pedagogical issues

Highlights

- The use of anthropomorphism increases learner engagement, identified motivation, and learning outcomes.
- The use of anthropomorphism decreases external motivation
- There is no significative difference in cognitive load between groups exposed to rich and neutral multimedia versions.

• Engagement plays a mediating role between external motivation and learning outcomes.

Introduction

In recent years, e-learning has gained increasing popularity as an educational tool due to its accessibility, flexibility, and cost-effectiveness (Dhika et al., 2019). However, the effectiveness of e-learning is highly dependent on the design and implementation of the learning material.

One approach to designing effective e-learning materials is the Cognitive Theory of Multimedia Learning (CTML), which suggests that learners can process and retain information more effectively when it is presented in a multimedia format that aligns with their cognitive architecture (Mayer, 2005). This theory emphasizes the importance of designing e-learning materials that integrate words and pictures, use relevant and meaningful content, and provide guidance and feedback to learners. Building on the CTML, the Cognitive and Affective Theory of Multimedia Learning (CATML) incorporates emotional and affective dimensions, positing that affective states also trigger a deeper cognitive processing (Moreno, 2006). Emotional design, based on the CATML, refers to the intentional use of design elements to elicit emotional responses in users (Um et al., 2012a), which can significantly impact learning outcomes (Plass et al., 2014; Plass and Kaplan, 2016). This framework is based on the principle that emotions can positively or negatively influence online learning (Um, Plass, Hayward, & Homer, 2012). The two most studied design principles in this framework are color (Bellizzi and Hite, 1992; Wolfson and Case, 2000; Gorn et al., 1997) and anthropomorphism, referring to the design technique of imbuing human-like attributes into learning elements (Liew et al., 2022).

Anthropomorphic attributes in multimedia learning materials are often cute and funny characters with personalized dialogs. They allow a direct description of the link between elements. Anthropomorphism has been shown to improve memorization (Brom et al., 2018; Plass et al., 2014; Schneider et al., 2016), comprehension (Plass et al., 2014; Kumar et al., 2016; Um et al., 2012b), and transfer (Park et al., 2015; Gong et al., 2017a; Um et al., 2012b) of the learned content. It was also suggested that anthropomorphic graphics can enhance learners' positive emotions (Mayer and Estrella, 2014; Gong et al., 2017b; Park et al., 2015). However, their effectiveness may depend on the type of learning material, the learners' prior knowledge, and the learners' cultural background (Stárková et al., 2019). While anthropomorphism has been shown to have a significant impact on e-learning outcomes, motivation, and emotions, it is important to note that, to the best of our knowledge, none of the studies in this area had focused on measuring engagement *(*Brom et al., 2018). Engagement can be defined as a positive psychological state characterized by emotional, cognitive, and behavioral investment in an activity (Fredricks et al., 2004). Well-designed multimedia materials can foster learner engagement, and higher engagement could result in better understanding, retention, and transfer of the learned material (Skuballa et al., 2018).

Some researchers, such as Shernoff (2013), have introduced the concept of situational engagement, which refers to the dynamic and context-dependent nature of engagement during a specific learning activity. This notion emphasizes the importance of considering the interaction between the learner, the learning environment, and the specific task at hand when examining engagement in multimedia learning. In the case of multimedia learning, situational engagement is particularly relevant, as it refers to the engagement that learners experience over a short time (Shernoff, 2013). Situational engagement is frequently analyzed in studies investigating multimedia learning (Sun and Rueda, 2012). However, it is worth noting that the concept of engagement is not widely incorporated within the theoretical frameworks of the Cognitive Theory of Multimedia Learning (CTML). CTML primarily focuses on cognitive processes and instructional design principles, whereas engagement is often regarded as a complementary construct used in empirical studies to better understand learners' experiences and outcomes in multimedia learning environments. We think that the measure of situational engagement could be more sensitive to anthropomorphism than the measure of emotions, as engaged learners are more likely to process information more deeply and retain information better (Schunk and Mullen, 2012).

To better understand the potential effects of anthropomorphism on engagement, it is important to distinguish it from motivation, as these two concepts, while closely related, have some notable differences that warrant attention. Engagement is an action, whereas motivation is an intention (Christenson et al., 2012). Motivation is defined as the "will" to learn, while engagement is seen as the "competence" to effectively carry out the tasks involved in learning (Cleary and Zimmerman, 2012). In other words, motivation is the initial driving force that leads to engagement, which involves the actual execution of learning tasks. The extant literature suggests that the use of anthropomorphism may have a significant impact on e-learning outcomes (Brom et al. 2018), mental load, i.e., the cognitive resources required for processing information (Sweller, 1988), and motivation, i.e., the intention to learn. It is still unknown whether it impacts learner engagement, i.e., the actual investment in the task. Furthermore, even though anthropomorphism seems to impact motivation to learn (Gong et al, 2017; Kumar et al; 2016; Hsu et al., 2019), and motivation can lead to greater engagement *(*Sharek and Wiebe, 2015), it is also unknown whether motivation may mediate the relationship between anthropomorphism and engagement.

Therefore, in the present study we propose to explore the effect of anthropomorphism on mental load, motivation, engagement, and learning outcomes, as well as their relationships. Here, we compare the effectiveness of a rich version of e-learning, which includes anthropomorphism, to a neutral version, which does not include this feature. Participants used an e-learning about immunotherapy, completed subjective questionnaires about their engagement, motivation, mental load, and answered comprehension and memorization questions about what they had learned on immunotherapy. We employed subjective and objective measures of mental load, motivation, engagement, and learning performance. It is expected that anthropomorphism increases learners' mental load, motivation, and engagement, as well as learners' comprehension and memorization compared to a neutral version that does not incorporate this principle (H1). It is further expected that the effect of anthropomorphism on engagement will be mediated by motivation (H2). Finally, the effect of anthropomorphism on engagement would mediate the relationship between the e-learning version (rich vs. neutral) and learning performances, where the rich version leads to higher engagement, in turn leading to better learning performances (H3).

1. Methodology

1.1. Participants

Our study involved a total of 76 participants, recruited from psychology students or adults undergoing career changes. The average age of the participants is 31 years (*SD* = 9.9). They were 42 women and 34 men. All the participants were native French speakers. The participants were randomly assigned into the rich group (20 women and 18 men; M = 29 years; *SD* = 9.4)) and the neutral group (22 women and 16 men; M = 32 years; *SD* = 10.3). We performed a statistical power calculation to determine the sample size needed for our study based on results from the meta-analysis of Brom et al. (2018). Based on a desired power of 90 % to detect an effect *(*one-sided), a total of 71 subjects are required. With a sample of 76 participants, our statistical power is calculated to be 91%. This indicates that we have a high likelihood of detecting the anticipated effect with the given sample size and statistical power.

In accordance with the ethical standards set forth in the Helsinki Declaration, this experiment was guided by a commitment to the well-being and dignity of all individuals involved. This declaration emphasizes the importance of obtaining informed consent, minimizing harm, and protecting the rights of all research participants. Participants in this study were fully informed about the experimental procedures. All participants were given an in-depth explanation of the experiment, including its purpose, procedures, risks, and benefits, prior to providing their written consent. The participants were also assured of the confidentiality of their data and their right to withdraw from the study at any time.

1.2. Material

1.2.1. E-learning material and devices

The multimedia material consisted of an e-learning on immunotherapy. This e-learning allows us to understand the functioning of the immune system (innate and adaptive), how it reacts to cancer and presents immunotherapy. The support takes the form of 5 modules with a set of 30 short, animated videos with a voice-over. The plan of this e-learning is as follows: Module 1: The Immune System; Module 2: Immune System and Cancer; Module 3: Immunotherapy; Module 4: Differences between Immunotherapy and Chemotherapy; Module 5: Learn more about Immunotherapy. In e-learning, videos are animated, and a voice explains the content. There is also a text below the video for reading the content. The environment relies on a main interface that presents the different e-learning modules (see Fig.1). The user can click directly on one of the blocks to access its content. Once in the module, there is a video, with a next button, a previous button, and a back to menu button allowing the learner to navigate between videos (see Fig.1)

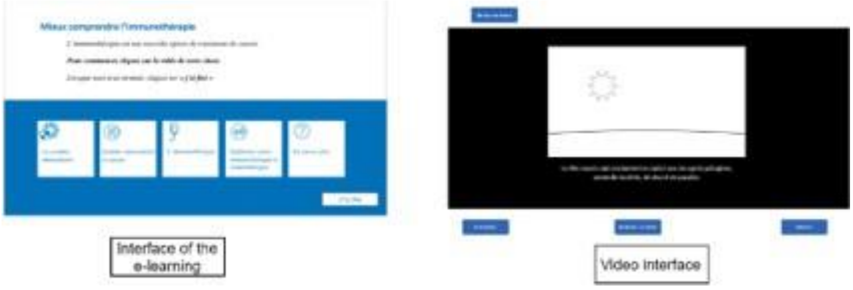

Figure 1: Learning materials. Interfaces of the e-learning environment: on the left the main interface to access the modules, on the right a screenshot of the interface to navigate in the video

Two versions of this e-learning were developed, one with anthropomorphism called the "rich version", and one without, called the "neutral version". In the neutral version, all the videos are in black and white, videos are composed of symbols that represent cells, and arrows with action verbs are used to present the links between these elements. The symbols are meant to visually convey the different cells and their functions in the immune system. For example, (see fig.2) the use of action verbs on the arrows is intended to show the dynamic interactions between the cells and how they work together to fight cancer. Additionally, it should be noted that this neutral version is presented in black and white. In the rich version, cells have emotions and facial expressions and express intentions allowing transposing of information. For example, we used an unhappy defense cell to indicate a sense of aggression towards a virus, and weapons to indicate the notion of attack by a defense cell. Night vision goggles express the ability of a defense cell to recognize a cancer cell. Videos are fully colored to understand the distinction between different elements.

Figure 2: Learning materials. Screenshots of two videos: the neutral version on the left and the rich version on the right

The experiment was conducted on a Dell Latitude 5530 laptop computer with a screen size of 13.3 inches. The e-learning material was displayed in full-screen mode, and the video was presented within a virtual rectangle of 1280 x 720 pixels. The display resolution was set to 1366 x 768 pixels, and the video was played using the default media player of the computer. The computer was operated on battery power, and the ambient lighting conditions of the room were consistent throughout the experiment. The angle of the screen was set at 90 degrees to the table on which the laptop computer was placed, with the viewer positioned directly in front of the screen, at 50 cm. The orientation of the screen ensured that the video was displayed in its optimal aspect ratio, with no distortion or loss of detail due to the screen angle.

1.2.2. Measures

Control measures.

Interest in the subject. We included a set of three questions to gauge learners' interest in the subject of immunotherapy (for example, "I was already interested in immunotherapy before the course"). Participants were asked to indicate their level of agreement with each statement on a Likert scale ranging from 1 (do not agree) to 5 (completely agree).

Prior Knowledge. We included one question to assess the participants' level of expertise on the subject of immunotherapy on a scale from novice to expert.

Measure of interest.

Mental load — Task Load Index. We use the Task Load Index (Hart and Staveland, 1988) to measure the cognitive load of the task, by evaluating six different dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. The questionnaire consists of rating scales for each dimension and asks participants to rate the task they just completed based on their perceived level of each dimension. The scores obtained from the NASA-TLX questionnaire are calculated based on a weighted average of the six dimensions mentioned above. The weighting of each dimension is determined by the individual participant, who is asked to rank the importance of each dimension for the specific task. The score ranges from 0 to 100, with higher scores indicating a greater perceived workload.

Motivation – Situational Motivation Scale (SIMS). We used the SIMS questionnaire developed by Guay et al. (2000) and translated and validated in French by Fontaine et al. (2019). This questionnaire consists of 16 statements rated on a 7-point Likert scale (1 = totally disagree; 7 = totally agree), which allows evaluation of four types of learners' motivation about the task they carried out: intrinsic motivation ("because I think that this activity is interesting," 4 items), identified regulation ("because I am doing it for my own good," 4 items), external regulation ("because I am supposed to do it," 4 items), and amotivation ("there may be good reasons to do this activity, but personally, I don't see any," 4 items). For each type of motivation, the score ranges from 4 to 28.

Engagement – User Engagement Scale (UES). All participants filled out the short form of the UES (O'Brien and Toms, 2010), translated and validated in French by Fontaine et al. (2019) The participant indicates on a 5-point Likert scale the degree of agreement *(*1 = totally disagree; 5 = totally agree) to each statement presented. The scale is composed of 4 dimensions, each one composed of 3 items: Focused attention (feeling of absorption in the interaction, and loss of track of time), Aesthetic appeal (attractiveness, and visual appeal of the interface), Perceived usability (negative affect experienced because of the interaction, and the degree of control and effort), and reward (amusement, novelty, interest, and overall satisfaction of the subject). The code for the items from perceived usability are reversed. A global engagement score can be calculated by adding all the items together and dividing by twelve. The higher the score is, the more the subject is engaged.

Engagement - Subjective Perception of Time (SPT). As this measure is relatively new in engagement research, we chose to follow the recommendations outlined by Hancock et al. (2019). We employed a retrospective paradigm in which learners were asked at the end of the course how much time (in minutes) they believed they spent navigating the e-learning modules, without consulting their phones or other devices. This perceived time was then divided by the actual time spent in the e-learning, with a higher score indicating an overestimation of the time elapsed.

Learning performances. We used a set of 10 comprehension and 10 memorization questions to evaluate learning performance after the task. Subjects were required to indicate whether the statement presented was correct or incorrect by answering Yes/No. An example of a memorization question is "The tears and the lymphocytes are the initial protections against the intrusion of bacteria". An example of a comprehension question is "The mechanism allowing lymphocyte education is slower than the innate immune response". Comprehension and memorization questions are two different types of assessment methods for learning performance. Comprehension questions assess a learner's understanding of concepts, while memorization questions assess the ability to recall specific details.

Learning Analytics. We used learning analytics to gather data on the learners' navigation patterns. This information was collected through tracking the learners' movements within the course, including their clicks on links, the time they spent on each page, and the pages they revisited.

1.3. Procedure

The experiment was conducted under standard laboratory conditions. Initially, participants were randomly assigned to one of the two conditions (neutral vs. rich). They filled out a first questionnaire to measure demographic data and initial interest in immunotherapy. Prior to the experiment, the participants' level of expertise on the subject matter of immunotherapy was assessed on a scale from novice to expert. Then, depending on the condition, the subjects used the e-learning on immunotherapy, composed of 5 modules, for a total minimum duration of 15 minutes. During this use, the navigation logs were collected, such as the number of clicks and the time spent on a particular module. The participants then completed all questionnaires described above to measure temporal absorption (SPT), engagement *(*UES), motivation (SIMS), and cognitive load (NASA-TLX). The questions for measuring comprehension and memorization were intermixed and randomly presented to each participant after they filled out all the scales mentioned above. The experimental sessions lasted about 30 minutes. Participants were instructed to engage with the e-learning material as if they were studying for a test, and to answer the comprehension and memorization questions to the best of their abilities. They were also instructed to rate their level of engagement, motivation, temporal absorption, and cognitive load, based on their personal experience during the e-learning session using the provided scales.

2. Statistical Analyses

In this study, we carried out a series of analyses. To address H1, we examined the differences in scores across various variables through a series of independent samples *t*-tests comparing the neutral and the rich groups. Cohen's *d* was used for the size effect. The significant *t*-test results are comprehensively presented in Table 1; (ii) The proportion of correct responses, the *d'* sensitivity index, and the *C* criterion index from the Signal Detection Theory (Macmillan & Creelman, 2004), were used as dependent variables for analysis of the learning outcomes. For these analyses, we conducted mixed-design ANOVAs with the question type (comprehension *vs.* memorization) as the within-participants factor, and version type (neutral *vs.* rich) as the between-subjects factor. The Signal Detection Theory was used to assess participants' responses by calculating two indexes: *d'* and *C*[2].(iii) Subsequently, we performed correlation analyses to explore the associations between mental load, motivation, engagement, and learning performance (see Table 2). (iv) To further investigate H1, we conducted a principal component analysis (PCA) to elucidate the underlying structure of the data and identify potential patterns among the variables being investigated (see Table 3). After conducting the PCA, the data were standardized and normalized to ensure comparability of the variables. The resulting standardized data were then used to compute composite scores for each of the four principal components, with higher scores indicating greater levels of the underlying constructs. We subsequently conducted *t*-tests to examine whether there were significant differences in these scores between the neutral and the rich versions. (v) In order to address H2 and H3, mediation analyses were performed to scrutinize the interconnections among the identified variables and to assess whether the impact of one variable on another was mediated by a third variable. More specifically, this research probed whether the influence of anthropomorphism on engagement was moderated through motivation.

3. Results

The age of participants assigned to each group was not different (rich M =32.50, SD = 10.34; neutral M = 29.57, SD = 9.41; t (74) = - 1.28; p = 0.20), and no differences were found as far as the number of women and men in each group was concerned, χ^2 (1) = 0.21, p > .05.

3.1. Effect of the version of support

We detail in this section all results regarding the effects of the version of support (neutral *vs.* rich). Only significant differences are reported in Table 1.

Control measures. There was no significant difference in the declared participants' degree of expertise of the content (neutral M = 1.48, SD = 0.24;

[2] The d' and C indexes are derived from the z-transformed hit and false alarm rates. Specifically, d' is computed as the difference between the z-transformed hit rate and the z-transformed false alarm rate, representing sensitivity to signal. Larger values of d' reflect greater sensitivity, i.e., an enhanced ability to discriminate signal from noise. The C index, or criterion, is calculated as the average of the z-transformed hit and false alarm rates, but with a negative sign applied to the false alarm rate. This reflects a response bias, with positive values indicating a bias towards claiming 'no signal' (or favoring noise responses) and negative values indicating a bias towards claiming 'signal' (or favoring signal responses).

rich M = 1.62, SD = 0.31; t (74) = 1.30, p = .19). There was no significant difference in the declared participants' initial interest in the content of the e-learning (neutral M = 2.60, SD = 0.90; rich M = 2.85, SD = 0.64; t (74) = 1.36, p = .17).

UES. The results showed a significant difference between the two groups (t (74) = 8.41, p <.001; Cohen's d = -1.93), with participants who received the neutral version having significantly lower scores (M = 2.37; SD = 0.43) than those who received the rich version (M = 3.23; SD = 0.45). Significant differences were observed between the two groups for the 3 factors:, i.e., focused attention (neutral M = 2.07, SD = 0.62; rich M = 3.32, SD = 0.79; t (74) = 7.58, p < .01, Cohen's d = 1.74), perceived usability (neutral M = 2.36, SD = 0.62; rich M = 3.12, SD = 0.67; t (74) = 5.04, p < .01, Cohen's d = 1.1), and aesthetic appeal (neutral M = 2.03, SD = 0.78; rich M = 3.34, SD = 0.82; t (74) = 7.05, p < .01, Cohen's d = 1.61).

SPT. The results revealed a statistically significant difference between the two groups (t (74) = 2.4, p = 0.02, Cohen's d = 0.55). Participants who received the neutral version overestimated the duration of the course (M = 1.42; SD = 0.79) more than those who received the rich version (M = 1.05; SD = 0.49).

SIMS. Results showed significant differences in the mean scores of the identified regulation (neutral M = 3.53, SD = 1.04, rich M = 4.13, SD = 0.94; t (74) = 2.58, p = .01, Cohen's d = 0.35), amotivation (neutral M = 5.12, SD = 0.96; rich M = 4.57, SD = 0.91; t (74) = 2.53, p = .01, Cohen's d = 0.35), and external regulation (neutral M = 5.22, SD = 0.83; rich M = 4.46, SD = 1.06; t (74) = 5.07, p < .01, Cohen's d = 0.63) between the two groups.

NASA-TLX. We did not observe any significant group difference in the cognitive load. Furthermore, we did not observe any significant effect of the version of support on the other subscales of the NASA-TLX, including mental demand, physical demand, temporal demand, performance, and effort.

Learning Outcomes. The analysis showed that there was no significant main effect of the type of question (F (1, 74) = 2.73, p = 0.10), and no significant interaction effect between the type of question and version of the support (F (1, 74) = 0.86, p = .35). Moreover, there was a significant main effect of support version (F (1, 74) = 5.66, p = .02; η_p^2 = 0.07). The participants in the neutral version condition (M = 0.58, SD = 0.02) had a lower proportion of correct responses compared to those in the rich version condition (M = 0.65, SD = 0.01).

Concerning the Signal Detection Theory indexes, the results showed that there was no significant main effect of question type on the d' sensitivity index (F (1, 74) = 2.899, p = .09), and no significant interaction effect between question type and support version (F (1, 74) = 0.706, p = .40). However, there was a significant main effect of support version (F (1,74) = 12.51, p < .01; η^2 = 0.10, neutral M = 0.53; SD = 0.02; rich M = 1.11; SD = 0.13). We did not observe any significant effect of the version on the C index.

Learning Analytics. The interactions of the learners with e-learning support were not analyzed due to lack of sufficient data that prevented us from conducting adequate analyses.

Table 1: Summary of significant differences between the two versions of support

Measure	Statistic	p-value	Effect Size (Cohen's d	Neutral version, M (SD)	Rich version M (SD)

			or η2)		
UES – Mean Score	*t* =-8.41	< .001	1.93	*M* = 2.37; *SD* = 0.43	3.23 (0.45)
UES – Focused Attention Factor	*t* =-7,58	< .001	1.74	*M* = 2.07, *SD* = 0.62	*M* = 3.32, *SD* = 0.79
UES – Perceived Usability Factor	*t* =-5.04	< .001	1.15	*M* = 2.36, *SD* = 0.62	*M* = 3.12, *SD* = 0.67
UES - Aesthetic Appeal Factor	*t* =-7.05	< .001	1.61	*M* = 2.03, *SD* = 0.78	*M* = 3.34, *SD* = 0.82
Subjective Perception of Time	*t* =2.41	<.05	0.55	*M* = 1.42; *SD* = 0.79	*M* = 1.05; *SD* = 0.49
SIMS - Identified Regulation	*t* =2.58	<.05	0.59	*M* = 3.53, *SD* = 1.04	*M* = 4.57, *SD* = 0.91
SIMS- External Regulation	*t* =5.07	< .001	1.16	*M* = 5.22, *SD* = 0.83	*M* = 4.46, *SD* = 1.06
SIMS - Amotivation	*t* =-2.53	<0.05	0.58	*M* = 5.12, *SD* = 0.96	*M* = 4.57, *SD* = 0.91
Proportion of Correct Responses	*F*(1,74) = 5.66	<0.05	0.03	*M* = 0.58, *SD* = 0.02	*M* = 0.65, *SD* = 0.01
d' indicator	*F*(1,74) 12.51	<0.01	0.05	*M* = 0.53; *SD* = 0.02	*M* = 1.11; *SD* = 0.13

3.2. Correlation Analysis

To address H1, Table 2 presents the Pearson correlation coefficients between various variables, including UES (focused attention, perceived usability, aesthetic appeal, and reward), SPT *(*subjective perception of time), mean d', mean C, and the four types of motivation (intrinsic motivation, identified regulation, external regulation, and amotivation). The results show several significant correlations among the variables studied.

Variable	1	2	3	4	5	6	7	8	9	10	11
1. NASA-TLX	--										
2. Subjective	-0.079	--									

Perception of Time											
3.UES – Focused Attention	-0.002	-0.262*	--								
4. UES – Perceived Usability	-0.026	-0.276*	0.417***	--							
5. UES - Aesthetic Appeal	0.137	-0.068	0.486***	0.458***	--						
6. UES - Reward	0.154	0.083	0.132	0.223	0.358**	--					
7. SIMS - Intrinsic Motivation	0.004	-0.072	0.021	0.061	0.025	0.148	--				
8. SIMS - Identified Regulation	0.037	-0.047	0.188	0.030	0.334**	0.296**	0.088	--			
9. SIMS - External Regulation	-0.068	0.134	-0.265*	-0.144	-0.332**	0.013	0.084	-0.386***	--		
10. SIMS - Amotivation	0.121	-0.005	-0.162	-0.176	-0.076	-0.055	0.028	0.068	0.105	--	
11. Mean d'	0.254*	-0.045	0.177	0.069	0.257*	-0.005	-0.070	0.042	-0.148	-0.002	--
12. Mean C	-0.018	0.142	0.096	-1.591e-4	0.044	-0.012	0.030	0.007	0.132	0.188	-0.089
13. Mean UES	-0.027	-0.159	0.531***	0.359***	0.510***	0.171	-0.052	0.222	-0.378***	-0.165	0.364**

Note: * p <0.05, ** p <0.01, *** p <0.001

Table 2 Correlation analysis between variables.

Regarding the motivational and engagement variables, there is a significant positive correlation between the global score on UES and mean d' (r = 0.364, p < .01), and a significant negative correlation between the global score on UES and SIMS – external regulation (r = -0.37, p < .01). We found a significant positive correlation between mean d' and UES – aesthetic appeal (r = 0.25, p < .05). Concerning the subjective perception of time, it showed significant negative correlations with focused attention (r = -0.26, p < .05) and perceived usability (r = -0.27, p < .01). Among the motivational factors, SIMS - Identified regulation demonstrated a significant positive correlation with UES - Aesthetic appeal (r = 0.33, p < .01) and UES – Reward (r = 0.29, p < .05). Conversely, SIMS - External regulation exhibited a significant negative correlation with UES – Focused attention (r = -0.26, p < .05), UES – Aesthetic appeal (r = -0.33, p < .05) and Identified regulation (r = -0.38, p < .01). These significant correlations among the variables provide evidence for the need to conduct a principal component analysis to further explore the underlying dimensions and relationships within the dataset.

3.3. ***Principal component analysis (H1).***

The PCA was conducted on a diverse set of variables that incorporated the four types of motivation from SIMS, the four factors of UES, NASA-TLX, mean d' and mean C indexes, and SPT. The PCA was executed with a set eigenvalue of 1 and a factor loading of 0.5, using an orthogonal varimax rotation to maximize the variance of the factor coefficients. This analysis resulted in the identification of six principal components, each encapsulating a unique cluster of variables. The subsequent sections provide a detailed interpretation of these six components. Refer to Table 3 for a complete overview of the principal components.

	Cognitive Effort	**Engagement**	**Internal Regulation**	**Non-valuing**	**Interest**	**Absorption**
------------------	--------	--------	--------	--------	--------	--------
Proportion of correct responses	0.95					
Mean d'	0.94					
NASA-TLX	0.51					
UES – Focused attention		0.78				
UES – Perceived usability		0.78				
UES - Aesthetic appeal		0.65				
SIMS - Identified regulation			0.77			
SIMS - External regulation			-0.76			
Mean C				0.82		

SIMS – Amotivation				0.69		
SIMS - Intrinsic motivation					0.77	
UES - Reward					0.61	
Subjective perception of time						0.87
Explained Variance (%)	16.3	15.6	12.6	9.4	9.3	8.8
Cumulative Explained Variance (%)	**16.3**	**31.9**	**44.5**	**53.9**	**63.2**	**72.0**

Table 3: Component Characteristics and Explained Variance

PC1: Cognitive Effort. This component encapsulates the learning outcomes and the perceived level of multimedia content difficulty. It appears to represent the overall cognitive effort exerted by the learner in the task.

PC2: Engagement. This component is highly correlated with variables related to perceived usability, aesthetic appeal, and focused attention. It seems to capture the participants' subjective engagement with multimedia learning.

PC3: Internal Regulation. This component is characterized by identified regulation and negative external regulation, signifying a trend towards self-determined behaviors.

PC4: Non-valuing. This component appears to be predominantly composed of variables connected to response bias and a lack of motivation. It indicates a state of disinterest or boredom in the learner, which could result in biased responses.

PC5: Interest. This component represents the extent to which individuals find interest with multimedia content internally rewarding. It implies that the user experiences a gratifying interaction with the multimedia content and an intrinsic motivation for the task, denoting a profound interest and satisfaction derived from the learning environment.

PC6. Absorption. This component seems to reflect the extent to which learners become so absorbed in the task that they lose track of time.

Following identification of the principal components, we proceeded with data preprocessing. This involved standardizing and normalizing the data for each variable within the components. After standardizing and normalizing the data, we computed the average for each principal component. Subsequently, we conducted independent samples *t*-tests based on the version of the support (neutral vs. rich). For PC1, which represents cognitive effort, we observe a highly significant difference (neutral M = 0.30, SD = 0.15, rich M = 0.67, SD = 0.16; t (74) = -10.19, p <.001, Cohen's d = -2.35) between the two groups. PC2, representing engagement, also shows a significant difference (neutral M = 0.44, SD = 0.25, rich M = 0.55, SD = 0.19; t (74) = -2.04, p =.04, Cohen's d = -0.47). PC4, which represents non-valuing, shows a significant difference (neutral M = 0.55, SD = 0.2, rich M = 0.44, SD = 0.22; t (74) = 2.2, p = .03, Cohen's d = 0.51).

For PC3 and PC5, which represent internal regulation and interest, respectively, we did not observe any statistical significance (p = .32; p = .11). Concerning the learner's absorption component, it was analyzed in the Section about t-test analysis, and we observed a significant difference (t (74) = 2.4, p = .02, Cohen's d = 0.55; neutral M = 1.42; SD = 0.79; rich version M = 1.05; SD = 0.49).

3.4. Statistical test for H2 and H3

Mediation analysis. To address H2, we conducted a mediation analysis between the four types of motivation as predictor, UES as mediator, and mean d', mean C, and mean proportion of correct responses as outcomes. The direct effect of external regulation on mean d' was found to be non- significant (β = 0.01, z = 0.10, p = .91). The indirect effect of external regulation on mean d' through UES is significant (estimate = -0.10, SE = 0.04, z = -2.34, p = .01). The 95% confidence interval for this effect ranges from -0.184 to -0.016. The total effect of external regulation on mean d' was found to be non-significant (β = 0.12, z = 1.30, p = .19). To sum up, our findings suggest that the influence of external regulation on mean d' is primarily exerted through the UES variable.

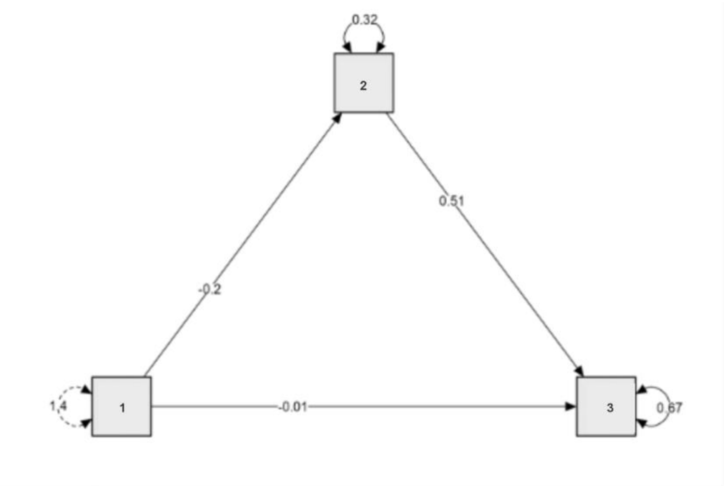

Fig 3. Path plot mediation of (1) External Regulation (2) UES and (3) mean d'

To address H3 (the impact of the support version as a predictor), we conducted a serial mediation, with the version of the support as a predictor, motivation as the first mediator, UES as the second mediator, and learning outcomes as the dependent variables. We did not find a significant indirect effect of the version on different types of motivation, engagement, and learning outcomes (p >.05). A simple mediation revealed a trend towards a significant indirect effect of the version of support on mean d' through UES as a mediator (estimate = 0.32, p = .08), but not with other outcome variables (mean C estimate = 0.09; p= .31; and mean proportion of correct responses estimate = .02, p = 0.8).

4. Discussion

This study aimed to explore the impact of anthropomorphized support versus neutral support on learner engagement, motivation, cognitive load, and learning outcomes in an e-learning setting (H1). A secondary objective was to elucidate how engagement mediates the effect of motivation on learning outcomes (H2). Lastly, the study investigates whether a serial mediation exists between the version of the support used, learner motivation, engagement, and subsequent learning outcomes (H3).

Regarding our first hypothesis, contrary to previous studies, our findings showed no significant impact of the version of support on intrinsic motivation (Gong et al., 2017a; Kumar et al., 2016). Instead, the rich multimedia condition was associated with higher identified regulation and lower amotivation and external regulation. According to Ryan and Deci (2000), motivation exists along a continuum, ranging from intrinsic motivation at one end, through identified regulation, to external regulation, and finally amotivation at the other end. This suggests that the enriched multimedia environment might not directly enhance learners' intrinsic motivation in the learning task itself, but it significantly supports learners in recognizing the personal relevance or value of the learning task (identified regulation), while also decreasing both the absence of motivation (amotivation) and the reliance on external rewards or punishments (external regulation). Moreover, the component analysis revealed a significant difference between the two versions of the support regarding the learner non-valuing component, the neutral version leading to an increase in disinterest. Thus, while not influencing intrinsic motivation *per se*, rich multimedia support appears to promote more autonomous forms of extrinsic motivation and diminish non-autonomous motivation, aligning with the overarching goal of the Self-Determination Theory to foster greater autonomy in learning. In the studies analyzed in a recent meta-analysis by Brom (2018), intrinsic motivation was typically measured through established questionnaires (e.g., Isen & Reeve, 2005) or researcher-developed short questionnaires. Considering these methodologies, it may be beneficial to explore motivation more thoroughly and distinguish more finely between different types of motivation, as many studies utilize ad hoc questionnaires that may not capture all nuances of motivation.

Regarding engagement, we observed noteworthy differences between the two groups in terms of the global UES score and across three components of the User Engagement Scale (UES) - focused attention, perceived usability, and aesthetic appeal. It underlines the potential of such support to not only draw learners' attention but also to improve the usability of the e-learning environment and create a visually appealing learning experience. Our empirical findings align well with and augment prior research. For instance, Park et al. (2015) and Um et al. (2012b), proposed that augmentations like rich multimedia and anthropomorphized elements can significantly improve the e-learning experience by stimulating affective-motivational states. This list includes key factors such as enjoyment, flow, and situational interest. Building on these findings, our study suggests the addition of a further significant factor - learner engagement with the task. An intriguing aspect of our findings is the observed negative correlation between external regulation and the overall score on the User Engagement Scale (UES). More specifically, the aesthetic appeal factor of the UES emerged as particularly significant in this relationship. This highlights the importance of the aesthetic aspect of the e-learning environment, with our findings suggesting that an attractive, engaging environment may lead learners to perceive the task as less constraining. Some studies, however, fail to measure engagement in a comprehensive manner, overlooking the subtler dimensions of this concept. For instance, according to the meta-analysis by Brom et al. (2018), the measurement most aligned with engagement is often encapsulated in terms of liking or enjoyment. This is typically quantified using researcher-created questionnaires, usually consisting of 1-3 items, which evaluate the appeal, likability, and/or enjoyment of the lesson or materials.

While these measures are effective in capturing an overall sense of enjoyment and likability, they might lack the intricacy required to fully comprehend the multidimensional nature of learner engagement. Notably, evaluation of the aesthetic appeal of the environment—a factor that our study has revealed to significantly impact engagement—may not be thoroughly represented in this conventional approach. Consequently, it might be beneficial to incorporate additional and complementary measures of engagement. For instance, our study employed the subjective perception of time as a potential supplemental measure. Such measures could provide a more nuanced understanding of engagement in e-learning environments.

In fact, we observe a lower subjective perception of time among participants exposed to the rich multimedia support, hinting at an increased immersion in the learning experience. Furthermore, we found a significant correlation between this subjective perception of time and focused attention and perceived usability from the UES. In essence, the more absorbed participants were in their tasks (as indicated by higher scores on the focused attention factor) and the more usable they found the e-learning platform (evidenced by higher scores on the perceived usability factor), the less aware they were of the time elapsed. Indeed, it is not surprising to find a correlation between the subjective perception of time and the focused attention factor of the UES, considering that the items used to measure focused attention are inherently based on the concept of absorption. For instance, one of the statements used in focused attention is: "I was so involved in this experience that I lost track of time.". The correlations we found between this perception and both the focused attention and perceived usability factors of the UES suggest that rich multimedia support can indeed foster a state of flow in e-learning environments. These results underscore the relevance of subjective time perception as a complementary measure to engagement metrics in multimedia learning (Hancock et al., 2018) and illustrate the sensitivity of engagement and flow state to anthropomorphism (Kahu, 2013; Crawford and Simons, 2021; Schilling, 2009).

Our findings confirm the effect of anthropomorphism on learning outcomes, with participants exposed to rich support outperforming those in the neutral condition. This resonates with earlier research advocating the enhancement of learning outcomes by rich multimedia support through the facilitation of new information integration into existing mental schema (Mayer, 2009). The rich version appeared to aid learners in identifying pertinent information from the content, subsequently elevating the proportion of correct responses without altering response strategies. In addition, we found a positive correlation between the overall score on the UES and the mean d’ from the Signal Detection Theory. This suggests that as user engagement increases, learners’ sensitivity to distinguishing relevant signals from noise also improves, which could potentially enhance their performance on tasks. In other words, high engagement seems to enhance learners' ability to identify and process pertinent information accurately, reflecting the construct's crucial role in optimizing learning outcomes. Although no significant effect of the type of the version of support on cognitive load was detected, contrary to other studies (see the meta-analysis from Brom et al.2018), we observed a clear association between cognitive load and learning outcomes in the PCA. This relationship hints at the interdependency of these factors in the learning process. Moreover, we found that the "version" significantly impacts the composite cognitive effort factor. Interestingly, while learners perceived the rich support version as more challenging, they simultaneously achieved superior learning outcomes. This observation aligns with the Cognitive Load Theory Sweller (1988) and the desirable difficulties concept *(*Anderson et al., 1994).

Finally, our mediation analysis revealed a significant indirect effect of external regulation on learning outcomes through engagement. When a learner is primarily motivated by external factors (such as rewards or punishments), their engagement levels decrease, which in turn results in lower learning outcomes. This suggests that the engagement of learners plays a crucial role in transforming motivation into learning outcomes. Our findings resonate with the model proposed by Christenson et al. (2012), which suggests that engagement is the outcome of motivation and is distinct from motivation itself. In this perspective, engagement is the external and visible manifestation of motivation. It emphasizes that the type of motivation (rather than the amount) significantly influences the quality of engagement and, subsequently, the quality of learning. According to this perspective, being motivated by external factors (external regulation) is less likely to lead to deep, meaningful engagement compared to more autonomous forms of motivation. To sum up, our findings validate our second hypothesis, demonstrating that learner engagement serves as a crucial bridge between external regulation and learning outcomes. However, our analysis did not provide evidence to validate this third hypothesis. The support version did not appear to affect learning outcomes through a sequential mediation process involving both motivation and engagement. This underscores the need for further research to elucidate the complex relationships between the support version, learner motivation, engagement, and learning outcomes in e-learning contexts.

To conclude, our study emphasizes the role of robust multimedia support in learner engagement, motivation, and learning outcomes. While we partially align with previous research highlighting the significant role of multimedia in enhancing learner engagement and fostering motivation (Liu et al., 2009), we introduce important nuances. We utilized more detailed engagement and motivation questionnaires, specifically addressing the different types of motivation as delineated by Deci and Ryan, and engagement as per O'Brien's formulation. Our findings demonstrated improved learning outcomes among participants using the enriched multimedia version of the learning tool. This reinforces the potential of well-designed multimedia support to boost learner performance, in line with Mayer's multimedia learning theory (Mayer, 2009). Additionally, the interplay between cognitive load and multimedia support deserves further exploration, in pursuit of a more nuanced understanding of the cognitive processes involved in the learning experience (Paas, Renkl, & Sweller, 2004).

It is essential to point out the need for incorporating additional measures in future research endeavors. Learning is a multifaceted and complex process that demands a more rigorous methodology. Solely relying on questionnaires, especially unvalidated ones, limits the depth of data collected and can lead to skewed or incomplete results. Therefore, it is paramount to diversify the assessment tools used in future studies to gain a more comprehensive understanding of learners' experience and the influence of multimedia elements on their learning trajectory. Additionally, our investigation of learning outcomes should extend beyond immediate recall or comprehension tests. Memory retention over an extended period is a crucial aspect of effective learning, and an immediate post-task test may not provide a complete picture of the learner's cognitive achievement. Therefore, future research should endeavor to evaluate long-term memory retention to better understand the lasting effects of multimedia-supported e-learning. These future investigations will be essential in corroborating our current findings and will provide a more in-depth understanding of the impact of multimedia on various aspects of the learning process. While our study provides compelling evidence for the beneficial impact of rich support on cognitive and engagement factors, it is not without limitations.

Primarily, our findings are situated in a university student population within a controlled experimental setting, which may limit the generalizability of the results (Hays et al., 2016). The short duration of the experiment might also limit the representation of long-term effects of rich support on learning outcomes. Thus, future studies should take these limitations into account, designing more longitudinally oriented experiments to assess the influence of enriched support more accurately on learning outcomes. Ultimately, our research contributes valuable insights for future design and implementation of enriched multimedia support in e-learning environments, aiming to maximize learner engagement and success. Unfortunately, our study faced limitations with learning analytics, as the e-learning environment did not offer sufficient interactive opportunities to discern differences between the two groups. Future research should address this by exploring more interactive e-learning settings.

5. References

Anderson, M. C., Bjork, R. A., and Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of experimental psychology: Learning, memory, and cognition*, 20(5):1063.

Bellizzi, J. A. and Hite, R. E. (1992). Environmental color, consumer feelings, and purchase likelihood. *Psychology & Marketing*, 9(5):347–363.

Brom, C., Stárková, T., and Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review*, 25:100–119.

Christenson, L., Reschly, S., L, Wylie, A., Widiasani, C., and A (2012).

Cleary, T. J. and Zimmerman, B. J. (2012). A Cyclical Self-Regulatory Account of Student Engagement: Theoretical Foundations and Applications. *Handbook of Research on Student Engagement*, pages 237–257.

Dhika, H., Destiawati, F., Sonny, M., and Surajiyo (2019). Study of the use and application of the moodle e-learning platform in high school. *Journal of Physics : Conference Series*, 1175 :12219–12219.

Fontaine, G., Cossette, S., Maheu-Cadotte, M.-A., Mailhot, T., Lavoie, P., Gagnon, M.-P., Dubé, V., and Côté, J. (2019). Traduction, adaptation et évaluation psychométrique préliminaire d'une mesure d'engagement et d'une mesure de charge cognitive en contexte d'apprentissage numérique. *Pédagogie Médicale*, 20(2):79–90.

Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74(1):59–109.

Gong, S., Shangguan, C., Zhai, K., and Guo, Y. (2017a). The effects of emotional design on multimedia learning. *Acta Psychologica Sinica*, 49(6):771–771.

Gorn, G. J., Chattopadhyay, A., Yi, T., and Dahl, D. W. (1997). Effects of Color as an Executional Cue in Advertising: They're in the Shade. *Management Science*, 43(10):1387–1400.

Guay, F., Vallerand, R. J., and Blanchard, C. (2000). *Motivation and Emotion*, 24(3):175–213.

Hancock, P. A., Kaplan, A. D., Cruit, J. K., Hancock, G. M., Macarthur, K. R., and szalma, J. L. (2019). A meta-analysis of flow effects and the perception oftime. *Acta Psychologica*.

Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, 52:62386–62395.

Hays, D. G., Wood, C., Dahl, H., and Kirk-Jenkins, A. (2016). Methodological rigor in Journal of Counseling & Development qualitative research articles: A 15-year review. *Journal of Counseling & Development*, 94(2):172–183.

Kumar, J. A., Muniandy, B., and Yahaya, W. A. J. W. (2016).

Liew, T. W., Pang, W. M., Leow, M. C., and Tan, S. M. (2022). Anthropomorphizing malware, bots, and servers with human-like images and dialogues: The emotional design effects in a multimedia learning environment. *Smart Learning Environments*, 9(1).

Liu, M., Toprac, P., and Yuen, T. T. (2009). What factors make a multimedia learning environment engaging: A case study.

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. *The Cambridge handbook of multimedia learning*, pages 31–48.

Mayer, R. E. (2009). Multimedia Learning. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, *33*, 12–18. https://doi.org/10.1016/j.learninstruc.2014.02.004

Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, 22(3):149–158.

O'brien, H. L. and Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69.

Park, B., Knörzer, L., Plass, J. L., and Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: An eye tracking study on the use of anthropomorphisms. *Computers & Education*, 86:30–42.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., and Um, E. (2014).

Plass, J. L. and Kaplan, U. (2016). Emotional Design in Digital Media for Learning. *Emotions*,

pages 131–161.
Ryan, R. M. and Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation. *Social Development, and Well-Being. American Psychologist*, pages 11–11.

Schneider, S., Nebel, S., and Rey, G. D. (2016). Decorative pictures and emotional design in multimedia learning. *Learning and Instruction*, 44:65–73.

Schunk, D. H. and Mullen, C. A. (2012). Self-efficacy as an engaged learner. *Handbook of research on student engagement*, pages 219–235.

Sharek, D. and Wiebe, E. (2015). Investigating Real-time Predictors of Engagement: Implications for Adaptive Videogames and Online Training. *International Journal of Gaming and Computer-Mediated Simulations*, 7(1):20– 37.

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement*. Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-7089-2Skuballa, I. T., Dammert, A., and Renkl, A. (2018).

Stárková, T., Lukavský, J., Javora, O., and Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning, jcal.12359*.

Sun, J. C.-Y. and Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British journal of educational technology*, 43(2):191–204.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):90023– 90030.

Um, E., Rachel, Plass, J. L., Hayward, E. O., and Homer, B. D. (2012a). Emotional design in multimedia learning. *Journal of Educational Psychology*, 104(2):485–498.

Um, E., Rachel, Plass, J. L., Hayward, E. O., and Homer, B. D. (2012b). Emotional design in multimedia learning. *Journal of Educational Psychology*, 104:485–498.

Wolfson, S. and Case, G. (2000). The effects of sound and color on responses to a computer game. *Interacting with Computers*, 13(2):183–192.

Transition

À partir de cette première expérimentation, nous avons pu atteindre plusieurs objectifs de recherche. Premièrement, les résultats ont démontré que l'engagement situationnel de l'apprenant, en tant qu'état dynamique au cours de la tâche d'apprentissage, est un indicateur pertinent pour mesurer l'impact du design émotionnel d'un contenu multimédia. Notre protocole combinant mesures subjectives et objectives de l'engagement s'est avéré suffisamment complet et sensible pour appréhender finement les liens entre engagement, motivation, charge cognitive et performances d'apprentissage. Nos analyses ont en effet mis en évidence que l'engagement joue un rôle médiateur crucial entre la motivation extrinsèque et les performances d'apprentissage. Un niveau élevé d'engagement permet aux apprenants de mieux détecter les informations pertinentes et de favoriser leur intégration en mémoire.

De plus, conformément à notre modèle intégratif, nous avons pu confirmer que les caractéristiques de la tâche, ici le niveau d'enrichissement multimédia, influencent significativement le niveau d'engagement des apprenants et leurs performances d'apprentissage. Cependant, notre modèle postule également que l'engagement situationnel dépendrait non seulement des facteurs motivationnels et des caractéristiques de la tâche, mais aussi des capacités attentionnelles propres à chaque apprenant (Cowan, 2014). Ces capacités influenceraient leur aptitude à réguler et déployer leurs ressources attentionnelles pour sélectionner les informations pertinentes, inhiber les distractions et gérer la charge cognitive (Lavie et al., 2004 ; Sweller et al., 2011).

La question qui se pose alors est de comprendre comment ces capacités attentionnelles individuelles peuvent moduler l'impact des caractéristiques de la tâche sur l'engagement réel au cours de l'apprentissage multimédia (Kizilcec et al., 2017 ; Mayer, 2018). Un apprenant disposant de faibles ressources attentionnelles pourrait en effet ne pas parvenir à s'engager pleinement malgré un support pédagogique attractif et motivant. À l'inverse, de fortes capacités attentionnelles non mobilisées par un contenu peu engageant limiteraient l'efficacité de l'apprentissage (Pekrun & Linnenbrink-Garcia, 2012).

Chapitre 4 : Le rôle des capacités attentionnelles sur l'impact des agents pédagogiques et l'engagement

Pour explorer ces interactions, nous avons choisi d'étudier un principe de design spécifique issu de la théorie cognitive de l'apprentissage multimédia (Mayer, 2005) : l'utilisation d'un agent pédagogique anthropomorphe. Reconnu pour son potentiel à favoriser l'engagement des apprenants (Schroeder & Adesope, 2014 ; Veletsianos & Russell, 2014), l'agent pédagogique peut aussi générer des situations d'attention divisée délétères (Lester et al., 1997 ; Moreno et al., 2001).

La notion de "force" de l'agent pédagogique est un axe de recherche important dans ce domaine. Elle se réfère au degré d'animation et d'expressivité des agents. Il est suggéré que des agents dotés d'une grande expressivité et capables d'imiter les comportements humains, tels que les expressions faciales ou les gestes, peuvent créer une connexion plus forte avec les apprenants (Abdullah et al., 2017). Par exemple, Mayer et DaPra (2012) ont montré que les étudiants apprenaient mieux à partir d'une animation narrée lorsque le narrateur était représenté par un agent pédagogique au comportement non-verbal naturel (expressions faciales, gestes) plutôt que par une simple voix. De même, une méta-analyse de Schroeder et Adesope (2014) a révélé que les agents pédagogiques ayant une apparence humaine réaliste et exprimant des émotions positives favorisaient davantage l'apprentissage que les agents au design plus minimaliste ou neutre.

Nous avons donc conçu une nouvelle expérimentation faisant varier la présence et le niveau d'expressivité de l'agent pédagogique (absent, faible, fort) au sein de l'environnement multimédia (Clark & Mayer, 2016). En conservant le protocole de l'étude précédente, nous y avons ajouté des mesures objectives des capacités attentionnelles des participants (Underwood & Everatt, 1996) afin d'examiner leurs liens avec l'engagement situationnel et les performances d'apprentissage. Les objectifs sont multiples:

- Répliquer les effets positifs de l'agent pédagogique sur l'engagement et la motivation (e.g., Johnson & Lester, 2016),
- Mesurer l'impact de l'agent sur le déploiement des ressources attentionnelles.
- Clarifier les relations entre capacités attentionnelles, engagement situationnel et performances.

Cette deuxième étude empirique devrait ainsi nous permettre de mettre à l'épreuve les prédictions de notre modèle intégratif concernant les interactions entre les caractéristiques de l'environnement d'apprentissage (ici la présence et l'expressivité de l'agent pédagogique), les processus attentionnels et motivationnels des apprenants, et leurs performances d'apprentissage (Domagk et al., 2010). En examinant ces interactions complexes, notre objectif est d'apporter un nouvel éclairage sur les conditions d'optimisation du design des agents pédagogiques dans les environnements d'apprentissage multimédia, en fonction des profils attentionnels et motivationnels des apprenants (Heidig & Clarebout, 2011 ; Schroeder et al., 2013).

The Impact of Pedagogical Agent on Learners' Attentional Abilities, Performances, Motivation and Engagement in a Multimedia Learning Environment

Y. Minne[1,2,3], E. Lavoué[3] C. Thiery[2] and G. A. Michael[1]

[1]Unité de Recherche Étude des Mécanismes Cognitifs, Université de Lyon, Université Lyon 2, Lyon, France
[2] Sydo, 69001 Lyon, France
[3] iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, Université de Lyon, Université Lyon 3, Lyon, France

Abstract

Background: The integration of pedagogical agents in multimedia learning environments has shown varied implications for learners, particularly in terms of their engagement, motivation, comprehension, and memorization. There are few studies who explore how the impact of these agents interacts with learners' attentional capacities.

Objective: This study aims to unravel the differential impact of pedagogical agents on learners' engagement, motivation, comprehension, and memorization in multimedia learning environments. It also seeks to explore how individual differences in attentional abilities modulate these effects.

Methods: Ninety-four participants were exposed to an e-learning module on immunotherapy, presented in three versions: without any agent, with a weak agent (lacking movements or expressions), or with a strong agent (featuring rich movements and expressions). The study assessed post-task performance in comprehension and memorization, and measured engagement. Attentional abilities such as inhibition of task-irrelevant information and the simultaneous processing of multiple relevant information were evaluated using a computerized task.

Results: The presence of a strong pedagogical agent significantly enhanced learners' engagement and reduced external motivation, but not intrinsic motivation. We also found that pedagogical agents had distinct effects on comprehension and memorization. The mere presence of an agent was detrimental for memorization, while comprehension was impaired with a weak agent and improved with a strong agent. Finally, we found a differential effect of attentional abilities on performance independently from the APA version.

Conclusion: Pedagogical agents in multimedia learning environments present complex challenges and benefits. Their impact on learner engagement, motivation, comprehension, and memorization is multifaceted.

1. Introduction

Traditional teaching in the classroom allows teachers to enhance learners' attention towards the learning contents (Wang, 2019). However, remote multimedia learning often lacks these inherent attention-guiding mechanisms found in physical classrooms (Borokhovski et al., 2016). Additionally, there is a high dropout rate in such environments (Rostaminezhad et al., 2013), Consequently, learners' engagement is crucial to push learners to complete the learning session.

Animated pedagogical agents (APA) have garnered significant interest in the context of multimedia learning (Mayer, 2005). APA are virtual entities integrated into the instructional software to enhance learning via interactions with learners. These agents can mimic human educators, using communication media like voice, facial cues, and gestures (Baylor & Kim, 2005; Schroeder, Adesope, & Gilbert, 2013). APA have been shown to attract up to 56% of learners' gaze (Van Mulken, Andre, & Muller, 1998) and to amplify intrinsic motivation in learners, this being the inner drive to engage with the material for its own sake, rather than for some external reward (Ryan & Deci, 2000). Meta-analyses on the impact of animated pedagogical agents (APA) in multimedia learning environments have demonstrated their significant influence on learning outcomes, intrinsic motivation, and engagement (Davis, 2018, Martha & Santoso, 2019, Schroeder et al., 2013, Schroeder & Adesope, 2014).

For example, Moreno (2005) found that when pedagogical agents speak in a conversational tone and display enthusiasm, learners report higher interest and enjoyment in the task, which are key components of intrinsic motivation. APA can also improve learners' engagement by incorporating emotions in human-agent interactions (Romero-Hall, 2016). Chin et al. (2016) demonstrated that integrating APA into multimedia learning systems not only enhances students' engagement in learning activities but also impacts learners' feelings, reduces anxiety, and creates a safe learning environment leading to engagement. A study conducted by D'Mello, et al. (2008) on affect-sensitive APA found that when APA responded to learners' emotions appropriately, it resulted in improved engagement and better learning outcomes. APA can heighten students' perceptions of the relevance and value of tasks, which in turn positively influences their engagement to the learning process and the effort they are willing to invest (Wigfield & Eccles, 2000). When students perceive their APA as supportive and informative, their belief in their capabilities to learn and understand the content increases (Meij, Meij, and Harmsen, 2015). This enhanced confidence may lead to a higher level of engagement with the learning material, as students are more likely to take on challenges, persist in the face of difficulty, and ultimately achieve better learning outcomes (Pajares, 1996). Wang, Li, Mayer, and Liu (2018) more particularly showed that when APA use hand gestures to emphasize important points or to represent abstract concepts, learners tend to be more engaged because the material becomes easier to visualize and comprehend. Our study builds on these previous findings to investigate the impact of APA on learners' engagement and motivation in a multimedia learning environment.

APA also enhance the cognitive dimension of learning like memorization and comprehension of content (Martha & Santoso, 2019). The ability of APA to aid in memorization and content comprehension is substantial, as evidenced by studies demonstrating improved retention and understanding when learners interact with these agents (Martha & Santoso, 2019While these agents undeniably play a pivotal role in learning outcomes, they also significantly impact cognitive dimensions such as guiding attention (Li et al, 2023). APA stimulate selective attention by drawing the learners' focus on essential instructional content through their dynamic gestures and vocal modulations (Schroeder, Adesope, & Gilbert, 2013). Attention can be shifted, focused, and redirected towards specific stimuli, filtering out other extraneous information (Posner, 1980). Animated Pedagogical Agents (APA) guide attention by leveraging human-like gestures, vocal intonations, and facial expressions, and strategically guide learners' attention towards key aspects of instructional content (Baylor & Kim, 2009).

However, at times, the mere presence, animation, and complexity of APA may act as a distractor, potentially diverting learners' attention from the primary instructional content (Schroeder & Adesope, 2013) and thus compromising optimal learning outcomes (Park, 2015; Ayres & Sweller, 2005). APA may also prove challenging in that they introduce additional cognitive load, where learners have to split their focus between the agent and the content (Sweller, 1988). Effective APA should guide attention towards relevant content without introducing extraneous and distracting elements that would require extra processes, such as inhibition, to control potential interference. The need for inhibition may deplete available attentional resources (Michael et al., 20013), therefore reducing learning efficiency.

In the landscape of multimedia learning, the differentiation between "strong" and "weak" APA has emerged as a focal point of research (Krämer and Bente, 2010). A "strong" APA is distinguished by its rich array of emotional expressions, dynamic movements, and pronounced interactivity. Much research showed that such agents adeptly guide learners' attention and foster deeper engagement (Abdullah et al., 2017). Their animated behaviors not only captivate learners but also emphasize critical points in the content, facilitating better comprehension (Liew, Tan, and Kew, 2022; Romero-Hall, 2016). Conversely, "weak" APA are often motionless and emotionless, serving merely as visual placeholders. Their lack of dynamism and expressiveness has been shown to decrease learning performance (Craig, Gholson, and Driscoll, 2002), most probably by inadvertently drawing attention away from the learning content or through difficulties in eliciting sufficient engagement (Lester et al., 1997).

Therefore, the difference between strong and weak APA highlights the pivotal role of agent design in either augmenting or hindering the efficacy of multimedia content. Given that instructional programs already incorporate visual elements, introducing APA might affect both learners' engagement and attention (Walker, Sproull, & Subramani, 1996).

We note in the literature that, as yet there are no studies that focus on the impact of APA on learners' engagement depending on attentional capacities. In the broader field of multimedia learning, research exploring the impact of individual differences considers variability in learners' skills, aptitudes, preferences, cognitive characteristics such as working memory capacity, processing speed, visual attention span, and inhibitory control (Antonenko, Dawson, Cheng, & Wang, 2019). Effects of individual differences in cognitive processing have been recently explored in empirical research on levels of interactivity and selective attention (inhibitory control) as predictors of learning in multimedia chemistry simulations (Homer & Plass, 2014). Both studies compared two versions of a simulation, one that was interactive requiring learners to generate and test hypotheses, and the other - less interactive - based on the worked example approach (Paas & van Gog, 2006), in which learners went through a step-by-step procedure of actual hypothesis testing conducted by an expert. One of the studies reported by Homer and Plass (2014) revealed that, while selective attention differences in learners did not interact with comprehension (lower-level learning), selective attention and inhibitory control did interact with the more cognitively challenging transfer outcome. Specifically, students with lower levels of selective attention demonstrated better transfer outcomes in the worked simulation condition, whereas learners with higher levels of selective attention had better transfer outcomes in the exploratory multimedia simulation condition. These results suggest that in situations with higher levels of complexity, exploratory multimedia may not be suitable for all learners, specifically those with lower selective attention ability.

In this context, we conducted a study that aims to evaluate the impact of the APA's degree of animation on learners' motivation and engagement. We also analyzed the effects on learning outcomes of the APA's degree of animation on learners' attentional capacities, such as the ability to inhibit irrelevant information and the ability to allocate attentional resources. Participants had to study a specific content accompanied by a weak APA, a strong APA, and no APA at all. Their individual attentional capacities of inhibition and resource allocation were assessed beforehand, while their engagement and motivation related to the learning content were assessed at the end of the session. Finally, they all received questions assessing their memory and comprehension of the content. Our primary hypothesis is that the presence of a strong animated pedagogical agent significantly elevates levels of learners' engagement and motivation compared to scenarios where the agent is absent or weak (H1). Furthermore, we hypothesize that a strong animated pedagogical agent enhances learning performances (H2). Moreover, learners endowed with superior attentional capacities are likely to experience enhanced learning outcomes, regardless of the presence or sophistication of the pedagogical agent. This suggests that such individuals are equipped to benefit more effectively from the learning environment due to their ability to manage attentional resources. We thus hypothesize that the potential distractions presented by the strong pedagogical agent will be more disruptive for participants with lower attentional capacities than for those with higher attentional abilities (H3). This aspect of our research seeks to illuminate how individual differences in attentional resources can interact with the impact of the design of animated pedagogical agents to affect educational outcomes.

2. Methodology

2.1 Participants

Our study involved a total of 94 participants, recruited among psychology students or adults undergoing career changes. The average age of the participants was 28 years (SD = 9.5). There were 58 women and 38 men. All the participants were native French speakers. The participants were randomly assigned into the absent APA control group (19 women and 12 men; M= 28 years; SD = 9.6)), the weak APA group (21 women and 13 men; M= 27 years; SD = 7.1) and the strong APA group (16 women and 13 men; M= 29 years; SD = 11). In accordance with the ethical standards set forth in the Helsinki Declaration, this experiment was guided by a commitment to the well-being and dignity of all individuals involved. This declaration emphasizes the importance of obtaining informed consent, minimizing harm, and protecting the rights of all research participants. Participants in this study were fully informed about the experimental procedures. All participants were given an in-depth explanation of the experiment, including its purpose, procedures, risks, and benefits, prior to providing their written consent. The participants were also assured of the confidentiality of their data and their right to withdraw from the study at any time.

2.2. Material
2.2.1. E-learning Material and devices

The multimedia material consisted of an e-learning on immunotherapy. This e-learning allows us to understand the functioning of the immune system (innate and adaptive), how it reacts to cancer and presents immunotherapy. The support is in the form of 5 modules with a set of 30 short, animated videos with a voice-over. The content of this e-learning is divided as follows: Module 1: The Immune System; Module 2: Immune System and Cancer; Module 3: Immunotherapy; Module 4: Differences between Immunotherapy and Chemotherapy; Module 5: Learn more about Immunotherapy. Module 5 is facultative and provides learners with an additional resource. Videos are animated, and a voice explains the content, accompanied by text below the video presenting the content. The environment relies on a main interface that presents the different e-learning modules (see fig.1). The user can click directly on one of the blocks to access its content. Once in the module, there is a video, with a next button, a previous button and a back to menu button allowing the learner to navigate between videos (see fig.1)

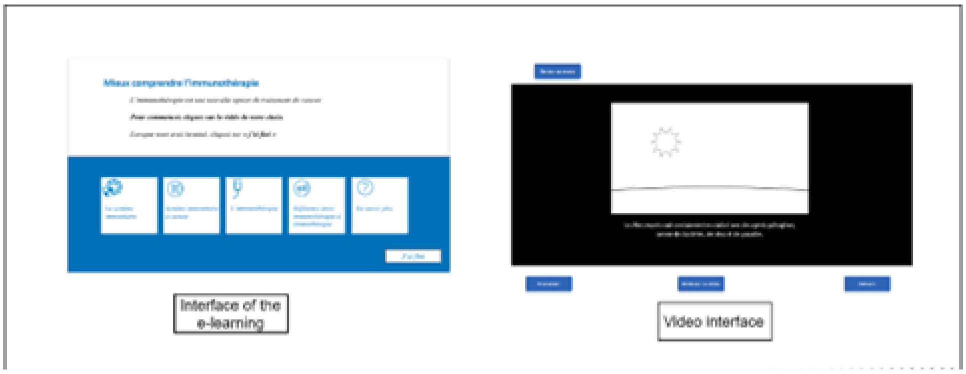

Figure 1: Learning materials. Interfaces of the e-learning environment: on the left the main interface to access the modules, on the right a screenshot of the interface to navigate in the video.

Three versions of this e-learning were developed, one without APA called Absent version (fig2), one with a weak APA (fig2), called Weak version and one with a strong APA called Strong version (fig2). The Absent version displays content via symbols that represent cells, and arrows with action verbs are used to present the links between elements. The symbols are meant to visually convey the different cells and their functions in the immune system. The weak APA version maintains the same content and format but incorporates a weak APA. This weak APA does not exhibit any movement or expressions, serving as a basic, non-interactive guide throughout the e-learning process. Its presence is consistent throughout the program, offering a stable point of reference for learners as they navigate through the different modules. The intention behind this version is to observe the impact of a minimal APA on the learning process. The Strong APA version goes a step further by incorporating a highly expressive APA. This agent exhibits movement and expressions, actively guiding learners through the e-learning program. The APA's actions, such as pointing to specific parts of the screen or displaying expressions that reflect the content, are carefully designed to align with the narration. This dynamic interaction is meant to engage the learners more deeply and to facilitate comprehension by bringing attention to key elements in the modules.

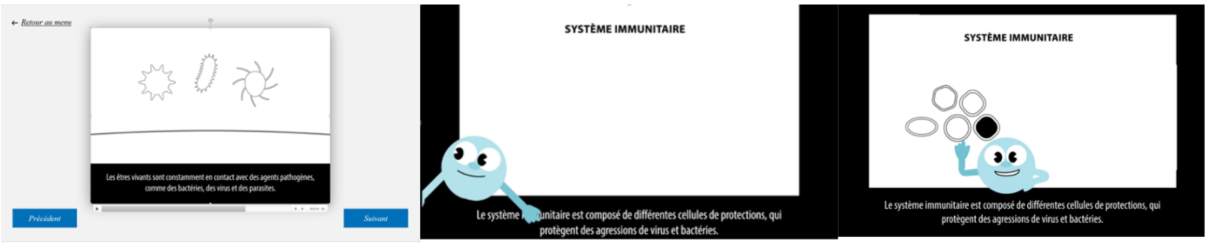

Figure 2: Learning materials. Screenshots of the three videos: the absent version on the left, the weak version on the middle and the strong version on the right.

The experiment was conducted on a Dell Latitude 5530 laptop computer with a screen size of 13.3 inches. The e-learning material was displayed in full-screen mode, and the video was presented within a virtual rectangle of 1280 x 720 pixels. The display resolution was set to 1366 x 768 pixels, and the video was played using the default media player of the computer. The computer was operated on battery power, and the ambient lighting conditions of the room were consistent throughout the experiment. The angle of the screen was set at 90 degrees to the table on which the laptop computer was placed, with the viewer positioned directly in front of the screen, at 30 cm. The orientation of the screen ensured that the video was displayed in its optimal aspect ratio, with no distortion or loss of detail due to the screen angle.

2.3 Measures

We have measured several control variables to ensure that they do not interfere in the experimental situation. Those variables are gender, age, interest in the subject and prior knowledge about it.

Interest in the subject. We included three home-made statements to assess the learners' interest in the subject. The participant must indicate on a Likert scale the degree of agreement (1= totally disagree; 5=totally agree) to each of the three statements presented ("I was already interested in immunotherapy before the course.";" This course was new to me"; "I knew about immunotherapy before seeing this course"). The score on the three statements were averaged to obtain an interest score. The higher the score is, the more the participant was interested in the topic before the course.

Prior Knowledge. We included one question to assess the participants' level of expertise on immunotherapy and scored on a scale going from novice to expert ("How do you rate your knowledge of immunotherapy?"). Prior knowledge has been shown to have a strong impact on the effectiveness of pedagogical design (Mayer, 2012).

Our independent variables are the attentional abilities and the type of pedagogical agent.

MAM – Attentional abilities. To assess participants' attentional capacities, we adopted a dual-task approach, as outlined in earlier studies (Michael, Dorey, et al., 2020; Michael et al., 2007; Michael, Pannetier, et al., 2014, Salgues, S et al, 2023). In the MAM test, the primary task involved a tailored computerized visual search requiring participants to find a predefined target and inhibit distractors. Simultaneously, an auditory detection task was presented, challenging participants to adeptly manage cognitive resources across both tasks. To heighten the challenge, the auditory task's difficulty was varied, requiring participants to shift their cognitive resources from the visual task. These two components offer a comprehensive, multi-faceted evaluation of attention, specifically targeting inhibitory control and the management of attentional resources. We follow the methodology used in recent studies (Michael, Dorey, et al., 2020; Salgues S et al 2023), and described in more detail in Appendix 1.

Our dependent variables are learners' learning performances, motivation and engagement. They are measured using different types of methods.

Learning performance. We used a set of 10 comprehension and 10 memorization questions to evaluate learning performance after the task. Comprehension and memorization questions are two different types of assessment methods for learning performance. Comprehension questions assess the learner's understanding of concepts, while memorization questions assess the ability to recall specific details. Participants were required to indicate whether the statement presented was true or false. An example of a memorization question is "The tears and the lymphocytes are the initial protections against the intrusion of bacteria." An example of a comprehension question is "The mechanism allowing lymphocyte education is slower than the innate immune response". Half of the questions required a “correct” answer whereas the other half required an “incorrect” answer. Performance was reflected in the proportion of correct responses, but also in two signal detection indices[3], *d’* which reflects the ability to distinguish true from false statements, and the *C* representing the response tendencies of participants towards “correct” or “incorrect” responses independently of whether the responses are correct or not.

Motivation – Situational Motivation Scale (SIMS). We used the Situational Motivation Scale (SIMS) developed by Guay et al. (2000) and translated and validated in French by Fontaine et al. (2019). This questionnaire consists of 16 statements which allow to evaluate four types of learners’ motivation: intrinsic motivation ("because I think that this activity is interesting," 4 items), identified regulation ("because I am doing it for my own good," 4 items), external regulation ("because I am supposed to do it," 4 items), and amotivation ("there may be good reasons to do this activity, but personally, I don't see any," 4 items). Participants responded on a Likert scale spanning from 0 (Disagree) to 7 (Totally agree). The higher the score of the first three types of motivation, the more likely learners are motivated. On the contrary, a high score in amotivation shows that learners are not willing to continue interacting with the learning material.

Engagement – User Engagement Scale (UES). All participants filled out the short form of the User Engagement Scale (UES) (O’Brien & Toms, 2010) translated and validated in French by Fontaine et al. (2019). They indicate on a Likert scale the degree of agreement (1 = totally disagree; 5 = totally agree) to each statement presented. The scale is composed of 4 dimensions, each one composed of 3 items: Focused attention (feeling of absorption in the interaction and loose of track of time, Aesthetic appeal (attractiveness and visual appeal of the interface),

[3] The d' and C indexes are derived from the z-transformed hit and false alarm rates. Specifically, d' is computed as the difference between the z-transformed hit rate and the z-transformed false alarm rate, representing sensitivity to signal. Larger values of d' reflect greater sensitivity, i.e., an enhanced ability to discriminate signal from noise. The C index, or criterion, is calculated as the average of the z-transformed hit and false alarm rates, but with a negative sign applied to the false alarm rate. This reflects a response bias, with positive values indicating a bias towards claiming 'no signal' (or favoring noise responses) and negative values indicating a bias towards claiming 'signal' (or favoring signal responses).

Perceived usability (negative affect experienced because of the interaction and the degree of control and effort), Reward (amusement, novelty, interest, and overall satisfaction of the subject). The scores of the Perceived usability dimension items are reversed as recommended by O'Brien and Cairns (2018). The higher the score is, the more the subject is engaged.

Engagement - Subjective Perception of Time (SPT). Following the recommendations outlined by Hancock (2019), we employed a retrospective paradigm in which learners were asked at the end of the course how much time they believed they spent navigating in the e-learning modules, without consulting their phones or other devices. This perceived time was then divided by the actual time spent in the e-learning, with a higher score indicating an overestimation of the time passed. It is inferred that learners who overestimate the time spent are likely experiencing boredom, while those who underestimate are more engaged and absorbed in the learning activity.

Engagement - Learning Analytics. We used learning analytics to gather data on the learners' navigation patterns. This information was collected through tracking learners' interactions with the learning materials within the course, including the time they spent on each of the 5 modules, and the number of modules they revisited. Moreover, we collected information about the click on module 5, which was not mandatory and thus considered as an indicator of interest.

2.3. Procedure

The experiment was conducted under standard laboratory conditions. Initially, participants were randomly assigned to one of the three conditions (absent, weak APA and strong APA). All the participants carried out the MAM task (10/15min). Then they filled out a first questionnaire to measure the level of expertise and initial interest in immunotherapy. Then the participants consulted the e-learning on immunotherapy for a total minimum duration of 15 minutes. During the experience, the interaction traces were collected. At the end, participants completed all measures of interest described above. Finally, the questions for measuring comprehension and memory were intermixed and randomly presented to each participant. The experimental sessions lasted about 30 minutes. Participants were instructed to engage with the e-learning material as if they were studying for an exam and to answer the comprehension and memory questions to the best of their abilities.

3. Results

3.1. Learners' characteristics

There were no significant age differences among participants assigned to each version (Absent $M = 28.1$, $SD = 8.3$; Weak $M = 25.3$ $SD = 6.5$; Strong $M = 29.3$ $SD = 13.9$; $F(1,94) = 0.484$, $p = 0.6$). Additionally, the distribution of men and women across groups was not significantly different (Absent H = 19, F = 12; Weak H = 21, W = 13; Strong M = 16, F = 13; $\chi^2(1) = 0.339$, $p > .05$). Regarding participants' self-reported expertise of the content, no significant group effects were observed (Absent $M = 1.1$, $SD = 0.37$; Weak $M = 1.14$, $SD = 0.35$; Strong $M = 1.0$, $SD = 0.25$; $F(2,94) = 0.651$, $p = .52$). Similarly, there were no significant differences in participants' interest in the topic across version (Absent $M = 2.39$, $SD = 0.74$; Weak $M = 2.18$, $SD = 0.65$; Strong $M = 2.24$, $SD = 0.74$; $F(2,94) = 0.758$, $p = .47$).

3.2. Effect of the Support Version on engagement and motivation (H1)

We carried out a series of analyses to explore the differences in scores across various variables that were examined through a series of ANOVAs comparing groups based on the type of APA version received (Absent, Weak and Strong). Cohen's *d* and η^2p were used for the size effect.

Engagement

UES - A notable difference between the three groups was identified in the average UES score of the four factors of the UES. ($F(2,94) = 3.58$, $p < .05$, $\eta^2p = .07$). Bonferroni's post hoc test revealed that participants with the strong APA declared a higher engagement (Strong $M = 3.44$, $SD = 0.52$) compared to participants without any APA (Absent $M = 3.02$, $SD = .57$). Taking each factor separately, this significant difference was found to be consistent across three of the four factors: Focused Attention ($F(2,94) = 3.56$, $p < 0.05$, $\eta^2p = 0.07$; Absent $M = 2.48$, $SD = 0.85$; Strong $M = 3.15$, $SD = .97$), Perceived Usability ($F(2,94) = 4.34$, $p < 0.05$, $\eta^2p = 0.08$; Absent $M = 3.1$, $SD = 0.89$; Strong M_ = 3.74, $SD = 0.90$), and Visual Appeal ($F(2,94) = 3.56$, $p < 0.05$, $\eta^2p = 0.07$; Absent $M = 2.48$, $SD = 0.85$; Strong $M = 3.14$, $SD = .97$). No significant differences were observed between absent and weak APA conditions, neither between the weak and strong APA conditions. No significant effect was observed for the Reward dimension.

SPT - A main effect of the condition on the subjective perception of time was found, $F(2,94) = 8.85$, $p < 0.001$, $\eta p^2 = 0.16$. The Bonferroni post hoc test indicated that participants underestimated the time spent in the Strong APA ($M = 0.9$, $SD = 0.22$) and the Weak APA ($M = 0.90$, $SD = 0.19$) conditions compared to the Absent APA condition ($M = 1.1$, $SD = .24$).

Learning Analytics - We observed a main effect of the APA version on the total time spent on e-learning ($F(2,94) = 6.85$, $p < .01$, $\eta^2 p = .131$). The Bonferroni post hoc test indicated that participants spent more time in the Strong APA version ($M = 25.27$, $SD = 2.72$) than in the Absent APA version ($M = 23.41$, $SD = 2.40$). Delving deeper into which part of the e-learning was observed more, we found differences in module 3 ($F(2,94) = 3.17$, $p < .05$, $\eta^2 = .06$; Absent $M = 5.41$, $SD = 1.36$; Weak $M = 6.14$, $SD = 1.45$; Strong $M = 6.24$, $F = 1.38$), but no differences were identified through the post hoc tests.

Motivation - The effect of the version on learners' motivation was significant only for external regulation ($F(2,94) = 5.44$, $p < 0.01$, $\eta^2 p = 0.107$): participants declared less extrinsic motivation in the Strong APA version ($M = 3.15$, $SD = 1.10$) than in the Absent APA version ($M = 4.04$, $SD = 0.89$). No other differences were observed.

UES and Subjective Perception of Time. A negative correlation ($r = -0.254$, $p = 0.013$) was found between the SPT and the overall UES score. This suggests that as learners' engagement increases, their perception of time spent on the e-learning platform tends to be underestimated. There was a notable negative correlation between the subjective perception of time and the UES score for Focused Attention ($r = -0.334$, $p < .001$).

3.3 The combined effect of pedagogical agent and attentional abilities on learning outcomes (H2/H3)

In order to assess whether attention abilities combine with the APA version to determine performance outcome, we conducted a mixed-design ANOVA on the proportion of correct responses to the questions participants received about the e-learning content. The question type (comprehension vs. memorization) was used as the within-participants factor, the APA version

(Absent, Weak and Strong) as the between-subjects factor, and the attentional abilities (RD and ID) as covariates and moderators.

Proportion of correct responses

The analysis revealed a main effect of the question type on the proportion of correct responses ($F(1,94) = 14.94$, $p < .001$, $\eta^2 p = 0.14$). Participants achieved better results in memorization ($M = 0.67$, $SD = 0.10$) than in comprehension ($M = 0.62$, $SD = 0.10$). We also observed a main effect of the APA version on performance ($F(2,94) = 5.69$, $p < .01$, $\eta^2 = 0.11$). Bonferroni's post hoc tests showed that participants in the weak APA version performed less well ($M = 0.61$, $SD = 0.08$) than those in the absent APA version ($M = 0.67$, $SD = 0.05$) and the strong APA version ($M = 0.65$, $SD = 0.09$). No difference was found between the absent and strong APA versions. We also identified an interaction effect between the question type and APA version ($F(2,94) = 6.99$, $p < .001$, $\eta^2 = 0.13$). Bonferroni's post hoc tests revealed that participants in Absent APA version scored higher in memorization ($M = 0.72$, $SD = .07$) than in comprehension ($M = 0.62$, $SD = .08$). A similar effect was found in the weak APA version ($F(2,94) = 6.99$, $p < .001$, $\eta^2 = 0.13$, Memorization $M= 0.72$; $SD= .07$; comprehension $M = 0.62$, $SD = .08$). No differences were observed between comprehension and memorization in the Strong APA version (Memorization $M= 0.64$, $SD = .10$; Comprehension $M = 0.66$, $SD = .11$). When inspecting memorization performance, lower scores were observed between the Absent and the Weak APA conditions, as well as between the Absent and the Strong APA. As far as comprehension is concerned, we observed lower performance in the Weak APA version ($M = 0.58$, $SD = 0.11$) as compared to the Strong APA version ($M = 0.66$, $SD = 0.11$).

Signal detection theory indices.

The analysis of the sensitivity index *d'* showed similar main effects as those reported above. We found a main effect of the question type ($F(1,94) = 5.65$, $p < .05$, $\eta^2 p = 0.05$). The average *d'* was higher in memorization ($M = 0.90$, $SD = 0.54$) than in comprehension ($M = 0.69$, $SD = 0.77$). A main effect of the APA version was observed ($F(2,94) = 4.96$, $p < .01$, $\eta^2 p = 0.09$) with lower performance for the weak version compared to the two other versions (Absent $M = 0.89$, $SD = 0.29$; Weak $M = 0.58$, $SD = 0.47$; Strong $M = 0.94$, $SD = 0.68$). We also observed an interaction effect between the APA version and the question type ($F(2,94) = 13.86$, $p < .001$, $\eta^2 p = 0.234$). Bonferroni's post hoc tests show a significant difference between memorization and comprehension for the absent version (Memorization $M = 1.18$, $SD = 0.38$;

Comprehension M = 0.61, SD = 0.45) and weak version (Memorization M = 0.76, SD = 0.51; Comprehension M = 0.39, SD = 0.45). When inspecting memorization, performance did not differ between APA versions (Absent M= 1.18; SD=.38; Weak M=.76; SD= .51; Strong M=.75; SD=.61). Concerning comprehension, we observed a higher performance for the strong APA condition compared to the absent APA condition and weak APA condition (Absent M = 0.61, SD = 0.45; Weak M = 0.39, SD = 0.61; Strong M = 1.13, SD = 1).

The analysis of the C index revealed a main effect of the question type ($F(1,94) = 21.64$, $p < .001$, $\eta^2p = 0.19$). Participants had a higher C value for memorization (M = 0.06, SD = 0.372) than for comprehension (M = -0.18, SD = 0.40). We also observed a main effect of the APA version ($F(2,94) = 5.71$, $p < .01$, $\eta^2p = 0.11$). Bonferroni's post hoc tests revealed a higher C index in absent version (M = 0.06, SD = 0.28) compared to the Strong APA version (M = -0.17, SD = 0.34). No interaction effect was observed.

Role of attentional abilities

We conducted several tests to examine the validity of the MAM task and to obtain two indices that we use as indicators of attentional abilities. We obtained the Inhibition Difficulty (ID) index, representing the learner's difficulties in inhibiting distractors, and the Resource Management Difficulty (RD) index, representing the difficulty to manage attentional resources by sharing them between the visual and the auditory components of the MAM task (as discussed in Michael, Dorey, et al., 2020). For details about the methods used to calculate the two indices and the psychometrics properties of the MAM task, see Appendix 1.

Learning outcome: Proportion of correct responses

We found no main effect of index type and no interaction between index type and APA version on the mean proportion of correct responses. We observed an interaction effect between the type of question and the ID score ($F(1,94) = 6.03$, $p < .05$, $\eta^2p = 0.69$) and an interaction effect between the type of question and the RD score ($F(1,94) = 5.56$, $p < .05$, $\eta^2p = 0.06$). Upon further analysis of these interactions through regression analyses, we found that the ID index negatively predicted memory performance ($\beta=-.029$, SEM = 0.013, $t(92) = -2.195$, $p=.031$, $R^2=.05$), but not comprehension performance ($\beta=.011$, SEM = .016, $t(92) = .71$, $p=.47$, $R^2=.006$). We also found that the RD index positively predicted memory performance ($\beta=.041$, SEM = .016, $t(92) = 2.57$, $p=.012$, $R^2=.07$), but not comprehension performance ($\beta=-.003$, SEM = .019, $t(92) = -.17$, $p=.86$, $R^2=0$).

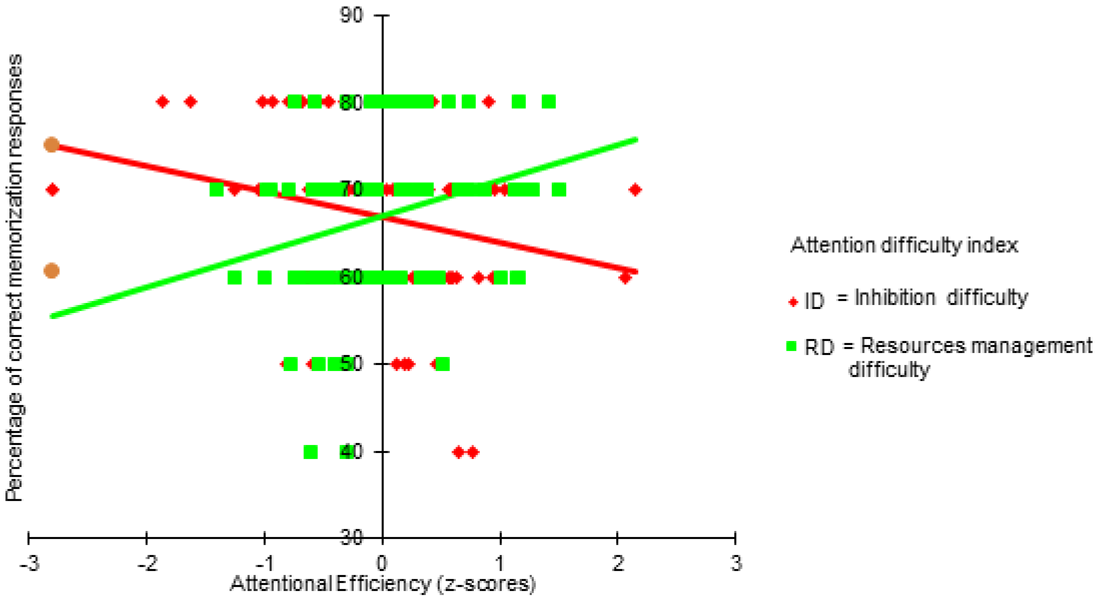

Figure 3: Interaction between ID and memorization, and RD and memorization. The higher the index, the less the subject has good associated abilities.

d' index

No main effect of attention index type and no interaction between index type and APA version were found. We observed an interaction between the type of question and the ID index (F (1,94) = 4.42, $p < .05$, $\eta^2p = 0.05$). Upon further analysis of this interaction through a regression analysis, we found that the ID index negatively predicted d' for memorization ((β=−.016, SEM = 0.07, t (92) =−2.22, p=.02, R^2=.05), but not comprehension performance (t=0.71, p=0.47).

C index

Regarding the C index, our analysis did not yield any significant results.4.

4. Discussion

In this study, we analyzed the impact of pedagogical agents, and their version (weak or strong) on learners' engagement, motivation, comprehension and memorization. Furthermore, we investigated the effect of attentional capacities on memorization and comprehension performance. Our findings unveil a complex picture, where the presence and characteristics of pedagogical agents can either hinder or facilitate different aspects of the learners' experience.

4.1 Motivational and engaged outcomes.

In agreement with the existing literature (Davis, 2018), we found that a strong APA led to higher learners' engagement, compared to the two other APA versions where the agent was absent or weak. This suggests that when learners interact with a strong pedagogical agent, they become absorbed in the interaction, losing track of time, and find the interface more attractive and visually appealing (O'Brien et Cairns 2018). Interestingly, this change in engagement did not extend to the reward dimension, suggesting that it was not driven by extrinsic rewards. In addition to these findings, the study also sheds light on the subjective perception of time in relation to the presence of pedagogical agents, adding another dimension to our understanding of learners' engagement (Hancock, 2019). Participants in conditions with strong and weak pedagogical agents were found to underestimate the time spent in the learning activity compared to those in the absent agent condition. We observed that the subjective perception of time correlates with the subjective measure of engagement. It may indicate that a strong engagement, particularly in terms of focused attention, contributes to creating an immersive learning experience for learners. It is coherent with the positive correlation between the subjective perception of time and the engagement score, showing that engaged learners underestimate the time spent in the e-learning. The data suggest that the combination of several measures of learners' engagement could extends beyond traditional measures of the impact of APAs, analyzing not only on the usability and aesthetic aspects of the multimedia learning environment, but also how it may affect the learner's psychological state, perception of the learning process and effective duration of the experience.

Regarding learners' motivation, we expected an effect of the APA on intrinsic motivation, in line with previous studies (Liew, Zin, & Sahari, 2017, Lang, Xie, Gong, Wang, & Cao, 2022). Unfortunately, we did not observe such an effect, which suggests that the manipulation of APA features was not more valued by learners, nor perceived as more interesting and enjoyable (Deci et Ryan 1991; Gay, 2008). The literature is not unequivocal as far as the effect of APAs on intrinsic motivation is concerned. Indeed, the study conducted by Schroeder and Adesope (2013) found that the presence of a pedagogical agent does not have any significant effect on learning or affective outcomes, including motivation (Schroeder &

Adesope, 2013). Additionally, Heidig and Clarebout (2011) argue that while pedagogical agents may facilitate learner motivation and learning, their effectiveness and design remain unclear due to a lack of control group (Heidig & Clarebout, 2011). In our study, we have included a control group (absent version), but we still have not observed an effect on intrinsic motivation.

We could explain this result by the fact that, in our study, the animations and visuals of the APA may not have been sufficiently varied, influencing learners' perceptions. Indeed, some studies hypotheses that the complexity of results when measuring agent persona might be attributed to the quality of agents employed (Veletsianos & Russell, 2014) or to the contextual relevance of the agents (Veletsianos, 2007). The visual aspect of the APA, whether in terms of realism or a drawn avatar, could play a crucial role. Realistic APAs may be more engaging than simplistic or cartoon avatars, thereby affecting intrinsic motivation (Gulz & Haake, 2006). It is also possible that the limited role of the APA, without interactive features such as navigation, reduced its impact, unlike other studies where APAs had more dynamic and interactive roles (Lester & Stone, 1997). In addition, the duration of the e-learning experience could have influenced the effectiveness of APA (30 min in average in this study). Longer exposure to APA might have been too long to motivate learners as suggested in a recent meta-analysis from Davis, 2018 who found that APA had a larger weighted effect size with videos under 6 min.

Finally, and contrary to our hypothesis, we observed that the presence of strong pedagogical agents was associated with a lower degree of external regulation. In the context of Self-Determination Theory (Deci et Ryan, 1990), external regulation refers to behaviors influenced by external factors such as rewards, punishments, or other forms of external control. This outcome might suggest that learners felt less constrained in their learning experience. One possible explanation could be that, in the presence of Animated Pedagogical Agents (APAs), students might feel more engaged. In fact, the reduction in perceived constraint might be linked to the concept of flow and engagement. When learners are engaged or in a state of flow, they are absorbed in the activity to the extent that they are less aware of external pressures or constraints (Csikszentmihalyi, 1990). This could mean that the presence of APAs, by fostering a more engaging and interactive learning environment, indirectly reduces the weight of extrinsic constraints felt by the learners. Furthermore, this result can be interpreted as indicating a shift in the social nature of the learning experience. In the experiment, without Animated Pedagogical Agents (APAs), learners engage with the educational content alone, where the only authoritative or guiding presence is the experimenter. Veletsianos, Miller, and Doering (2009) have observed that “interactions between humans and computers [are] expected to approximate social interactions between humans and humans” (p. 174). This perspective is in line with our findings; the APAs seem to provide social cues that diminish the overt external influence of the experimenter. Learners might perceive the animated pedagogical agent as the primary authority in the learning experience, which could explain the reduced sense of external regulation.

4.2 Cognitive outcomes

We observed lower performance in memorization associated with the presence of pedagogical agents. This resonates with Mayer's Cognitive Theory of Multimedia Learning (CTML), which postulates that the human cognitive system has limited processing capacity (Mayer, 2009) and that introduction of pedagogical agents, regardless of their sophistication, could introduce extraneous cognitive load. This would impede the encoding of information into long-term memory leading to decrease performance. This aligns with prior research suggesting that extraneous details, even when related to the content, can detract from essential processing (Sweller, 1994). According to this theory, the additional visual and auditory information provided by an APA might increase the extraneous cognitive load, especially if this information does not directly support the learning objectives or if it competes for the learners' attention with the primary material (Sweller, 1994). Specifically, in the context of memorization, where the task is to recall information verbatim, the presence of additional stimuli from the APA might lead to a 'split-attention' effect (Park, S.,2015). Learners may find themselves dividing their cognitive resources between the content and the APA's cues, which could hamper the process of memorization. Moreover, our findings align with the redundancy principle, which suggests that presenting the same information in multiple forms (e.g., text and narration) can be detrimental if learners have sufficient cognitive capacity to understand the material from a single source (Mayer & Fiorella, 2014). This idea seems confirmed by the observation that higher difficulties in inhibition as measured through the attentional dual task were associated with poorer memory outcomes (Hasher & Zacks, 1988). Moreover, the difficulties in managing attentional resources by considering multiple sources of information seem to positively impact memory performance. In other words, rather than efficiently dividing their attention across multiple stimuli or tasks, participants may focus more intensively on a single task or piece of information at a time that enhance their ability to encode and retrieve information, leading to better memory performance. These two effects of attentional abilities on memorization performance are coherent, as focusing attention and reducing distractions are known to enhance encoding and retrieval of information on memory (Eagle & Ortof, 1967).

A strong APA appears to bolster overall comprehension by facilitating a deeper engagement with the content, encouraging learners to process information at deep cognitive levels. The non-verbal cues and the presence of a well-designed APA direct learners to important elements and support the construction of mental models that are essential for deep comprehension (Mayer, 2021; Moon and Ryu, 2021). We did not observe an effect of attentional abilities on comprehension, this result may reflect that the APA's effect might be more related to information retrieval rather than encoding. It indicates that APAs help consolidate learning by making retrieval processes more efficient, rather than enhancing the initial absorption of information. We did not observe an interaction effect between the type of pedagogical agent and the attentional capacities on performance outcomes, but we observed the same coherent effect of these variables on memorization performance. In one hand, the strong APA with animation, as we could see it as a distractor, is associated to a lower memorization, and on another hand, the learner's ability to focus and inhibit distractor was associated with higher memorization. In contrast to memorization, comprehension is only positively impacted by the APA's degree of animation. This implies that the potential distraction induced by the presence of the strong APA did not impact comprehension. It's coherent with Kintsch's (1996) model that indicates deeper levels of processing are less susceptible to disruption.

To conclude, our findings suggest a nuanced landscape where the presence and characteristics of APAs can variously impede or enhance different facets of the learning experience. In terms of motivational and engagement outcomes, our findings partially align with existing literature, highlighting that a strong APA enhances learner engagement, leading to a more immersive and visually appealing learning environment. However, this increased engagement did not translate into enhanced intrinsic motivation, but instead led to a perception of learning experience less constrained. Concerning cognitive outcomes, the presence of APAs, regardless of their sophistication and learner's attentional abilities, were associated with reduced memorization performance. This outcome resonates with theories proposing limited cognitive capacity and the potential for extraneous cognitive load introduced by APAs. While APAs can enhance certain aspects of the learning experience, their impact on cognitive outcomes is complex and multifaceted. The results underscore the importance of considering both the design features of APAs and individual learners' differences when assessing their efficacy in educational settings. Furthermore, given the differential effects on comprehension and memorization, the deployment of APAs might be more beneficial for learning objectives focused on understanding concepts rather than for tasks requiring verbatim recall, such as memorizing laws.

One of the primary limitations of our study lies in the generalizability of the findings. The experimental setup and the characteristics of the pedagogical agents used may not fully encapsulate the diversity of educational contexts and digital learning environments. Furthermore, the study focuses on short-term engagement and cognitive outcomes, and does not address the long-term retention of information or the sustained motivational impact of APAs. The sample size and demographic characteristics of participants may also limit the applicability of our results across different age groups, educational levels, and cultural backgrounds. Additionally, the reliance on self-reported measures for engagement and motivation can introduce biases, as these perceptions may not accurately reflect actual learning outcomes or behaviors. Another notable aspect that warrants attention is the exploration of the direct impact of attentional capacities on APA design effectiveness. Our study observed parallel effects of attentional abilities and APA design on learning outcomes without establishing a direct correlation between the two. This gap suggests a potential oversight in assessing how individual differences in attentional capacities could directly influence the efficacy of different APA designs. The assumption that all learners would respond similarly to APA features may not hold true across diverse attentional profiles, highlighting a limitation in the current research approach.

Future works should explicitly investigate the direct relationship between learners' attentional capacities and the effectiveness of APA design. This could involve varying the complexity and interactivity of APAs and examining how these variations may impact learners with different attentional abilities. For instance, APAs with more subtle cues and interactions may be more effective for learners with higher distractibility, whereas those with stronger attentional control may benefit from more complex and engaging APA features.

5.References

Abdullah, A., Adil, M., Rosenbaum, L., Clemmons, M., Shah, M., Abrahamson, D., & Neff, M. (2017). *Pedagogical agents to support embodied, discovery-based learning*. 1–14.

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. *The Cambridge Handbook of Multimedia Learning*, *2*, 135–146.

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, *25*(2), 450–457.

Baylor, A. L., & Kim, Y. (2005). Simulating Instructional Roles through Pedagogical Agents. *International Journal of Artificial Intelligence in Education*, *15*(2), 95–115.

Borokhovski, E., Bernard, R. M., Tamim, R. M., Schmid, R. F., & Sokolovskaya, A. (2016). Technology-supported student interaction in post-secondary education: A meta-analysis of designed versus contextual treatments. *Computers & Education*, *96*, 15–28. https://doi.org/10.1016/j.compedu.2015.11.004

Chin, K.-Y., Hong, Z.-W., Huang, Y.-M., Shen, W.-W., & Lin, J.-M. (2016). Courseware development with animated pedagogical agents in learning system to improve learning motivation. *Interactive Learning Environments*, *24*(3), 360–381. https://doi.org/10.1080/10494820.2013.851089

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*(2), 428.

Csikszentmihalyi, M. (2000). *FLOW: The Psychology of Optimal Experience*.

Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review*, *24*, 193–209. https://doi.org/10.1016/j.edurev.2018.05.002

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and Education: The Self-Determination Perspective. *Educational Psychologist*, *26*(3–4), 325–346. https://doi.org/10.1080/00461520.1991.9653137

D'Mello, S., Jackson, T., Craig, S., Morgan, B., & Chip, P. (n.d.). *AutoTutor Detects and Responds to Learners Affective and Cognitive States*.

Eagle, M., & Ortof, E. (1967). The effect of level of attention upon "phonetic" recognition errors. *Journal of Verbal Learning and Verbal Behavior*, *6*(2), 226–231. https://doi.org/10.1016/S0022-5371(67)80101-4

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, *39*(2), 75–85. https://doi.org/10.1016/j.cedpsych.2014.01.001

Fontaine, G., Cossette, S., Maheu-Cadotte, M.-A., Mailhot, T., Lavoie, P., Gagnon, M.-P., Dubé, V., & Côté, J. (2019). Traduction, adaptation et évaluation psychométrique préliminaire d'une mesure d'engagement et d'une mesure de charge cognitive en contexte d'apprentissage numérique. *Pédagogie Médicale*, *20*(2), 79–90.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation : The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*, 175–213.

Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents—A look at their look. *International Journal of Human-Computer Studies*, *64*(4), 322–339. https://doi.org/10.1016/j.ijhcs.2005.08.006

Hancock, P. A., Kaplan, A. D., Cruit, J. K., Hancock, G. M., MacArthur, K. R., & Szalma, J. L. (2019). A meta-analysis of flow effects and the perception of time. *Acta Psychologica*, *198*, 102836. https://doi.org/10.1016/j.actpsy.2019.04.007

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, *6*(1), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004

Johnson, C. I., & Mayer, R. E. (2009). A testing effect with multimedia learning. *Journal of Educational Psychology*, *101*(3), 621–629. https://doi.org/10.1037/a0015183

Krämer, N. C., & Bente, G. (2010). Personalizing e-learning. The social effects of pedagogical agents. *Educational Psychology Review*, *22*, 71–87.

Lang, Y., Xie, K., Gong, S., Wang, Y., & Cao, Y. (2022). The Impact of Emotional Feedback and Elaborated Feedback of a Pedagogical Agent on Multimedia Learning. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.810194

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: Affective impact of animated pedagogical agents. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 359–366. https://doi.org/10.1145/258549.258797

Li, J., Antonenko, P. D., & Wang, J. (2019). Trends and issues in multimedia learning research in 1996–2016: A bibliometric analysis. *Educational Research Review*, *28*, 100282.

Li, W., Wang, F., & Mayer, R. E. (2023). How to guide learners' processing of multimedia lessons with pedagogical agents. *Learning and Instruction*, *84*, 101729.

Liew, T. W., Mat Zin, N. A., & Sahari, N. (2017). Exploring the affective, motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment. *Human-Centric Computing and Information Sciences*, *7*(1), 9. https://doi.org/10.1186/s13673-017-0089-2

Liew, T. W., Tan, S.-M., & Kew, S. N. (2022). Can an angry pedagogical agent enhance mental effort and learning performance in a multimedia learning environment? *Information and Learning Sciences*, *123*(9/10), 555–576.

Martha, A. S. D., & Santoso, H. B. (2019). The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. *Journal of Educators Online*, *16*(1). https://eric.ed.gov/?id=EJ1204376

Mayer, R. E. (2014). *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*(3), 247–288. https://doi.org/10.1080/01638539609544975

Michael, G. A., Dorey, J.-M., Rey, R., d'Amato, T., Fabre, D., Brunet, S., & Padovan, C. (2020). Attention in schizophrenia: Impaired inhibitory control, faulty attentional resources, or both? *Psychiatry Research*, *290*, 113164.

Michael, G. A., Lété, B., & Ducrot, S. (2013). Trajectories of attentional development: An exploration with the master activation map model. *Developmental Psychology*, *49*(4), 615.

Michael, G., Pannetier, C., Chastain, V., Couffe, C., Padovan, C., Fabre, D., Dorey, J., & Neuschwander, P. (2014). Transnosographie des processus attentionnels basée sur le modèle Master Activation Map : Étude préliminaire d'un continuum neuropsychiatrique. *Kinésithérapie Scientifique*, *550*, 41–45.

Moon, J., & Ryu, J. (2021). The effects of social and cognitive cues on learning comprehension, eye-gaze pattern, and cognitive load in video instruction. *Journal of Computing in Higher Education*, *33*(1), 39–63. https://doi.org/10.1007/s12528-020-09255-x

Moreno, R. (2005). Multimedia Learning with Animated Pedagogical Agents. In *The Cambridge handbook of multimedia learning* (pp. 507–523). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.032

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, *112*, 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, *61*(1), 50–69. https://doi.org/10.1002/asi.21229

O'Keefe, P. A., Letourneau, S. M., Homer, B. D., Schwartz, R. N., & Plass, J. L. (2014). Learning from multiple representations: An examination of fixation patterns in a science simulation. *Computers in Human Behavior*, *35*, 234–242.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology*, *21*(4), 325–344.

Park, I., Kim, D., Oh, J., Jang, Y., & Lim, K. (2015). Learning effects of pedagogical robots with programming in elementary school environments in Korea. *Indian Journal of Science and Technology*, *8*(26), 1–5.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160.

Romero-Hall, E., Watson, G. S., Adcock, A., Bliss, J., & Adams Tufts, K. (2016). Simulated environments with animated agents: Effects on visual attention, emotion, performance, and perception. *Journal of Computer Assisted Learning*, *32*(4), 360–373. https://doi.org/10.1111/jcal.12138

Rostaminezhad, M. A., Mozayani, N., Norozi, D., & Iziy, M. (2013). Factors Related to E-learner Dropout: Case Study of IUST Elearning Center. *Procedia - Social and Behavioral Sciences*, *83*, 522–527. https://doi.org/10.1016/j.sbspro.2013.06.100
Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020
Salgues, S., Plancher, G., & Michael, G. A. (2023). Attention and body awareness: The role of inhibition and the management of cognitive resources in the perception of spontaneous sensations. *Psychology of Consciousness: Theory, Research, and Practice*.
Schroeder, N. L., & Adesope, O. O. (2014a). A Systematic Review of Pedagogical Agents' Persona, Motivation, and Cognitive Load Implications for Learners. *Journal of Research on Technology in Education*, *46*(3), 229–251. https://doi.org/10.1080/15391523.2014.888265
Schroeder, N. L., & Adesope, O. O. (2014b). A Systematic Review of Pedagogical Agents' Persona, Motivation, and Cognitive Load Implications for Learners. *Journal of Research on Technology in Education*. https://www.tandfonline.com/doi/abs/10.1080/15391523.2014.888265
Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013a). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*(1), 1–39. https://doi.org/10.2190/EC.49.1.a
Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013b). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*(1), 1–39. https://doi.org/10.2190/EC.49.1.a
Sproull, L., Subramani, M., Kiesler, S., Walker, J. H., & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction*, *11*(2), 97–124.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.
Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5
van der Meij, H., Van der Meij, J., & Harmsen, R. (2015). Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment. *Educational Technology Research and Development*, *63*, 381–403.
van Mulken, S., André, E., & Müller, J. (1998). The Persona Effect: How Substantial Is It? In H. Johnson, L. Nigay, & C. Roast (Eds.), *People and Computers XIII* (pp. 53–66). Springer. https://doi.org/10.1007/978-1-4471-3605-7_4
Veletsianos, G. (2007). Cognitive and Affective Benefits of an Animated Pedagogical Agent: Considering Contextual Relevance and Aesthetics. *Journal of Educational Computing Research*, *36*(4), 373–377. https://doi.org/10.2190/T543-742X-033L-9877
Veletsianos, G., & Russell, G. S. (2013). What Do Learners and Pedagogical Agents Discuss When Given Opportunities for Open-Ended Dialogue? *Journal of Educational Computing Research*, *48*(3), 381–401. https://doi.org/10.2190/EC.48.3.e
Veletsianos, G., & Russell, G. S. (2014). Pedagogical Agents. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 759–769). Springer. https://doi.org/10.1007/978-1-4614-3185-5_61
Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology*, *110*(2), 250.
Wang, Q. (2019). *Research on the Role of Teachers in Traditional Teaching Mode and "MOOC Plus Flipped Classroom" Teaching Mode**. 1018–1022. https://doi.org/10.2991/iccese-19.2019.224
Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.
Yusoff, M. D. M., Yusof, A. M., & Jamaludin, N. S. A. (2022). *Pedagogical Agent Element and Design for Learning: A Systematic Literature Review*. http://myscholar.umk.edu.my/handle/123456789/4130

Appendix 1 Description of the MAM task

Visual Detection Task: Each trial began with a central fixation point displayed for one second. This was followed by the search display where participants were tasked with identifying the gap's location in a target square. The squares were drawn in white (37.37 cd/m²) on a black background (0.034 cd/m².). From a viewing distance of 30 cm, these squares had an angular size of 0.5° on each side. Each square had a distinct orientation: 0°, 90°, 180° or 270°, with a notable gap. The primary targets were positioned at 90° or 270°, while the secondary or distractor squares were oriented at 0° or 180°. The primary targets were accentuated by being enclosed within a 1° white diamond of matching brightness. In contrast, the distractors were underscored with a white circle, also 1° in radius and of equal luminance. The entire visual setup was showcased using a Dell Latitude 5310 equipped with an Intel I7 processor.

Concurrent Auditory Task. The auditory component of the MAM task presents participants with a sequence of numbers ranging from 0 to 9. These numbers are randomly played via speakers at a rate of one per second, arranged in sets of 10. The rapid presentation frequency of the auditory items, at one per second, was designed to intensively consume cognitive resources, as evidence suggests that when stimuli are delivered at such a pace, multitasking becomes less efficient. The presentation is randomized and non-repetitive to avoid pattern recognition and response strategy formulation by the participants. The auditory stimuli are played through a dedicated computer setup.

Procedure. The task initiates with 10 introductory trials to familiarize participants with all possible scenarios. Each trial begins with a central focal point, which remains visible for 1.000 ms prior to unveiling the primary display. Participants are required to identify the target's gap location (either right or left) and respond using designated keys on the keyboard as quickly as possible. Both reaction times (RTs) and errors are recorded. The task incorporates two distinct manipulations: (i) display size, where either 6 or 10 items are presented, each having an equal likelihood of appearance and (ii) distractor presence, categorized as either present or absent. In the baseline scenario, only the target square, accentuated by a noticeable diamond, occupies the central position, surrounded by other squares devoid of any diamond. Conversely, in the distractor scenario, while the target maintains its central positioning and diamond accentuation, an additional non-target item is highlighted with a salient circular surround. Each trial concludes with a 500 ms interval after the participant's response. Summarily, the visual task comprises 240 trials, segmented into three consistent blocks corresponding to the auditory task's varying difficulty levels. The auditory task modulates the multitasking challenge through three conditions, thereby varying the difficulty. Each of these blocks consists of 20 trials for each combination of distractor presence (either present or absent) and display size (either 6 or 10).

Participants concurrently engaged in an auditory task. Before beginning, a separate trial for the auditory aspect was conducted to familiarize participants with its specifics and to introduce the three distinct conditions. During this auditory task, participants were prompted to say "stop" upon identifying a designated target. The task's difficulty was modulated by altering the number of auditory targets to be detected. Initially, participants listened to an mp3 file that played a series of numbers without any directives, resulting in no designated auditory targets. In the second condition, the sole auditory target to be detected was number 2. The third condition challenged participants to three auditory targets: numbers 2, 4, and 6. Throughout these conditions, the auditory targets were presented unpredictably, with participants expected to deliver swift verbal responses: responses within a 2-item delay (i.e., target item + 2 sec) were deemed acceptable based on a pilot study (Michael, Mizzi et al.,2014). We measured participants' accurate and erroneous responses. Correct responses pertained to successful detection of target numbers and non-detection of non-target numbers. Conversely, incorrect responses referred to false identifications of non-targets and missed detections of targets. To assess auditory conditions' influence, participants repeated the identical visual task three times, each corresponding to a distinct auditory setting (0, 1, or 3 targets). Uniformity of the visual task was maintained as the primary goal was to discern the influence of the auditory task's level of difficulty on RTs and errors. To counteract any sequence bias, the three auditory conditions (0, 1, 3) were systematically varied among participants, using a Latin-square design (i.e., sequences of 013, 130, and 301). Participants were instructed that both tasks held equal significance and that they should not prioritize one over the other. Drawing from the methodologies outlined by Michael, Dorey, et al. (2020), Michael, Mizzi, et al. (2014) and Salgues, S., Plancher, G., & Michael, G. A., 2023, this approach was designed to shift attention away from the visual task, allowing for assessment of the effects of constrained resource allocation.

We conducted a preliminary processing of the median reaction times and errors from the visual task. We excluded response times that were below 200 ms and above 2000 ms, as times below 200 ms indicate anticipation and times above 2000 ms signify inattention. The excluded response times accounted for 2.3% of the total responses. To assess the impact of (i) the difficulty of the dual task (0, 1, or 3 targets), (ii) the presence of distractors (absent vs. present), and the number of items (6 or 10 items) on response time and accuracy rate, we conducted a repeated measure ANOVA considering these 3 within-participant factors on error rate and response times.

Average error rate. The average error rate for the visual task was observed to be 4.4%. We observed a main effect of task difficulty, $F(2.94) = 7.08$, $p < .001$, $\eta 2\ p = .07$. Bonferroni's post hoc test revealed a significantly higher error rate when participants had to detect 3 auditory targets ($M = 5.1\%$, $SD = 0.61\%$) compared to 0 targets ($M = 3.68\%$, $SD = 0.41\%$, $p < .001$). There were no differences between 1 and 3 targets. We did not observe any main effects of the number of items presented (6 or 10) or the presence of distractors (with vs. without).

For participants' response times, we applied the Greenhouse-Geisser correction to the data, since Mauchly's test of sphericity indicated that the assumption of sphericity was violated ($p < .05$). We observed a main effect of the auditory task difficulty, $F(1.88,94) = 6.80$, $p < .001$, $\eta 2\ p =$

.07. Bonferroni's post hoc test showed a significant increase in response time when participants had to detect 3 targets (M = 1031, SD = 289, p < .001) compared to 0 targets (M = 1004, SD = 290).

We also found a main effect of the presence of distractors (present vs. absent) on reaction time, F (1.94) = 1135.65, p < .001, η2 p = 0.92, with a longer reaction time in the presence of distractors (M = 1087, SD = 295) than without distractors (M = 947, SD = 289). We did not observe any main effect of the number of items presented (6 or 10) or any interaction effects.

Auditory Task Analysis

From the auditory task, we derived two primary metrics: d′ and C[1]. An ANOVA (Analysis of Variance) was conducted to examine the influence of task difficulty on these metrics. We observed a main effect of difficulty (F (1.94) = 8.35, p < .005, η2 p = .08) on a participant's ability to detect the target among the list (d′). Participants exhibited better sensitivity when there was 1 target (M = 5.17, SD = 0.08) compared to when there were 3 targets (M = 4.72, SD = 0.09). We also noticed an effect of difficulty on the participant's response bias (C). Participants were more prone to say that the target was present when presented with 3 targets (M = 1.14, SD = 0.04) than with a single target (M = 1.24, SD = 0.03). The results indicate that the MAM task effectively highlighted variations in reaction times within the visual activity due to the presence of a distractor and the increasing complexity of the secondary task. Moreover, differences in detection accuracy within the auditory task became apparent, influenced by the number of auditory targets to be identified.

Visual search task	Errors (%)	Mean median RT (ms)
Auditory task difficulty		
0 auditory target	3.68 (0.41)	1004 (290)
1 auditory target	4.44 (0.43)	1017 (303)
3 auditory targets	5.1 (0.61)	1031 (289)
Distractor		
Absent	4.5 (0.49)	947 (289)
Present	4.4 (0.47)	1087 (295)
Display size		
6 items	4.36 (0.53)	1016 (291)
10 items	4.68 (0.55)	1019 (292)
Auditory task	C	D'
Number of targets		
1	0.47 (0.06)	1.03 (0.13)
3	0.25 (0.03)	0.49 (0.06)

Table 1: Mean (SD) Values for Errors and Median RT in the MAM Visual Search Task, for Hits and False Alarms in the MAM Auditory Detection Task

Analyses complémentaires sur le rôle des capacités attentionnelles sur l'état d'engagement

Dans cette étude, nous avons principalement examiné l'impact des capacités attentionnelles sur les performances d'apprentissage des apprenants. Cependant, nous avons souhaité réaliser a posteriori des analyses complémentaires pour explorer le rôle de ces capacités attentionnelles sur l'engagement des apprenants. En effet, bien que notre objectif initial fût d'étudier le lien entre attention et performance, il nous a semblé pertinent d'approfondir également la relation entre attention et engagement, compte tenu de l'importance de cette dernière variable dans les processus d'apprentissage.

Nous avons donc émis l'hypothèse que les capacités attentionnelles, en particulier la gestion des ressources attentionnelles, pourraient prédire significativement l'engagement situationnel des apprenants. Pour tester cette hypothèse, nous avons conduit des analyses de régression en utilisant les indices attentionnels issus du MAM (indice RD pour la gestion des ressources et indice ID pour les capacités d'inhibition) comme prédicteurs, et les différentes dimensions de l'engagement mesurées par l'UES (User Engagement Scale) comme variables dépendantes.

Engagement

Nos analyses ont révélé un effet significatif de l'indice RD (Resource Difficulty) du MAM sur le score moyen d'engagement mesuré par l'UES ($F(2\ ;91)=19.29$, $p < .001$, $\eta^2p=0.19$). Plus précisément, l'analyse de régression montre que l'indice RD prédit positivement le score moyen d'engagement ($t =4.55$, $p < .001$), mais pas l'indice ID (Inhibition Difficulty) ($t=-0.17$, $p=0.86$). En examinant chaque facteur de l'UES séparément, nous avons observé le même effet positif du RD sur les quatre dimensions : Attention focalisée ($t=3$, $p < .01$), Convivialité perçue ($t=2.36$, $p < .05$), Attrait visuel ($t=2.29$, $p < 0.05$) et Récompense ($t=4.3$, $p < .001$). En revanche, aucun effet significatif n'a été trouvé sur la perception subjective du temps.

Ces résultats suggèrent que les apprenants présentant un score RD élevé, indicateur d'une plus grande difficulté à gérer le partage des ressources attentionnelles en situation de double tâche, mais aussi d'une attention plus focalisée, semblent plus enclins à s'engager dans l'activité d'apprentissage. Ils rapportent une attention plus soutenue, une appréciation accrue de l'interface et une satisfaction intrinsèque plus marquée. Cette observation pourrait s'expliquer par le fait qu'une attention plus focalisée, même si elle peut rendre plus difficile le traitement simultané de multiples sources d'information, favorise néanmoins une implication plus profonde dans la tâche en cours (Ophir et al., 2009).

Motivation

Concernant la motivation, nous avons observé un effet principal de l'indice RD sur la motivation intrinsèque ($F(1,91)=6.04$, $p < .05$, $\eta^2p=0.07$). L'analyse de régression révèle que le RD prédit positivement le niveau de motivation intrinsèque ($t=2.24$, $p<0.05$). Nous avons trouvé le même effet sur l'amotivation ($F(1,91) =7.18$, $p< .01$, $\eta^2p=0.09$), avec un RD prédisant positivement le niveau d'amotivation ($t=2.89$, $p<0.01$). En revanche, aucun effet significatif n'a été mis en évidence pour la régulation identifiée et la régulation externe.

Ce pattern motivationnel ambivalent associé à un score RD élevé pourrait refléter une forme de dissonance entre, d'une part, l'intérêt accru suscité par un engagement attentionnel focalisé sur le contenu d'apprentissage, et d'autre part, un sentiment de moindre efficacité ou de découragement face à la difficulté perçue pour gérer la charge cognitive supplémentaire (Miele & Scholer, 2018). Ainsi, une attention très focalisée pourrait à la fois renforcer la motivation intrinsèque en favorisant une expérience d'apprentissage plus intense et gratifiante, mais aussi générer une certaine frustration ou démotivation lorsque les ressources cognitives viennent à manquer.

Pris ensemble, ces résultats mettent en lumière le rôle complexe et différencié des capacités attentionnelles individuelles dans la régulation de l'engagement et de la motivation au cours de l'apprentissage multimédia. Notre expérimentation a notamment permis de montrer qu'une capacité cognitive spécifique, à savoir la gestion des ressources attentionnelles telle que mesurée par l'indice RD, pouvait prédire significativement l'état d'engagement et certains types de motivation des apprenants face à un contenu pédagogique en ligne.

Plus précisément, nous avons constaté qu'un niveau élevé de difficulté à partager ses ressources attentionnelles en situation de double tâche, reflétant une attention plus focalisée, était associé à la fois à un engagement accru dans l'activité d'apprentissage (attention soutenue, appréciation de l'interface, satisfaction) et à un pattern motivationnel ambivalent, combinant motivation intrinsèque et amotivation. Ces résultats suggèrent que les caractéristiques attentionnelles propres à chaque apprenant, en interaction avec les propriétés de l'environnement d'apprentissage, jouent un rôle clé dans l'émergence et l'évolution de ses états d'engagement et de motivation.

Nos résultats mettant en évidence le rôle prédictif des capacités attentionnelles sur l'engagement situationnel et la motivation des apprenants viennent confirmer une partie clé de notre modèle intégratif, présenté dans le chapitre 1 (voir Figure 1). En effet, ce modèle postule que les capacités attentionnelles individuelles, en tant que caractéristiques propres à chaque apprenant, influencent directement la quantité de ressources attentionnelles pouvant être engagées dans la tâche, ainsi que la façon dont ces ressources sont déployées. Nos données empiriques corroborent cette hypothèse en montrant qu'une plus grande difficulté à gérer le partage attentionnel en situation de double tâche (indice RD élevé), reflet d'une attention plus focalisée, prédit un engagement accru dans l'activité d'apprentissage, ainsi qu'un pattern motivationnel combinant motivation intrinsèque et amotivation. Ces observations valident donc le rôle des capacités attentionnelles en tant qu'antécédent de l'engagement situationnel dans notre modèle.

Transition

Notre expérimentation précédente a mis en évidence les effets complexes et différenciés de la présence et du degré d'animation d'un agent pédagogique sur l'engagement, la motivation et les performances d'apprentissage des apprenants dans un environnement multimédia. Nous avons notamment observé qu'un agent fortement animé favorisait l'engagement situationnel et la compréhension, mais pouvait interférer avec les processus de mémorisation, en particulier chez les apprenants présentant de plus faibles capacités attentionnelles d'inhibition des distracteurs et de gestion des ressources en situation de double tâche.

Ces résultats soulignent l'importance de concevoir des agents pédagogiques suffisamment expressifs et interactifs pour soutenir l'engagement et la compréhension, mais aussi suffisamment adaptés et pertinents pour ne pas surcharger cognitivement les apprenants et perturber l'encodage des informations en mémoire (Schroeder & Adesope, 2013 ; Wang & Antonenko, 2017). Ils appellent donc à poursuivre les recherches sur les caractéristiques spécifiques des agents qui sous-tendent ces différents effets, afin d'optimiser leur conception pédagogique.

Notre travail met en évidence l'intérêt d'une approche intégrative, articulant les niveaux d'analyse des capacités cognitives individuelles, des états psychologiques et des caractéristiques des supports d'apprentissage, pour mieux comprendre et soutenir l'engagement et la motivation des apprenants dans les dispositifs de formation en ligne. Il ouvre de nouvelles perspectives pour le développement d'environnements d'apprentissage personnalisés et adaptatifs prenant en compte ces différents aspects.

Chapitre 5 : Le rôle de l'interactivité dans l'apprentissage multimédia

Dans ce contexte, les avancées récentes en matière d'intelligence artificielle (IA) ouvrent de nouvelles perspectives prometteuses pour le développement d'agents pédagogiques plus adaptatifs, interactifs et personnalisés. En effet, l'intégration de modèles d'IA génératives, tels que GPT, pourrait permettre de concevoir des agents capables d'ajuster dynamiquement leurs comportements et leurs interventions en fonction des caractéristiques individuelles des apprenants, de leurs états attentionnels et motivationnels, et de leurs performances en temps réel.

De tels agents pédagogiques "intelligents" seraient ainsi potentiellement en mesure de détecter les moments de désengagement ou de surcharge cognitive des apprenants, et d'y remédier de manière ciblée en attirant leur attention sur les informations essentielles, en clarifiant les points difficiles, ou en introduisant des pauses ou des contenus de réengagement (Grawemeyer et al., 2017). Ils pourraient également fournir un soutien motivationnel et métacognitif personnalisé, en s'adaptant aux états affectifs et aux besoins spécifiques de chaque apprenant.

Cependant, l'efficacité réelle de tels agents pédagogiques "augmentés" par l'IA reste encore à démontrer empiriquement. Il convient notamment d'examiner dans quelle mesure leurs capacités accrues d'adaptabilité et d'interactivité permettent effectivement de mieux réguler l'attention et l'engagement des apprenants, de réduire la charge cognitive inutile, et *in fine* de favoriser des apprentissages plus profonds et pérennes (Martha & Santoso, 2019).

C'est précisément l'objectif que nous nous sommes fixés dans la dernière étude empirique de cette thèse. En nous appuyant sur le même environnement d'apprentissage que précédemment, nous avons cherché à comparer les effets de deux types d'agents pédagogiques. Un agent "classique", similaire à la condition "forte" de l'étude précédente et un agent "intelligent", disposant de la capacité de communiquer avec l'apprenant, de répondre à ses questions. En reprenant notre protocole de mesure de l'engagement, de la motivation et des performances d'apprentissage, ainsi que d'évaluation des capacités attentionnelles individuelles, notre but est de tester l'hypothèse selon laquelle un agent pédagogique fondé sur un modèle d’IA générative serait plus à même de stimuler et de maintenir l'engagement des apprenants tout en respectant les limites de leurs ressources cognitives et attentionnelles, comparativement à un agent traditionnel.

Beyond Embodiment: Exploring the impact of Generative AI in Animated Pedagogical Agents on multimedia learning.

Y. Minne[1,2,3], E. Lavoué[3] C. Thiery[2] and G. A. Michael[1]

[1]Unité de Recherche Étude des Mécanismes Cognitifs, Université de Lyon, Université Lyon 2, Lyon, France

[2] Sydo, 69001 Lyon, France

[3] Iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, Université de Lyon, Université Lyon 3, Lyon, France

Abstract

This study investigated the impact of an advanced, interactive, and embodied Animated and Conversational Pedagogical Agent on learners' engagement, motivation, and learning outcomes in a multimedia learning environment. The role of individual attentional capacities in moderating these effects was also examined. Ninety-three business school students were randomly assigned to either a control condition with a traditional Animated Pedagogical Agent (APA) or an experimental condition with an interactive agent powered by AI. The interactive APA, powered by GPT-4, allowed learners to interact with the agent, ask questions, and receive personalized responses, setting it apart from the traditional non-interactive APA used in the control condition. Both groups studied an e-learning module on immunotherapy. Learners' engagement, motivation, learning performance, and attentional capacities were assessed using various measures, including the User Engagement Scale (UES), Situational Motivation Scale (SIMS), comprehension and memorization questions, and the Master Activation Map (MAM) attentional test. The results partially supported the hypothesis that an interactive agent would enhance learner engagement compared to a non-interactive APA. Learners in the interactive agent condition reported significantly higher levels of engagement, particularly in terms of focused attention and perceived usability. However, no significant differences were found in learner motivation between the two conditions. The subjective perception of time (SPT) emerged as a noteworthy measure of engagement, with learners in the interactive agent condition significantly underestimating the time spent on the e-learning platform, suggesting a state of deep absorption and engagement. Contrary to expectations, learning outcomes did not

significantly differ between the interactive agent and control conditions. Learners in the control condition performed slightly better on memorization questions compared to those in the interactive condition. Learners' inhibition difficulties (ID index) negatively predicted performance outcome, highlighting the role of attentional control in learning performance.

A. Introduction

In 2021, a landmark advancement emerged in the artificial intelligence landscape with the advent of ChatGPT, signaling a potential paradigm shift in multimedia learning. The evolution of technology has continuously shaped the modern educational landscape, transforming instructional methods and tools. As the frontier of this technological revolution in education unfolds, two emerging theoretical frameworks centered on Animated Pedagogical Agents (APA) and Conversational Automated Agents (CA) offer profound insights. The former, based on the cognitive theory of multimedia learning, predominantly delves into the visual characteristics of the APA and their impacts on learning outcomes and motivation. In contrast, the second approach takes a technical perspective, exploring the potential interactions between the learner and the APA but without studies on the impact of psychological effect. We propose to combine these two frameworks to examine the impact of animated and conversational pedagogical agents.

1. Framework of Animated pedagogical agents

The Cognitive Theory of Multimedia Learning (CTML) is a comprehensive theory that tries to explain how individuals interpret and comprehend multimedia presentations. It was put forth by Mayer (2005), emphasizing that individuals have two primary channels for processing information: auditory-verbal and visual-pictorial. These channels operate concurrently, allowing for efficient information processing. However, they are limited in capacity, making the structure of multimedia presentations crucial to optimizing understanding (Mayer, 2009). At the heart of CTML lies the dual-coding theory (Paivio,1971), suggesting that information is processed in separate, yet simultaneous, verbal and visual channels. This leads to an enhanced understanding and recall when both visuals and text are combined compared to using either mode in isolation.

A pivotal principle within the CTML is the embodiment principle. This principle was explored in depth by Mayer and DaPra (2012) in their research on animated pedagogical agents. They investigated the influence of social cues, such as gesturing, facial expressions, eye gaze, and human-like movement, on multimedia learning. Their findings revealed an "embodiment effect," where learners performed better in a transfer test when a human-voiced agent displayed human-like gestures, facial expressions, eye gaze, and body movement than when the agent lacked these features. The APA impact was better when the agent spoke in a human voice but diminished with a machine voice. On the other hand, the voice principle underscores the significance of using a human, friendly voice in multimedia instructions. It enhances the learning process by making the content more relatable and engaging.

Mayer and colleagues (2005) further expanded on this by introducing the Social Agency Theory. This theory posits that social cues in multimedia messages, like human-like gestures and voice, prime a feeling of social partnership in learners. Such a feeling triggers deeper cognitive processing during learning, culminating in more meaningful learning outcomes. These priming pushes learner to exert more effort in understanding the presented material, leading to a more meaningful learning outcome. This theory aligns with the broader media equation theory by Reeves and Nass (1996), which posits that media are perceived as real people, and human responses to media are based on the rules of social relationships. There remains a significant gap in the application of these principles to animated pedagogical agents (APAs). Many current implementations of APAs in educational settings tend to be predominantly non-interactive or minimally interactive, acting largely as passive deliverers of content. This approach, while benefiting from the embodiment and voice principles, may not fully harness the potential of APAs in fostering deep and meaningful learning experiences. The true strength of APAs, especially when rooted in the Social Agency Theory, lies in their ability to engage learners in dynamic, two-way interactions, simulating the depth and nuance of human-to-human conversations.

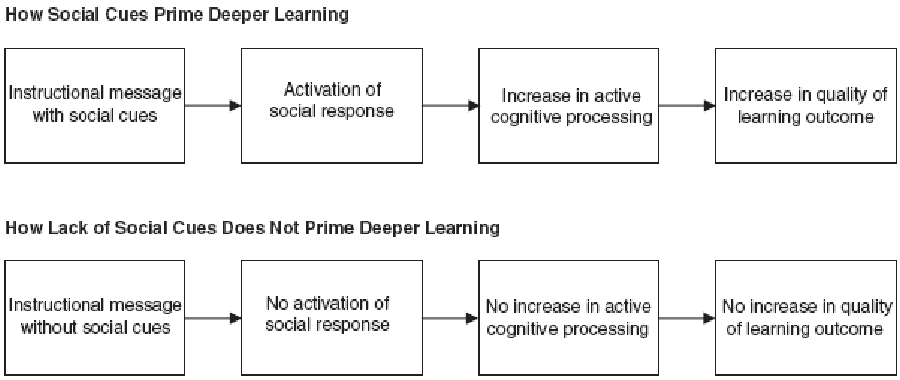

Figure 1: Theory of social agency from Mayer (2012)

The incorporation of pedagogical agents in multimedia learning has been shown to enhance learner engagement and motivation. Lester et al. (1997) introduced the "persona effect," suggesting that the mere presence of a lifelike character in a multimedia learning environment can have a positive impact on learners' perceptions and attitudes. Subsequent research has suggested that pedagogical agents can foster learner engagement by providing social cues, such as facial expressions, gestures, and voice intonation, which create a sense of social interaction and partnership (Mayer, 2014; Moreno et al., 2001). Moreover, pedagogical agents can influence learner motivation by providing personalized feedback, encouragement, and guidance (Kim & Baylor, 2016). Agents that are designed to be empathetic, supportive, and adaptive to individual learner needs have been found to enhance intrinsic motivation and promote a more positive learning experience (Kim & Baylor, 2016; Rosenberg-Kima et al., 2008).

However, poorly designed APAs can lead to situations that negatively impact learners' attentional processes, such as divided attention (Domagk et al., 2010; Lusk & Atkinson, 2007) and impaired inhibition, which may be detrimental for learning. When the design of the pedagogical agent is not well-integrated with the learning content or when the agent's behavior is not relevant to the learning task, learners' attention may be distracted and divided between the agent and the content. This division of attentional resources can increase cognitive load and hinder learning (Sweller, 2005). Moreover, the intensive aspect of attention (Hicks et al., 2021) and its influence on cognitive resources (Kahneman, 1973; Michael et al., 2006) can also affect inhibition (Michael et al., 2014), which is crucial for suppressing irrelevant information and maintaining focus on pertinent aspects of the learning material.

Individual differences in attentional capacity (Unsworth & Miller, 2021; Hicks et al., 2021) and inhibitory control can significantly influence how learners interact with and benefit from multimedia learning environments (Lusk & Atkinson, 2007; Sanchez & Wiley, 2006). Learners with higher attentional capacities and better inhibitory control are more adept at managing the cognitive demands of processing multiple sources of information simultaneously, such as the learning content and the pedagogical agent (Fenesi et al., 2015). They are also more effective at filtering out irrelevant information and maintaining focus on the most pertinent aspects of the learning material (Domagk et al., 2010).

In contrast, learners with lower attentional capacities and weaker inhibitory control may struggle to effectively process and integrate information when presented with multiple sources of information (Fenesi et al., 2015; Sanchez & Wiley, 2006). They are more susceptible to the detrimental effects of divided attention and may experience cognitive overload more easily (Domagk et al., 2010). As a result, poorly designed pedagogical agents can be particularly disruptive for learners with lower attentional capacities and weaker inhibitory control, as they may not have the cognitive resources to effectively manage the additional load imposed by the agent and suppress the processing of irrelevant information (Lusk & Atkinson, 2007).

2. Framework on conversational Agents

Conversational Agents (CAs) are tangible manifestations of artificial intelligence (AI) designed to interact with humans in a natural manner. These agents operate based on a set of rules and algorithms that guide their behavior, allowing them to mimic human conversations through voice commands, text chats, or a combination of both (Radziwill and Benton, 2016; Klopfenstein et al., 2017). Over the years, the educational sector has shown increasing interest in the potential applications of CAs. Researchers have explored various applications of these agents to achieve educational objectives. For instance, CAs have been employed for tutoring purposes (Heffernan and Croteau, 2004; VanLehn et al., 2007), answering questions (Feng et al., 2006), aiding language learning (Griol et al., 2014), and even fostering the development of metacognitive skills (Kerly et al., 2008). The overarching consensus from these studies suggests that CA-based services positively influence student satisfaction and learning outcomes (Kerly et al., 2007; Huang et al., 2017). Goel et al. (2015) highlighted the integration of a CA in an online course to manage forum posts, which led to heightened student engagement. Notably, some studies have delved into the use of such agents to foster productive peer dialogue and support collaborative learning scenarios, with promising outcomes (Kumar and Rose´, 2011; Walker et al., 2011).

Chatbots constitute a more specific subset of CAs. These are computer programs designed to provide services through dialogue. Recent years have seen a surge in research exploring the possibilities and impacts of chatbot applications in educational contexts (Ferrell & Ferrell, 2020). The benefits span from enhancing real-time interactions (Kim et al., 2019) to bolstering peer communication skills (Hill et al., 2015) and even optimizing learners' efficiency (Wu et al., 2020). These agents can undertake various roles such as tutoring, organizing, motivating, mentoring, and moderating (Gubareva & Lopes, 2020). For instance, a tutoring CPA's primary role is to deliver learning content (Taoum et al., 2019; Winkler, Hobert, Salovaara, et al., 2020), while the organizer focuses on academic integration tasks like university admissions (Al Muid et al., 2021), course management (Priadko et al., 2020), or career opportunity guidance (El Hefny et al., 2021). Motivating CPAs engage users to invest more time in learning, utilizing methods like gamification (Krassmann et al., 2019) or strategies to overcome procrastination (Rodriguez et al., 2019). Mentoring agents track student progress through quizzes (Kita et al., 2018) and self-assessments (Durall & Kapros, 2020). The moderator agent is a relatively rare and consists of facilitating group learning activities, including group work (David et al., 2019) and collaborative learning environments (Graesser et al., 2018). Interestingly, a single CPA can encompass multiple roles, creating a multi-faceted learning experience. Notably, the literature has shown a distribution of these roles as 31% tutor, 21% motivator, 21% organizer, 18% mentor, and 9% moderator (Winkler et al., 2020).

CPAs facilitate not only the recall of factual information but also the development of deeper, causal reasoning and problem-solving skills, essential in managing complex and ambiguous scenarios (Graesser et al., 2017). This approach allows CPAs to adapt their interactions to encourage more profound cognitive engagement, prompting learners to think critically and analyze information, thereby enhancing their understanding and retention of the material. Additionally, the integration of incremental learning theories into CPA design supports the breakdown of complex subjects into smaller, digestible learning units. This strategy, often referred to as micro-learning, aligns content delivery with the learner's cognitive capacity, making it easier to absorb and retain information (Yin et al., 2021). By delivering content in sequences that build upon each other, CPAs provide a form of cognitive scaffolding, aiding learners in gradually acquiring knowledge and skills without becoming overwhelmed.

Furthermore, CPAs play a crucial role in fostering collaborative and social learning environments. They are not merely facilitators but active participants in learning processes, engaging in dialogues that simulate interactions akin to those with human peers or instructors. This capability is particularly significant in collaborative problem-solving tasks, where the CPA can offer real-time feedback and guide the learner through complex learning activities (Graesser et al., 2018; Hayashi, 2018).

Beyond these specific roles, some research focuses on design principles for communication style (Wolfbauer et al., 2020) or avatar design (Taoum et al., 2019), indicating the breadth and depth of the field. Technologically, Conversational Pedagogical Agents (CPAs) manifest in a variety of formats and designs, each tailored to specific pedagogical needs and technological capabilities. The simplest form is the text-based CPA, which interacts with learners solely through textual dialogue. These agents, often resembling chatbots, are typically integrated within websites or messenger platforms, allowing for seamless interactions without the need for intricate visuals (Parthornratt et al., 2018). More sophisticated are CPAs with avatars, which not only communicate textually but also provide visual cues, gestures, and expressions, enhancing the embodiment principle and offering a more human-like interaction (Mayer & DaPra, 2012). Under the hood, many of these agents operate based on predefined rules, producing responses from a set list of possibilities. This "rules-based" approach, while straightforward and reliable, can sometimes lack the dynamism and adaptability of more advanced systems (Dahiya, 2017). As artificial intelligence (AI) and machine learning technologies advance, newer CPAs are starting to leverage these techniques, allowing for more fluid, context-aware interactions that go beyond the constraints of rules-based systems. These AI-driven agents utilize natural language processing (NLP) and natural language understanding (NLU) to comprehend and generate more nuanced responses, providing a richer, more personalized learning experience (Braun et al., 2017).

Central to this transformation is the emergence of the Transformer architecture, which underpins models like OpenAI's GPT (Generative Pre-trained Transformer) series. These models are known for their unparalleled capacity to generate coherent, contextually relevant textual content, simulating human-like linguistic patterns (Vaswani et al., 2017). Moreover, advancements in voice synthesis technologies, such as WaveNet and Tacotron, have enabled the creation of lifelike digital voices, further blurring the boundaries between human and machine communication (van den Oord et al., 2016; Wang et al., 2017). Historically, research on Animated Pedagogical Agents (APA) leveraging advanced technologies predominantly stemmed from computer science. However, the growing accessibility of these technologies presents an invaluable opportunity for scholars in the Cognitive Theory of Multimedia Learning (CTML) to delve deeper into the impact of pedagogical agent. This confluence of pioneering tools and pedagogical practices invites an examination of their psychological and cognitive impacts on learners.

3. Limitations of this framework

In their systematic literature review, Khosrawi-Rad et al. (2022) examine the use of Conversational Agents (CAs) in educational contexts on 252 articles, uncovering pivotal insights and notable limitations that hold significant implications for future research. This review highlights two critical areas where improvements are necessary: methodological robustness and theoretical grounding.

Khosrawi-Rad et al. (2022) identify a prevalent issue in CA research: the lack of methodological rigor. Many studies within the scope of their review do not employ mixed-methods approaches, which are crucial for gaining a comprehensive understanding of how CAs impact learning outcomes. The absence of such approaches results in a one-dimensional view that neglects the nuanced interplay between quantitative data (e.g., performance metrics) and qualitative insights (e.g., student satisfaction and engagement levels) (Ivankova et al., 2006). Furthermore, the review criticizes the short duration of most studies, which fail to assess the long-term effects of CAs on learners, such as sustained motivation and success over time (Salkind, 2010; Hobert, 2019a).

Another significant limitation highlighted by Khosrawi-Rad et al. (2022) is the insufficient theoretical foundation in the studies reviewed. Many PCA implementations are not grounded in established educational or psychological theories, which can lead to designs that are not aligned with pedagogical objectives or that fail to consider important psychological impacts on learners. This theoretical gap undermines the potential of CAs to be effectively integrated into educational systems and to be tailored to different learning contexts and needs (Gregor, 2002; Kuechler and Vaishnavi, 2008).

In this study, we delve into the effects of interactive, animated Pedagogical Agents on motivation, engagement, learning outcomes in multimedia learning. To draw a clear distinction between the potential impacts of agent interactivity, learners are presented with consistent content on immunotherapy across two experimental conditions. The mode of interaction with the pedagogical agent differs between the scenarios. In both conditions, learners encounter a pedagogical agent complete with voice and bodily attributes, ensuring a consistent visual and auditory experience. In the first condition, learners engage with a traditional APA, which, while offering a human-like visual and auditory presence, remains non-interactive in its approach. Conversely, the second condition introduces an animated and conversational pedagogical agent that, while retaining the same embodied attributes of voice and appearance, is powered by the GPT architecture.

This advanced agent is designed to accompany learners throughout their learning trajectory, actively partaking in dialogue, offering feedback, and addressing queries. To ensure the most human-like interaction experience possible, learners communicate with the agent using their voice via a microphone, fostering a more organic and immersive conversational flow. Notably, despite the enhanced interactivity in the GPT-driven agent condition, no additional informational content is introduced. The foundational content remains uniform across both conditions, enabling us to precisely evaluate the unique contributions and impacts of agent interactivity on the overall learning process. Our primary hypothesis (H1) suggests that an interactive agent will bolster learner engagement, motivation, and learning outcomes relative to non-interactive agents. In

addition to evaluating the differential impacts of the pedagogical agents on learning outcomes, our study also involved measuring participants' attentional capacities, specifically focusing on their abilities in inhibition and resource management. We posited an exploratory hypothesis (H2) that learners with enhanced attentional capacities—those better able to manage cognitive resources and inhibit distracting stimuli—would achieve superior learning outcomes. This hypothesis suggests a dual role for attentional capacities in educational settings: while strong attentional skills may mitigate the cognitive load introduced by complex pedagogical agents, thus enhancing learning, these same capacities may also support learners who otherwise struggle with poor attentional control by helping them manage the increased cognitive demands posed by advanced pedagogical agents.

B. Methodology

1. Participants

The research engaged 93 individuals, selected from a pool of students of the EM Lyon, a business school in Lyon, France. The mean age of the participants is 26 years (SD = 5.1). There are 59 women and 34 men. All the participants are native French speakers. The participants were randomly assigned into the APA control group (N =44; 27 women and 17 men; M= 25.4 years; SD = 5) and the interactive APA group (N = 49; 32 women and 17 men; M= 26.6 years; SD = 5.2). In line with the ethical guidelines of the Helsinki Declaration, which prioritizes the welfare and respect for all individuals participating in research, this study adhered strictly to ethical principles. These principles highlight the necessity of informed consent, the reduction of harm, and the safeguarding of participants' rights. Every participant received comprehensive information about the study's aims, methods, potential risks, and benefits before giving their written consent. Additionally, they were guaranteed data confidentiality and informed they could exit the study whenever they wished.

2. Material

2.1. E-learning Material and devices

The study employed an e-learning module centered on the topic of immunotherapy as the primary multimedia material. This comprehensive module delved into the complexities of the immune system, its intricate response to cancer, and the finer points of immunotherapy. The module was structured into five distinct sections: Module 1 focused on the immune system, Module 2 explored the relationship between the immune system and cancer, Module 3 provided an in-depth look at immunotherapy, Module 4 highlighted the key differences between immunotherapy and chemotherapy, and Module 5 offered additional insights into immunotherapy. The e-learning experience was enriched with animated videos accompanied by voice-overs narrated by the pedagogical agent. The user interface presented an overview of the various e-learning modules (fig.1). Learners could directly access the desired content by clicking on any of the modules. Within each module, users had the flexibility to navigate between videos using the strategically placed next, previous, and return to main menu buttons (see fig.1). For the purposes of the present study, two unique versions of the e-learning module were crafted. The control condition (depicted in Fig. 2) employed a symbology system to provide a visual representation of cells. The relationships between these cells were further depicted through the use of arrows and action verbs. Complementing this visual approach was a pedagogical agent that, while animated and passive, exhibited expressive gestures and facial expressions to enhance the learning experience. In contrast, the Interactive APA version (illustrated in fig3) preserved the same core content but introduced an Advanced Pedagogical Agent (APA) powered by a large language model and Eleven Labs to interact with the learner and answer their questions. To facilitate learner interaction with the APA through voice communication, a microphone button was seamlessly integrated into the interface. The two agents in the delivery content mode were meticulously designed to align with the narration, pointing to specific areas of the screen or displaying expressions that mirrored the content being presented.

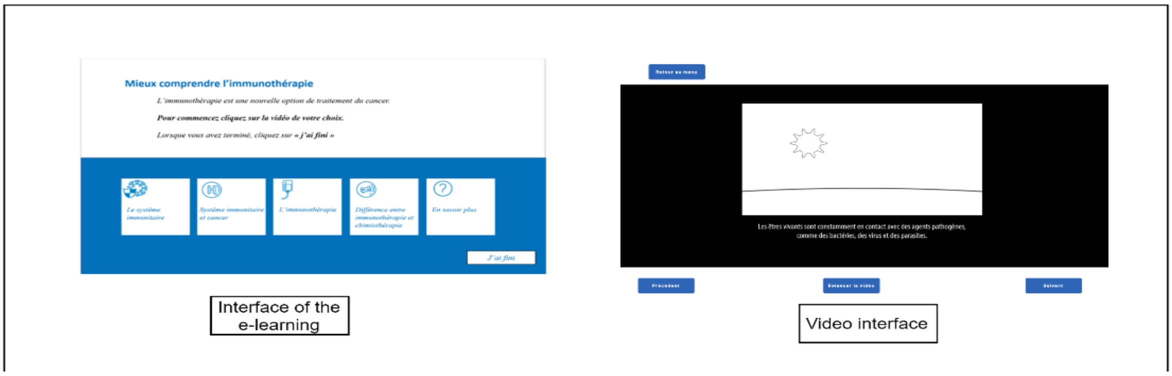

Figure 2: Learning materials. Interfaces of the e-learning environment: on the left the main interface to access the modules, on the right a screenshot of the interface to navigate in the video.

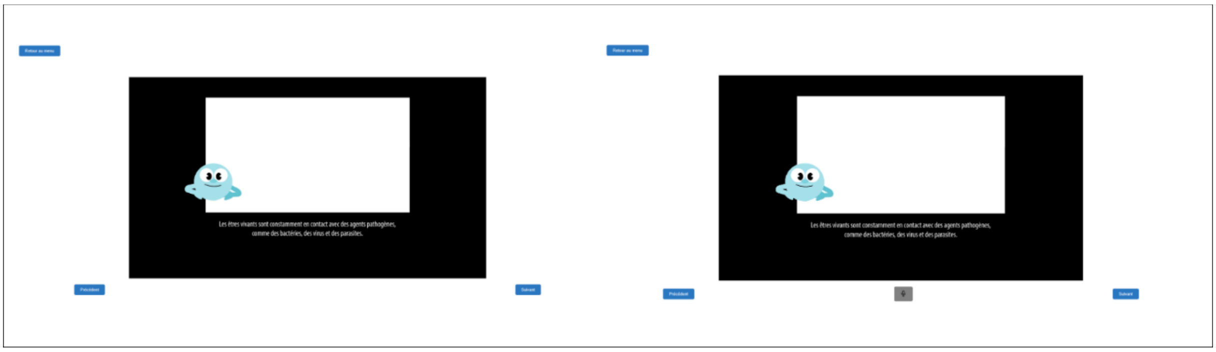

Figure 3: Learning materials. On the left, the video of the control APA condition and, on the right, the video of the interactive APA condition

2.2. Architecture of the interactive APA

The chatbot, a key component of the interactive APA version, was developed using a combination of modern tools and platforms. The chatbot architecture was built using Flowise, an open-source, no-code platform. The GPT-4 language model, hosted on Microsoft Azure for data privacy compliance, was used to generate the chatbot's responses. Chroma, a vector database, was employed for efficient data vectorization, transforming the learning content into a format that the GPT-4 model could process and respond to. To enhance the chatbot, Eleven Labs' text-to-speech technology was integrated, providing a realistic voice for content delivery.

The technical details of each tool and their integration process are further elaborated in appendix 1. The APA can provide contextualized support to learners as they progress through the e-learning content. It has access to all the documents and materials used in the e-learning modules, allowing it to draw upon this information when generating responses to learners' queries. Additionally, the chatbot is designed to track the learner's position within the e-learning platform, enabling it to provide relevant and timely assistance based on the learner's current context. The chatbot has been programmed to avoid redundancy and repetition of information already presented in e-learning content, which is known to be beneficial for memorization. Instead, it focuses on providing complementary explanations, examples, and guidance to enhance the learning experience without unnecessary duplication of content.

To ensure the quality of the answers and to prevent redundancy, we conducted extensive pre-testing. During these tests, we carefully evaluated the chatbot outputs to ensure that it does not repeat information already presented in the e-learning content. This pre-testing phase allowed us to refine the chatbot responses.

2.3 Hardware

The study was carried out using a portable computer, the Dell Latitude 5530, which boasts a compact 13.3-inch display. To enable seamless interaction with the interactive APA version, a microphone was strategically placed in front of the computer. The e-learning content was showcased in a full-screen format, with the video material confined within a digital rectangle spanning 1280 x 720 pixels. The display resolution was optimized to 1366 x 768 pixels, and the video content was rendered using the computer's built-in media player software. To maintain a standardized computing environment, the portable computer was powered solely by its battery throughout the duration of the experiment. The ambient lighting within the experimental space was meticulously controlled to remain consistent, mitigating any potential external influences. To create an optimal viewing scenario, the computer screen was carefully adjusted to form a perpendicular angle with the surface of the table. The participant was seated directly opposite the screen, preserving a consistent distance of 30 cm. This deliberate screen positioning guaranteed that the video material was displayed in its intended aspect ratio, eliminating any visual distortions or loss of detail that could arise from suboptimal screen angles.

2.4. Measures

Control measures.

We assessed several control variables to ensure that they do not interfere with the experimental situation. These variables include gender, age, interest in the subject, and prior knowledge about immunotherapy.

Interest in the subject. We incorporated three custom-made statements to gauge the learners' interest in the subject. Participants were asked to indicate their level of agreement (1 = strongly disagree; 5 = strongly agree) with each of the three presented statements on a Likert scale ("I was already interested in immunotherapy before the course."; "This course was new to me"; "I knew about immunotherapy before taking this course"). The scores from the three statements were averaged to obtain an overall interest score. A higher score indicates a greater level of interest in the topic prior to the course.

Prior Knowledge. We included a single question to assess the participants' level of expertise in immunotherapy, rated on a scale ranging from novice to expert ("How would you rate your knowledge of immunotherapy?"). Prior knowledge has been demonstrated to have a significant impact on the effectiveness of pedagogical design (Mayer, 2012).

Measure of interest.

MAM - Attentional abilities

To evaluate participants' attentional capacities, we employed a dual-task approach, as described in previous studies (Michael, Dorey, et al., 2020; Michael et al., 2007; Michael, Pannetier, et al., 2014, Salgues, S et al, 2023). In the MAM test, the primary task involved a customized computerized visual search that required participants to locate a predefined target while inhibiting distractors. Concurrently, an auditory detection task was presented, challenging participants to skillfully manage cognitive resources across both tasks. To increase the challenge, the difficulty of the auditory task was varied, requiring participants to shift their cognitive resources from the visual task. These two components provide a comprehensive and multifaceted assessment of attention, specifically targeting inhibitory control and the management of attentional resources. We followed the methodology used in recent studies (Michael, Dorey, et al., 2020; Salgues S et al 2023), which is described in more detail in Appendix 1.

Our dependent variables are learners' learning performance, motivation, and engagement.

Learning performance. We utilized a set of 10 comprehension and 10 memorization questions to assess learning performance after the task. Comprehension and memorization questions are two distinct types of assessment methods for evaluating learning performance. Comprehension questions assess the learner's understanding of concepts, while memorization questions evaluate the ability to recall specific details. Participants were required to indicate whether the presented statement was true or false. An example of a memorization question is "Tears and lymphocytes are the initial defenses against bacterial intrusion." An example of a comprehension question is "The mechanism allowing lymphocyte education is slower than the innate immune response." Half of the questions required a "correct" answer, while the other half required an "incorrect" answer. Performance was reflected by the proportion of correct responses, as well as by two signal detection indices, *d'*, which reflects the ability to distinguish true statements from false ones, and *C*, which represents participants' response tendencies towards "correct" or "incorrect" responses, regardless of whether the responses are actually correct or not.

Motivation – Situational Motivation Scale (SIMS). To assess learners' motivation, we employed the Situational Motivation Scale (SIMS) originally developed by Guay et al. (2000) and later adapted and validated in French by Fontaine et al. (2019). This instrument comprises 16 items that measure four distinct types of motivation: intrinsic motivation (e.g., "because I find this activity interesting," 4 items), identified regulation (e.g., "because I believe it's beneficial for me," 4 items), external regulation (e.g., "because I feel obligated to do it," 4 items), and amotivation (e.g., "I may have reasons for engaging in this activity, but personally, I don't see any," 4 items). Participants rated their agreement with each statement on a Likert scale ranging from 0 (Disagree) to 7 (Totally agree). Higher scores on the first three types of motivation indicate greater learner motivation, while elevated scores on amotivation suggest a lack of willingness to continue engaging with the learning material.

Engagement – User Engagement Scale (UES). To measure engagement, all participants completed the abridged version of the User Engagement Scale (UES) (O'Brien & Toms, 2010), which was translated and validated in French by Fontaine et al. (2019). Participants rated their level of agreement (1 = strongly disagree; 5 = strongly agree) with each presented statement on a Likert scale. The scale consists of four dimensions, each comprising three items: Focused Attention (sense of immersion in the interaction and losing track of time), Aesthetic Appeal (attractiveness and visual appeal of the interface), Perceived Usability (negative affect experienced due to the interaction and the degree of control and effort required), and Reward (enjoyment, novelty, interest, and overall satisfaction). As recommended by O'Brien and Cairns (2018), scores for the Perceived Usability dimension were reverse-coded. Higher overall scores indicate greater participant engagement.

Engagement - Subjective Perception of Time (SPT). Following the guidelines proposed by Hancock (2019), we utilized a retrospective approach in which learners were asked to estimate the amount of time they believed they spent navigating the e-learning modules at the end of the course, without consulting their phones or other devices. This perceived time was then compared to the actual time spent in the e-learning by dividing the perceived time by the actual time. A higher ratio suggests an overestimation of the time elapsed. It is assumed that learners who overestimate the time spent may be experiencing boredom, while those who underestimate are likely more engaged and immersed in the learning activity.

2.5. Procedure

The study was carried out in a controlled laboratory setting. At the outset, participants were randomly allocated to one of the two experimental conditions: control APA or interactive APA. All participants began by completing the MAM task, which lasted between 10 to 15 minutes. Following this, they responded to an initial questionnaire designed to gauge their level of expertise and pre-existing interest in the subject of immunotherapy. Subsequently, participants were given a minimum of 15 minutes to explore and interact with the immunotherapy e-learning platform.

Upon concluding the e-learning session, participants were asked to complete a series of measures assessing various aspects of their experience, as detailed in the previous section. Lastly, the comprehension and memory questions were presented to each participant in a randomized and intermixed order. The entire experimental session duration was approximately 30 minutes.

Prior to commencing the study, participants were provided with clear instructions regarding their engagement with the e-learning material. They were asked to approach the content as if they were preparing for an upcoming examination, and to respond to the comprehension and memory questions to the best of their abilities. This approach was chosen to ensure that participants remained focused and invested in the learning process throughout the experimental session.

C. Results

Learner's characteristics

We conducted one-way ANOVAs and chi-squared tests to control for the impact of individual characteristics on the results. There were no significant age differences among participants assigned to each version (Control M = 25.4, SD= 5; Interactive M = 26.6 SD=5.2; $F(1,91) = 1.12$, $p = 0.293$). Additionally, the distribution of men and women across groups was not significantly different (Control M = 17, W = 27; Interactive M = 17, F = 32; $\chi^2(1) = 0.155$, $p = 0.69$). Regarding participants' self-reported expertise of the content, no significant group effects were observed (Control M = 1.18, SD = 0.39; Interactive M = 1.14, SD = 0.35; $F(1,91) = 0.255$, $p = .61$). Similarly, there were no significant differences in participants' interest in the topic across version (Control M = 2.29, SD = 0.76; Interactive M = 2.19, SD = 0.68; $F(1,91) = 0.486$, $p = .48$).

Effect of the Support Version on engagement, motivation and performances.

We carried out a series of ANOVAs to explore the differences in scores across the dependent variables with the APA version (Control, Interactive) as a between-participants factor. Cohen's *d* and η^2p were used for the size effect.

Engagement and motivation

UES - A notable difference between the two groups was identified in the average UES score of the four factors of the UES, ($F(1,91) = 4.61$, $p < .05$, $\eta^2p = .04$). Participants with the interactive APA declared a higher engagement (Interactive M = 3.24, SD = 0.66) compared to participants with control APA (Control M = 2.94, SD = .67). Taking each UES factor separately, this difference was found to be consistent across two of the four factors: Focused Attention ($F(1,91) = 5.74$, $p < 0.05$, $\eta^2p = 0.06$; Control M = 2.36, SD = 0.93; Interactive M = 2.88, SD = 1.13), Perceived Usability ($F(1,91) = 17.37$, $p < .001$, $\eta^2p = 0.16$; Control M = 3.04, SD = 0.89; Interactive M_ = 3.76, SD = 0.78). No significant differences were observed between control and interactive APA condition for the reward and aesthetic appeal dimensions.

SPT - A main effect of the condition on the subjective perception of time was found, $F(1,91) = 57.69$, $p < .001$, $\eta p^2 = 0.38$. Participants underestimated the time spent in the Interactive APA condition ($M = 0.62$, $SD = 0.12$) compared to the control APA condition ($M = 0.93$, $SD = .25$). To further investigate whether the SPT scores in each condition significantly differed from the theoretical value of 1 (the situation where the learner accurately estimates the time spent), we conducted one-sample t-tests for each condition. The results showed that the SPT score in the control condition ($M = 0.93$, $SD = 0.25$) did not significantly differ from 1, $t(43) = -1.883$, $p = .066$. However, the SPT score in the interactive condition ($M = 0.62$, $SD = 0.12$) was significantly lower than 1, $t(49) = -20.944$, $p < .001$.

Motivation - The effect of the condition on learners' motivation was not significant on the four dimensions of SIMS.

UES and Subjective Perception of Time. A negative correlation ($r = -0.219$, $p = 0.03$) was found between the SPT and the overall UES score. This suggests that as learners' engagement increases, their perception of time spent on the e-learning platform tends to be underestimated. There was a notable negative correlation between the subjective perception of time and the UES score for both Focused Attention ($r = -0.268$, $p = .009$) and Perceived Usability ($r = -0.355$, $p<.001$).

Learning performance and attentional abilities

To assess whether attention abilities combine with the APA version to determine performance outcome, we conducted a mixed-design ANCOVA on the proportion of correct responses to the questions participants received about the e-learning content. The question type (comprehension vs. memorization) was used as the within-participants factor, the APA version (Control and Interactive) as the between-subjects factor, and the attentional abilities (RD and ID) as covariates and moderators.

We observed relatively low performance on the proportion of correct responses on the task for both memorization ($M = 0.57$, $SD = .09$) and comprehension ($M = 0.55$, $SD = 0.11$) questions. However, one-sample t-tests revealed that the performance was significantly better than chance (0.5) for both question types in both conditions (all $p < .05$). We observed an interaction between the type of question and the tested condition ($F(1,91) = 9.54$, $p < .01$, $\eta^2 p = .09$). Bonferroni's post hoc tests revealed that participants in the control APA version scored higher on memorization ($M = 0.6$, $SD = .08$) than on comprehension ($M = 0.53$, $SD = .11$; $t(91) = 3.0$, $p = 0.01$). We also observed that the performance on memorization was significantly higher in the control condition ($M = 0.60$, $SD = 0.08$) than in the interactive condition ($M = 0.54$, $SD = 0.09$; $t(91) = 2.94$, $p = 0.02$). Interestingly, we didn't observe any significant effects of condition or question type on the two signal detection theory indices, C and d'.

To assess the validity of the MAM task and derive two indices reflecting attentional abilities, we conducted several tests. The first index, Inhibition Difficulty (ID), indicates the participant's challenges in suppressing distractors, while the second index, Resource Management Difficulty (RD), represents the difficulty in allocating attentional resources between the visual and auditory aspects of the MAM task (Michael, Dorey, et al., 2020). Appendix 1 provides a detailed description of the methods employed to derive these indices and the psychometric properties of the MAM task. Our analyses focused on the proportion of correct answers, d', and C. We discovered a significant main effect of the ID index on mean d' ($F (1,92) =9.49$, $p<0.01$). Further regression analyses revealed that the ID index served as a negative predictor of mean *d'* ($\beta=-1.46$, SEM = 0.06, $t (91) =-2.282$, $p=.02$, _R_2=.05).

We also investigated the impact of these indices on various measures of motivation and engagement. While no significant effects were found on motivation, both the RD and ID indices were found to predict the subjective perception of time. Specifically, the RD index positively predicted SPT ($\beta=0.04$, SEM = 0.09, $t (92) =2.24$, $p=.02$, _R_2=.05), while the ID index negatively predicted SPT ($\beta=-0.07$, SEM = 0.03, $t (92) =30.29$, $p=.01$, _R_2=.06). Additionally, we observed that only the ID index significantly predicted the overall engagement score, as measured by the UES ($\beta=0.23$, SEM = 0.08; $t (92) =43.88$, $p<.001$).

D. Discussion

The present study aimed to investigate the effects of an interactive and Animated and Conversational Pedagogical Agent on learners' engagement, motivation, learning outcomes, and cognitive load in a multimedia learning environment. The study also sought to examine the role of individual attentional capacities in moderating these effects.

Despite this limitation, our first hypothesis (H1) posited that an interactive agent would enhance learners’ engagement, motivation, and performance compared to a non-interactive APA. The results supported this hypothesis only partially. Learners who interacted with the interactive reported significantly higher levels of engagement, as measured by the User Engagement Scale (UES), particularly in terms of focused attention and perceived usability. This finding aligns with previous research highlighting the potential of interactive pedagogical agents to foster learner engagement (e.g., Liew et al., 2017; Brandão et al., 2022). The APA’s ability to respond to learners' queries and provide contextualized support (APA know where the learner is when he asks something) likely contributed to a more immersive and engaging learning experience. Interestingly, while engagement was higher in the interactive condition, there were no significant differences in learners' motivation between the two conditions. This suggests that the enhanced interactivity may not directly translate into increased intrinsic motivation or identified regulation. This finding deviates from some previous studies that have reported positive effects of pedagogical agents on learner motivation (e.g., Liew et al., 2017; Hobert & Meyer von Wolff, 2019). The lack of significant motivational differences in our study might be attributed to the specific design and role of the interactive, as well as the relatively short duration of the learning intervention. Further research is needed to explore the nuances of how different interactive designs and roles influence various facets of learner motivation over extended periods.

The subjective perception of time (SPT) emerged as a noteworthy measure of engagement in our study. Learners in the interactive condition significantly underestimated the time spent on the e-learning platform compared to those in the control condition. This temporal distortion effect, known as "flow" (Csikszentmihalyi, 1991), suggests a state of deep absorption and engagement in the learning activity. The negative correlations between SPT and the focused attention and perceived usability dimensions of the UES further support this interpretation. These findings highlight the value of incorporating SPT as a complementary measure of engagement in multimedia learning research (O'Brien & Lomas, 2017; Yang et al., 2021). Interestingly, both the inhibition difficulties (ID) and resource management difficulties (RD) indices predicted learners' subjective perception of time. Learners with greater inhibition difficulties tended to underestimate the time spent, possibly because they were more aware of the distractors diverting their attention away from the relevant information, making them feel like they had less time to focus on what was pertinent. Conversely, learners who had difficulty managing cognitive resources may have experienced the learning activity as more effortful and time-consuming, leading them to overestimate the time spent on the e-learning material. When learners struggle to distribute their attentional resources effectively among relevant information, they may feel like they are wasting time, resulting in an overestimation of the time spent on the learning task.

These results align with recent findings by Chen and Yeh (2022), who demonstrated that individual differences in attentional capacities can significantly influence learners' engagement and perceived cognitive load in multimedia learning environments. Learners with lower attentional capacities may be more susceptible to distraction and struggle to manage the cognitive demands of the learning task, leading to a distorted perception of time. As O'Brien and Lomas (2017) argue, incorporating measures of attentional capacity and cognitive load alongside traditional engagement metrics can provide a more comprehensive understanding of learners' experiences in multimedia learning.

As mentioned, contrary to our expectations, learning outcomes did not significantly differ between the interactive and APA conditions. In fact, learners in the control APA condition performed slightly better on memorization questions compared to those in the interactive condition. This finding challenges the assumption that increased interactivity and embodiment necessarily lead to improved learning performance. It is possible that the additional cognitive load imposed by interacting with the interactive (e.g., formulating questions, processing responses) might have detracted from the cognitive resources available for encoding and retaining the core learning content (Adams et al., 2021). This interpretation aligns with the cognitive load theory (Sweller, 1988) and the potential for interactive elements to introduce extraneous cognitive load. Learners' inhibition difficulties (ID index) negatively predicted their ability to discriminate between true and false statements (d' index), suggesting that attentional control plays a crucial role in learning performance. This finding resonates with previous research linking attentional capacities to multimedia learning outcomes (e.g., Mayer & Fiorella, 2014; Szulewski et al., 2021). However, we did not observe significant interactions between attentional capacities and the APA condition in predicting learning outcomes or cognitive load. This suggests that the effects of the interactive on learning may be relatively consistent across learners with varying attentional abilities.

It is important to address a significant bias observed in the learning performance results. The overall performance on the learning tasks was notably lower compared to our previous studies using the same learning material. The student sample in this study differed slightly from the samples in our earlier investigations. The participants in the current study were business school students who were recruited during their classes and had to leave the classroom to participate. This difference in the sample characteristics and the experimental conditions may have introduced a considerable bias, potentially affecting the learning outcomes.

Furthermore, it is worth noting that the learners in this study exhibited generally low levels of motivation, which could partially explain the lower performance scores. The motivation measures indicated that the participants were not highly motivated to engage with the learning material, possibly due to the experimental context or the specific sample characteristics. This lack of motivation may have hindered their ability to fully invest in the learning process, leading to suboptimal performance outcomes.

Consequently, it would be necessary to replicate this study with a more specific target audience, under more controlled conditions, and with a sample that demonstrates higher levels of intrinsic motivation to draw definitive conclusions about the effects of interactive APAs on learning performance. Future research should aim to control for the potential confounding effects of motivation and ensure that the experimental design fosters a more engaging and motivating learning environment.

In conclusion, our study provides some insights into the relationship between attention and engagement in the context of advanced, interactive, and embodied Animated and Conversational Pedagogical Agents (ACPAs) in multimedia learning environments. The findings suggest that ACPAs have the potential to enhance learner engagement, particularly in terms of focused attention and perceived usability. However, the effects on motivation and learning outcomes appear to be more complex and may depend on various factors, such as the specific design and role of the ACPA, the characteristics of the learners, and the nature of the learning task. It is important to acknowledge that the present study has several limitations that prevent us from drawing definitive conclusions. The observed bias in learning performance, likely stemming from the differences in the sample characteristics and experimental conditions compared to our previous studies, raises concerns about the generalizability and reliability of the findings. Additionally, the lack of a more representative sample and fully controlled experimental setting limits our ability to make strong inferences about the effects of interactive on learning outcomes and engagement. Given these limitations, we didn't publish this research in its current form. Instead, we suggest replicating the study under more controlled conditions, with a sample that is representative of the target population. Future research should aim to minimize potential confounding factors and ensure that the experimental setting is optimized for assessing the impact of interactive pedagogical agent on learning and engagement.

E. References

Adams, D. M., Reilly, M., Sosu, E. M., & McLaughlin, P. J. (2021). Cognitive load theory, multimedia learning, and instructional design: A state of the art review. Journal of Computer

Assisted Learning, 37(6), 1715-1732.
Brandão, L. O., Nascimento, J. O., Brandão, A. A., & Brandão, G. O. (2022). The impact of pedagogical agents on learning outcomes: A meta-analysis. Computers & Education, 176, 104353.
Chen, C.-J., & Yeh, K.-H. (2022). Designing pedagogical agents for fostering students' engagement and learning outcomes in online learning: The moderating role of learner characteristics. Computers & Education, 183, 104539.
Csikszentmihalyi, M. (1991). Flow: The psychology of optimal experience. Harper Perennial.
Fontaine, S., Hoareau, C., & Dubois, M. (2019). Validation française de l'échelle de motivation situationnelle (SIMS) et de l'échelle d'engagement de l'utilisateur (UES). Psychologie Française, 64(2), 121-137.
Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). Motivation and Emotion, 24(3), 175-213.
Hancock, P. A. (2019). Time and time again. The American Journal of Psychology, 132(1), 1-25.
Hobert, S., & Meyer von Wolff, R. (2019). Say hello to your new automated tutor – A structured literature review on pedagogical conversational agents. Wirtschaftsinformatik 2019 Proceedings.
Liew, T. W., Zin, N. A. M., & Sahari, N. (2017). Exploring the affective, motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment. Human-centric Computing and Information Sciences, 7(1), 9.
Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (pp. 31-48). Cambridge University Press.
Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press.
Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. Journal of Experimental Psychology: Applied, 18(3), 239-252.
Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (2nd ed., pp. 279-315). Cambridge University Press.
Michael, G. A., Dorey, J. M., Rey, R., & Cabon, P. (2020). Master Activation Map and attentional resources management. Neuropsychologia, 142, 107459.
Michael, G. A., Mizzi, R., Couffe, C., & Gálvez-García, G. (2014). Dissociable yet tied

inhibitory processes: The structure of inhibitory control. Cognitive, Affective, & Behavioral Neuroscience, 14(3), 1026-1040.

Michael, G. A., Pannetier, C., Maljkovic, V., Brunel, L., Dormal, V., & Van der Linden, M. (2007). The Triple Master Activation Map: A neutral substrate for the assessment of attentional abilities. Clinical Neuropsychology, 21(6), 1022-1026.

O'Brien, H. L., & Cairns, P. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. International Journal of Human-Computer Studies, 112, 28-39.

O'Brien, H. L., & Lomas, T. (2017). Developing and evaluating a mindfulness-based program for self-regulated learning: A mixed-methods investigation. International Journal of Technology in Education and Science, 1(1), 46-60.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. Journal of the American Society for Information Science and Technology, 61(1), 50-69.

Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people and places. Cambridge University Press.

Salgues, S., Michael, G. A., Leman, S., & Baccino, T. (2023). How the Master Activation Map predicts attentional capacity during a lecture: An event-related potentials study. Brain Sciences, 13(3), 395.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257-285.

Szulewski, A., Braund, H., Egan, R., & Gegenfurtner, A. (2021). A critical narrative review of the relationship between cognitive load and clinical reasoning in medical education. Advances in Health Sciences Education, 26(2), 783-802.

Yang, X., Zhang, S., Zhang, H., & Liang, Y. (2021). Investigating the impact of learning behavior engagement on learning outcomes and subjective time perception in game-based learning using facial recognition technology. Computers & Education, 176, 104362.

Discussion générale

A. Synthèse

L'objectif central de cette thèse était d'examiner comment les caractéristiques des environnements d'apprentissage multimédia, telles que l'anthropomorphisme, le degré d'interactivité ou la présence d'agents pédagogiques, interagissent avec les caractéristiques des apprenants, notamment leurs capacités attentionnelles, pour influencer leur engagement, leur motivation, leur charge cognitive et leurs performances d'apprentissage.

Pour répondre à cette problématique, nous avons dans un premier temps proposé un modèle intégratif original articulant les notions d'attention et d'engagement dans le cadre de l'apprentissage multimédia (Chapitre 2). Ancré dans la théorie cognitive de l'apprentissage multimédia de Mayer (2005), mais intégrant aussi les apports des théories de l'engagement (Fredricks et al., 2004) et des modèles cognitifs de l'attention (Michael et al., 2006), ce modèle vise à capturer la dynamique complexe des états attentionnels et motivationnels des apprenants en fonction des caractéristiques des supports d'apprentissage et de leurs propres ressources cognitives.

Nous avons ensuite mis ce modèle à l'épreuve à travers trois études empiriques manipulant différentes caractéristiques de design des environnements multimédia et mesurant leurs effets sur un ensemble de variables d'engagement, de motivation, d'attention, de charge cognitive et de performance.

La première étude (Chapitre 3) s'est intéressée à l'impact de l'anthropomorphisme, montrant qu'un environnement d'apprentissage présentant des attributs visuels humains améliorait l'engagement, la motivation et la mémorisation, sans surcoût cognitif, nous permettant également de construire un protocole de mesure de l’engagement permettant de mesurer finement cet état.

La deuxième étude (Chapitre 4) a examiné les effets de la présence et du degré d'expressivité d'un agent pédagogique sur l'engagement et les performances, révélant des effets différenciés selon le type de performance (compréhension vs. mémorisation) et les capacités attentionnelles individuelles.

Enfin, la troisième étude (Chapitre 5) a testé l'apport de l'intelligence artificielle permettant une interaction plus poussée avec un agent pédagogique, soulignant l'intérêt de cette approche pour soutenir l'engagement des apprenants, mais aussi les limites de l'étude elle-même en termes de charge cognitive et de performance d'apprentissage. La synthèse et la discussion de ces différents résultats nous permettra de dégager les principaux apports empiriques et théoriques de la thèse quant aux déterminants de l'engagement et de l'attention en situation d'apprentissage multimédia. Nous verrons en quoi ils valident certaines prédictions de notre modèle intégratif tout en le précisant et le complexifiant. Nous en tirerons des implications pratiques pour le design d'environnements multimédias optimisant l'engagement et la gestion des ressources attentionnelles des apprenants. Les limites des présents travaux et les nouvelles questions qu'ils soulèvent seront aussi discutées, ouvrant des perspectives de recherche intégrant encore d'avantage les dernières avancées en matière d'intelligence artificielle et d'adaptativité des environnements d'apprentissage.

B. Implications théoriques

Nos travaux s'inscrivent dans le cadre des théories de l'apprentissage multimédia, en particulier la théorie cognitive de l'apprentissage multimédia de Mayer (2005) et son extension, la théorie cognitive-affective de l'apprentissage multimédia de Moreno (2006). Ces théories mettent en avant le rôle clé de l'attention et de l'engagement dans l'efficacité de l'apprentissage multimédia. Cependant, elles n'explicitent pas pleinement les relations entre ces deux concepts, ni leur articulation avec les caractéristiques des environnements d'apprentissage et des apprenants.

C'est précisément l'objectif du modèle intégratif de l'engagement et de l'attention que nous avons proposé (Chapitre 2). Ce modèle postule que l'engagement situationnel des apprenants résulte d'une interaction dynamique entre les caractéristiques des supports d'apprentissage (niveau d'interactivité, présence d'agents pédagogiques, etc.), les caractéristiques des apprenants (capacités attentionnelles, motivation, etc.) et les caractéristiques de la tâche (complexité, type de contenu, etc.).

1. L'engagement

Nos résultats expérimentaux valident partiellement les prédictions de notre modèle intégratif. À travers nos trois études, nous avons observé que des supports d'apprentissage plus riches sur le plan de l'anthropomorphisme (Étude 1), ou intégrant des agents pédagogiques expressifs (Étude 2) et interactifs (Étude 3) suscitent un engagement accru des apprenants. Cet engagement se manifeste notamment par une attention plus soutenue, captée par la dimension "Focused Attention" de l'échelle UES, ainsi que par un sentiment d'immersion plus marqué, reflété par la mesure de perception subjective du temps (SPT). Les apprenants rapportent également une appréciation esthétique et une satisfaction plus positive envers ces environnements d'apprentissage enrichis, comme en témoignent leurs scores plus élevés sur les dimensions "Aesthetic Appeal" et "Reward" de l'UES.

Plusieurs mécanismes peuvent être invoqués pour expliquer ces effets. Tout d'abord, l'utilisation d'éléments anthropomorphiques et d'APAs, en personnifiant l'interface d'apprentissage, tend à la rendre plus attractive et familière pour les apprenants (Mayer & DaPra, 2012). La présence de visages souriants, de personnages sympathiques ou d'animations amusantes capte l'attention et suscite des réactions émotionnelles positives (Um et al., 2012). Cette première réaction affective et attentionnelle positive agit comme un déclencheur de l'engagement, incitant les apprenants à interagir plus volontiers avec le support.

Ensuite, les APAs expressifs et interactifs, en mimant des comportements communicatifs naturels (regards, gestes, intonations), créent une illusion de présence sociale et de conversation avec l'apprenant (Fiorella et al., 2019). Cette perception d'un "dialogue" avec un partenaire d'apprentissage, même virtuel, renforce l'engagement en donnant aux apprenants le sentiment d'être guidés, encouragés et soutenus dans leur démarche (Mayer, 2014). Les rétroactions positives et les marques d'empathie exprimées par les APAs contribuent à maintenir la motivation et l'implication tout au long de l'activité.

De plus, les APAs interactifs, capables de répondre aux questions et d'adapter leurs interventions aux actions des apprenants, renforcent le sentiment de contrôle et d'autonomie, facteur clé de l'engagement (Ryan & Deci, 2000). En donnant aux apprenants la possibilité d'influencer le déroulement de l'interaction, les APAs interactifs favorisent une posture active et réflexive, propice à un apprentissage engagé (Moreno et al., 2001).

Enfin, sur le plan cognitif, les APAs bien conçus peuvent servir de guides attentionnels en attirant le regard des apprenants vers les informations essentielles, en posant des questions ciblées ou en fournissant des résumés intermédiaires (Louwerse et al., 2005). Ce guidage attentionnel permet de réduire la charge cognitive et de libérer des ressources pour un traitement plus profond du contenu, ce qui favorise en retour l'engagement cognitif (Mayer, 2014).

Ces différents constats débouchent sur plusieurs préconisations pour la conception de supports d'apprentissage engageants. Il s'agit d'abord de veiller à la qualité graphique et esthétique de l'interface, en utilisant des éléments de design émotionnel (couleurs, formes, visages) pour capter l'attention et susciter des émotions positives. L'intégration d'APAs expressifs, dotés de mimiques faciales, de gestes et d'une voix naturelle, permet de créer une impression de présence sociale et de soutien. Le développement d'APAs interactifs, capables de dialoguer avec les apprenants et de fournir des rétroactions adaptées, renforce l'engagement en donnant un sentiment de contrôle et de progression. Enfin, une réflexion sur le rôle pédagogique des APAs, conçus comme de véritables partenaires d'apprentissage plutôt que comme de simples avatars, permet d'optimiser leur potentiel de guidage attentionnel et cognitif.

Cependant, nos manipulations expérimentales n'ont pas eu d'impact significatif sur la dimension "Perceived Usability" de l'UES, et ce de façon constante à travers nos trois études. Cette observation suggère que l'engagement suscité par les choix de design étudiés (anthropomorphisme, expressivité et interactivité des agents pédagogiques) relève davantage d'une captation attentionnelle et émotionnelle des apprenants que d'une amélioration perçue de l'utilisabilité et de la prise en main du système d'apprentissage.

Ces résultats rejoignent l'idée d'envisager l'engagement comme un construit multifacette (Fredricks et al., 2004), dont les différentes dimensions peuvent être influencées de façon distincte par les caractéristiques de design des environnements d'apprentissage. Certaines propriétés, comme l'anthropomorphisme ou la présence d'agents pédagogiques, semblent agir principalement sur les aspects attentionnels (focused attention), émotionnels (aesthetic appeal) et expérientiels (reward) de l'engagement (O'Brien & Toms, 2008). En revanche, elles n'affectent pas nécessairement la perception de l'utilisabilité du système, qui renvoie à des aspects plus pragmatiques et fonctionnels de l'interaction (Zaharias & Poylymenakou, 2009).

Cette disparité pourrait s'expliquer par le fait que les leviers de design testés dans nos études visent prioritairement à capter l'intérêt et la curiosité des apprenants, à créer une expérience d'apprentissage immersive et plaisante (Mayer, 2014 ; Salar et al., 2020), mais ne modifient pas fondamentalement les modalités de navigation et de contrôle de l'interface. Ainsi, l'ajout d'éléments anthropomorphiques ou d'agents pédagogiques, même expressifs et interactifs, ne suffit pas à lui seul à améliorer le sentiment de prise en main et d'efficience dans l'utilisation du système (Davis & Antonenko, 2017 ; Kalyuga & Liu, 2015).

Ces constats nous amènent à affiner notre modèle intégratif en distinguant plus clairement les facettes expérientielle et instrumentale de l'engagement. Ils suggèrent que les choix de design axés sur l'anthropomorphisme et la présence sociale agissent principalement sur la dimension expérientielle, en stimulant l'attention, les émotions et le plaisir d'apprendre (Schroeder & Gotch, 2015). Pour influencer la dimension instrumentale, relative à l'utilisabilité perçue, il conviendrait sans doute d'explorer d'autres pistes de conception davantage centrées sur l'optimisation des interactions humain-machine (ergonomie, personnalisation des parcours, etc.) (Angeli, Howard et al., 2017 ; Zaharias, 2009).

À la lumière de ces réflexions, nous proposons d'enrichir notre modèle intégratif en y intégrant explicitement cette distinction entre les facettes expérientielle et instrumentale de l'engagement (Heidig et al., 2015). Concrètement, cela impliquerait de représenter l'engagement non plus comme un construit unidimensionnel, mais comme le produit de l'interaction dynamique entre deux sous-dimensions complémentaires :

- Une dimension expérientielle, qui renverrait à la qualité de l'expérience subjective vécue par l'apprenant (attention, émotions, plaisir), et qui serait prioritairement influencée par des choix de design visant à susciter l'intérêt, la curiosité et l'immersion (anthropomorphisme, présence sociale, esthétique, etc.) (Choi et al., 2017; Plass & Kaplan, 2016).
- Une dimension instrumentale, qui concernerait la perception de l'utilisabilité, de l'efficience et de la contrôlabilité du système, et qui dépendrait davantage de choix de conception centrés sur l'optimisation ergonomique et fonctionnelle des interactions (architecture de l'information, personnalisation, aide à la navigation, etc.) (Lee & Wong, 2014).

L'engagement "optimal" résulterait alors d'une articulation réussie entre ces deux facettes, permettant à l'apprenant de vivre une expérience d'apprentissage à la fois cognitivement et émotionnellement stimulante (dimension expérientielle) et fluide, efficiente sur le plan de l'utilisation du système (dimension instrumentale) (O'Brien & Cairns, 2016 ; Satar & Akcan, 2018). Différents paramètres de conception pourraient être actionnés pour équilibrer et optimiser ces deux dimensions en fonction des contextes et des profils d'apprenants (Keller & Suzuki, 2004 ; Paas & Sweller, 2014).

Cette distinction entre les dimensions expérientielle et instrumentale de l'engagement pourrait s'expliquer par le contexte spécifique de production des environnements d'apprentissage utilisés dans nos études. En effet, nos expérimentations ont été menées en collaboration avec l'entreprise Sydo, spécialisée dans la création de contenus e-learning. Grâce à cette collaboration, nous avons pu bénéficier de l'expertise d'une équipe pluridisciplinaire composée notamment de designers graphiques, d'illustrateurs et de motion designers. Cette configuration nous a permis de développer des supports d'apprentissage particulièrement soignés sur le plan esthétique et visuel, mettant en œuvre les principes du design émotionnel (Um et al., 2012 ; Plass et al., 2014). L'accent a été mis sur la création d'interfaces attractives, d'illustrations expressives et d'animations dynamiques, visant à capter l'attention et susciter des réactions affectives positives chez les apprenants.

Perception du temps et engagement

Un apport original de notre travail réside dans l'étude du lien entre engagement et perception subjective du temps (SPT). Dans nos trois études, nous avons observé une forte corrélation entre les scores d'engagement (UES) et la sous-estimation du temps passé sur la tâche. Plus les apprenants se déclaraient engagés, plus ils avaient l'impression que le temps passait vite. Cette distorsion temporelle était particulièrement liée aux dimensions "Focused Attention" et "Aesthetic Appeal" de l'UES.

Ces résultats suggèrent que la SPT pourrait être un indicateur pertinent des états de concentration intense et d'immersion, caractéristiques des moments de fort engagement (Csikszentmihalyi, 1990). Ils font écho aux travaux montrant que les expériences optimales (flow) s'accompagnent d'une perception accélérée du temps (Agarwal & Karahanna, 2000). Sur le plan méthodologique, ils soulignent l'intérêt de la SPT comme mesure écologique de l'engagement, complémentaire aux échelles auto-rapportées (Baldauf et al., 2009).

Nos analyses corrélationnelles ont également révélé des liens intéressants entre les capacités attentionnelles des apprenants (mesurées par les indices ID et RD du MAM) et leur engagement. Dans l'étude 2, les apprenants avec de plus grandes difficultés à gérer le partage attentionnel (RD élevé) ont rapporté des niveaux d'engagement plus élevés. Dans l'étude 3, l'indice RD prédisait positivement la sous-estimation du temps (SPT).

Ces résultats, bien que préliminaires, suggèrent que les apprenants avec une attention plus "focalisée" pourraient paradoxalement ressentir un engagement plus intense, peut-être en raison d'une plus grande réactivité aux stimuli captivants (Derryberry & Reed, 2002). Ils invitent à considérer les capacités attentionnelles non seulement comme modérateurs mais aussi comme antécédents directs de l'engagement dans notre modèle.

Nos trois études ont systématiquement inclus une mesure de la perception subjective du temps passé sur la tâche d'apprentissage, en complément des mesures auto-rapportées de l'engagement. Ce choix méthodologique s'inscrit dans la lignée des travaux récents suggérant que la distorsion de la perception temporelle peut être un marqueur pertinent des états de concentration intense et d'immersion dans une activité, caractéristiques des moments de fort engagement (Csikszentmihalyi, 1990 ; Engeser & Rheinberg, 2008). Traditionnellement, la SPT a surtout été utilisée dans la littérature comme un indicateur indirect de la charge cognitive (Block, Hancock & Zakay, 2010). Selon ce cadre, une tâche cognitivement exigeante, mobilisant fortement les ressources attentionnelles, conduirait à une sous-estimation du temps passé, car moins de ressources seraient disponibles pour le suivi temporel. Inversement, une tâche peu coûteuse se traduirait par une surestimation du temps, l'attention pouvant se porter davantage sur le passage du temps. Quelques études ont ainsi montré que des tâches complexes conduisent à une compression de la perception temporelle (Fink & Neubauer, 2001 ; Luthman, Bliesener & Staude-Müller, 2009). Cependant, des travaux plus récents dans le champ des expériences optimales (flow) ont mis en évidence que la distorsion temporelle pouvait aussi refléter un état d'engagement et d'absorption intense dans une activité, indépendamment de sa difficulté objective (Agarwal & Karahanna, 2000 ; Engeser & Rheinberg, 2008). Lorsque l'engagement est profond, l'attention est entièrement focalisée sur la tâche, au point de perdre la notion du temps qui passe. Cette altération de la perception du temps a été proposée comme une des dimensions de l'expérience optimale de flow, au même titre que le sentiment de contrôle et l'absence de préoccupation de soi (Nakamura & Csikszentmihalyi, 2009).

Dans ce contexte, nos résultats apportent un éclairage nouveau en montrant que, dans des situations d'apprentissage multimédia, la SPT est corrélée à l'intensité de l'engagement rapporté, et ce indépendamment de la charge cognitive perçue. Dans nos trois études, les apprenants ayant rapporté des niveaux d'engagement élevés sur l'échelle UES (User Engagement Scale) ont aussi significativement sous-estimé le temps passé sur l'activité. Plus précisément, la sous-estimation temporelle était fortement corrélée aux dimensions de l'attention focalisée, de l'usabilité perçue et de l'esthétique de l'UES. Une méta-analyse récente de Hancock (2019) a souligné l'intérêt de la perception subjective du temps (SPT) comme un indicateur non invasif et écologiquement valide de l'état de flow.

Ces constats suggèrent que, plus que la simple charge mentale, cela serait la captation attentionnelle et l'immersion dans l'interaction avec l'environnement d'apprentissage qui semble conduire à une distorsion de la perception du temps. Lorsque l'apprenant est pleinement engagé, absorbé par une interface qu'il juge à la fois stimulante, facile à utiliser et visuellement attractive, il tend à perdre la notion du temps, signe d'une expérience optimale.

Notre dernière étude (Chapitre 5) apporte un éclairage complémentaire en révélant des liens spécifiques entre les capacités attentionnelles des apprenants, mesurées par les indices ID (Inhibition Difficulty) et RD (Resource Management Difficulty) du MAM, et leur perception subjective du temps (SPT). Nous avons observé que les apprenants présentant de plus grandes difficultés à gérer le partage attentionnel en situation de double tâche (indice RD élevé), mais aussi ceux ayant de meilleures capacités d'inhibition des distracteurs (indice ID faible), tendaient à sous-estimer davantage le temps passé sur l'activité d'apprentissage.

Ces résultats suggèrent que la propension à une focalisation attentionnelle intense, qu'elle résulte d'une difficulté à désengager son attention ou d'une capacité à résister aux interférences, favorise une perception du temps accélérée, caractéristique des états d'engagement profond. Ils font écho aux travaux montrant que les individus avec un fort niveau d'absorption cognitive rapportent souvent une expérience altérée du temps lors d'activités captivantes (Agarwal & Karahanna, 2000). Sur le plan théorique, ces observations invitent à considérer les capacités attentionnelles individuelles comme des antécédents dispositions des états d'engagement optimal, au même titre que les caractéristiques des environnements d'apprentissage. Sur le plan méthodologique, ces observations confirment l'intérêt de l'utilisation de la mesure de la SPT en complément des mesures auto-rapportées (Baldauf et al., 2009 ; Schoenau-Fog, 2011).

En somme, l'étude du lien entre perception subjective du temps et engagement constitue une contribution originale de notre travail au champ de la recherche sur l'apprentissage multimédia. Elle ouvre de nouvelles perspectives, tant méthodologiques que théoriques, pour comprendre et façonner des expériences d'apprentissage optimales.

2. Motivation

Nos études ont également examiné l'impact des caractéristiques des supports d'apprentissage sur la motivation des apprenants, en utilisant l'échelle multidimensionnelle SIMS (Situational Motivation Scale, Guay et al., 2000). Cette échelle permet de distinguer différents types de motivation selon leur degré d'autodétermination : la motivation intrinsèque, la régulation identifiée, la régulation externe et l'amotivation.

Dans notre première étude portant sur l'anthropomorphisme (Chapitre 3), nous avons observé que la version riche du support, intégrant des éléments visuels anthropomorphiques, suscitait des niveaux plus élevés de régulation identifiée et plus faibles d'amotivation et de régulation externe, comparativement à la version neutre. Ces résultats suggèrent que l'anthropomorphisme, s'il ne renforce pas directement la motivation intrinsèque, favorise néanmoins des formes de motivation plus autodéterminées. Les apprenants semblent percevoir davantage la valeur et la pertinence personnelle de l'activité d'apprentissage lorsque l'interface présente des caractéristiques humanisées.

En revanche, dans nos études sur les agents pédagogiques (Chapitres 4 et 5), nous n'avons pas répliqué ces effets sur la motivation autodéterminée. La présence d'un agent pédagogique, qu'il soit expressif ou interactif, n'a pas entraîné de différences significatives sur les différentes dimensions de la motivation par rapport à une version sans agent. Ce résultat contraste avec certaines recherches antérieures ayant mis en évidence un impact positif des agents pédagogiques sur la motivation intrinsèque des apprenants (Liew et al., 2017 ; Park et al., 2015).

Plusieurs interprétations peuvent être avancées pour expliquer cette absence d'effet. Il est possible que l'influence motivationnelle des agents pédagogiques soit plus sensible à des variations fines de leur design, comme leur réalisme visuel ou leur proactivité (Baylor & Kim, 2005), dimensions que nous n'avons pas fait varier dans nos études. Il se peut également que l'impact des agents sur la motivation nécessite une interaction plus prolongée et personnalisée pour se manifester pleinement, là où nos expérimentations se déroulaient sur une durée relativement courte avec des agents non personnalisables.

Quoi qu'il en soit, ces résultats nuancent l'idée d'un effet robuste et systématique des agents pédagogiques sur la motivation des apprenants. Ils invitent à explorer plus finement les conditions d'efficacité motivationnelle de ces dispositifs, en examinant notamment l'influence de leurs caractéristiques spécifiques (apparence, comportements, rôle pédagogique) et des facteurs contextuels (durée d'interaction, personnalisation).

Au-delà des effets des supports, nos études ont mis en évidence des liens intéressants entre motivation et engagement. Dans le chapitre 3, nous avons observé une corrélation négative entre le score d'engagement et le niveau de régulation externe, ainsi qu'une médiation de l'effet de la régulation externe sur les performances d'apprentissage par l'engagement. Ces résultats suggèrent que la motivation extrinsèque peut nuire à l'engagement et, indirectement, aux performances, soulignant l'importance de soutenir des formes de motivation plus autodéterminées.

En somme, nos résultats sur la motivation offrent un tableau nuancé, avec des effets différenciés selon les types de supports et de mesures motivationnelles utilisés. S'ils confirment globalement l'intérêt d'un design anthropomorphique pour favoriser un engagement motivationnel dans l'activité, ils questionnent la généralisation de cet effet aux agents pédagogiques et appellent à poursuivre les recherches pour mieux cerner leurs conditions d'efficacité. Nos observations sur les liens entre motivation, engagement et performances soulignent quant à elles la pertinence d'une approche intégrative pour comprendre la dynamique motivationnelle à l'œuvre dans l'apprentissage multimédia.

3. L'attention

Nos études ont également mis en évidence le rôle clé des capacités attentionnelles individuelles dans la prédiction des performances d'apprentissage, au-delà de leur impact sur l'engagement situationnel. Plus précisément, nous avons observé que les apprenants présentant de meilleures capacités d'inhibition des distracteurs obtenaient de meilleurs scores de mémorisation et de compréhension, indépendamment des caractéristiques des supports d'apprentissage (Étude 2 et 3). Ces résultats convergent avec un large corpus de recherches en psychologie cognitive et en neurosciences montrant que les capacités de contrôle attentionnel, et notamment d'inhibition, sont des prédicteurs robustes de la réussite dans diverses tâches d'apprentissage (Unsworth & Spillers, 2010 ; Stevens & Bavelier, 2012). Les individus capables de résister à l'interférence des distracteurs et de maintenir leur focus attentionnel sur les informations pertinentes sont plus à même d'encoder et de manipuler ces informations en mémoire de travail, favorisant ainsi leur compréhension et leur rétention à long terme (Conway, Cowan & Bunting, 2001 ; Fukuda & Vogel, 2011).

Ces résultats convergent avec un large corpus de recherches en psychologie cognitive et en neurosciences montrant que les capacités de contrôle attentionnel, et notamment d'inhibition, sont des prédicteurs robustes de la réussite dans diverses tâches d'apprentissage (Unsworth & Spillers, 2010 ; Stevens & Bavelier, 2012).

Cette relation entre capacités attentionnelles et performances d'apprentissage a été particulièrement étudiée chez l'enfant. Des études longitudinales ont montré que les compétences attentionnelles précoces, telles que l'attention soutenue et le contrôle inhibiteur, prédisent significativement la réussite académique ultérieure (Steele et al., 2012). Par exemple, Steele et al. (2012) ont suivi une cohorte d'enfants de la maternelle jusqu'à la fin de l'école élémentaire et ont constaté que ceux qui présentaient de meilleures capacités d'attention soutenue et de contrôle inhibiteur à 5 ans obtenaient de meilleurs résultats scolaires à 11 ans, et ce même en contrôlant le niveau initial de compétences. De même, Rhoades et al. (2011) ont montré que l'attention exécutive (incluant l'inhibition) mesurée à 4 ans prédisait les performances en lecture et en mathématiques à 8 ans.

Ces résultats soulignent l'importance des capacités attentionnelles comme fondement de l'apprentissage, dès les premières années de scolarité. Ils suggèrent que les enfants capables de maintenir leur attention sur les tâches scolaires et de résister aux distractions sont plus à même de s'engager dans les activités d'apprentissage, d'acquérir les connaissances et compétences enseignées, et donc de réussir académiquement à long terme.

En transposant ces résultats à l'apprentissage multimédia chez l'adulte, nos études confirment la persistance de l'influence des capacités attentionnelles sur les performances d'apprentissage, même à des âges plus avancés. Elles montrent que les individus capables de résister à l'interférence des distracteurs et de maintenir leur focus attentionnel sur les informations pertinentes sont plus à même d'encoder et de manipuler ces informations en mémoire de travail, favorisant ainsi leur compréhension et leur rétention à long terme (Conway, Cowan & Bunting, 2001 ; Fukuda & Vogel, 2011).

Dans le contexte spécifique de l'apprentissage multimédia, plusieurs études ont montré que les apprenants avec de meilleures capacités d'attention sélective étaient plus aptes à discriminer les éléments visuels pertinents des détails superflus, et par conséquent à élaborer des modèles mentaux plus cohérents du matériel pédagogique (Hegarty & Kriz, 2008 ; Jarodzka, Scheiter, Gerjets & van Gog, 2010). Nos résultats sur le lien entre l'inhibition et les performances de mémorisation et de compréhension s'inscrivent dans la lignée de ces travaux.

Sur le plan théorique, ces constats renforcent l'idée que les capacités attentionnelles constituent une ressource cognitive fondamentale pour l'apprentissage, au même titre que les connaissances préalables ou les stratégies métacognitives (Cowan, 2014 ; Mayer, 2014). Ils invitent à considérer les différences attentionnelles inter-individuelles comme une variable explicative centrale dans les modèles de compréhension et de mémorisation, aux côtés des caractéristiques des apprenants plus classiquement étudiées comme la motivation ou les styles d'apprentissage.

4. Attention et engagement

Nos études mettent en lumière le rôle des capacités attentionnelles individuelles sur l'engagement. Ainsi, nous avons montré que les apprenants présentant de meilleures capacités d'inhibition des distracteurs et de gestion des ressources attentionnelles en situation de double tâche étaient plus à même de maintenir leur engagement et leurs performances face à des supports pédagogiques visuellement riches ou interactifs (Étude 2). À l'inverse, les apprenants avec de moindres capacités attentionnelles se sont avérés plus vulnérables aux effets délétères potentiels de ces mêmes caractéristiques, notamment en termes de charge cognitive et de mémorisation.

Nos analyses ont montré que les apprenants présentant un score élevé à l'indice RD, reflétant une plus grande difficulté à partager leurs ressources attentionnelles en situation de double tâche, rapportaient paradoxalement des niveaux d'engagement plus élevés sur les dimensions d'attention focalisée, d'appréciation esthétique et de satisfaction (Étude 2). Ce pattern s'est vu confirmé dans notre étude sur les agents pédagogiques interactifs (Étude 3), où l'indice de gestion des ressources prédisait positivement la perception subjective du temps, autre marqueur d'un engagement profond. Ces observations font écho à certains travaux antérieurs ayant suggéré qu'une propension à la focalisation attentionnelle, même si elle peut rendre plus difficile le traitement simultané de multiples informations, favorise néanmoins une expérience subjective d'immersion et de flow dans les activités (Csikszentmihalyi, 1997 ; Agarwal & Karahanna, 2000). Les individus avec une forte tendance à "l'absorption cognitive" rapportent ainsi souvent un engagement plus intense et gratifiant dans les tâches captivant leur attention (Agarwal & Karahanna, 2000).

De même, la recherche sur les différences individuelles dans le fonctionnement attentionnel a montré que les individus avec un contrôle attentionnel plus faible peuvent paradoxalement ressentir leur engagement dans les tâches de façon plus marquée, en raison d'une plus grande réactivité aux stimuli de l'environnement (Derryberry & Reed, 2002). Un parallèle peut être fait avec le concept de "présence" dans les environnements virtuels, où une allocation plus exclusive des ressources attentionnelles à l'environnement simulé (au détriment de la prise en compte du contexte réel) est associée à un sentiment plus fort d'engagement et d'immersion (Wirth et al., 2007).

Dans le contexte éducatif, quelques études ont mis en évidence un lien entre certains traits de personnalité associés à une focalisation attentionnelle, comme l'absorption ou la propension à la rêverie, et l'intensité de l'engagement ressenti dans les activités d'apprentissage (Agarwal & Karahanna, 2000 ; Boyle et al., 2016). Les apprenants avec ces dispositions rapportent souvent un engagement et un intérêt plus marqués lorsque les contenus pédagogiques captent leur attention. Le revers étant qu'ils peuvent aussi être plus enclins à se "perdre" dans des détails ou des réflexions parallèles au détriment de l'apprentissage (Hove & Corcoran, 2008). Ces différents constats entrent en résonance avec nos résultats montrant qu'une plus grande difficulté à désengager son attention et à inhiber les informations non pertinentes, si elle peut ponctuellement nuire aux performances, favorise néanmoins une expérience d'engagement plus intense dans les situations d'apprentissage. Ils confirment donc le rôle clé des capacités de contrôle attentionnel dans la régulation de l'engagement des apprenants, au-delà même des caractéristiques des supports.

Sur le plan théorique, ces résultats invitent à enrichir les modèles existants de l'engagement en y intégrant plus explicitement les différences attentionnelles comme antécédents directs de l'expérience d'engagement, et non seulement comme modérateurs de l'impact des facteurs contextuels (Kahu, 2013 ; Christenson et al., 2012). Ils plaident pour une prise en compte de la diversité des profils attentionnels des apprenants dans la compréhension de leurs trajectoires d'engagement.

5. Performances d'apprentissage

L'un des objectifs centraux de nos travaux était d'examiner l'impact des caractéristiques de design des supports d'apprentissage multimédia sur les performances des apprenants, en distinguant deux types de mesures de l'apprentissage : la mémorisation et la compréhension. En effet, si ces deux processus sont souvent liés, ils renvoient à des mécanismes cognitifs distincts (Mayer, 2014). La mémorisation correspond à la capacité à encoder, stocker et rappeler des informations factuelles, tandis que la compréhension implique une transformation des informations acquises pour construire une représentation mentale cohérente et applicable à de nouvelles situations (Kintsch, 1994). Nos études visaient donc à clarifier comment différents choix de design, tels que le degré d'anthropomorphisme (Étude 1) ou la présence et les caractéristiques d'un agent pédagogique (Études 2 et 3), pouvaient affecter différentiellement ces deux types d'apprentissage.

Dans notre première étude portant sur l'anthropomorphisme (Chapitre 3), nous avons observé que les participants ayant utilisé la version anthropomorphique du support obtenaient de meilleurs scores de performances peu importe le type de question, comparativement à ceux ayant utilisé la version neutre. Ce résultat suggère que la présence d'éléments visuels humanisés a facilité l'encodage et la rétention des détails du contenu, potentiellement en les rendant plus concrets et distinctifs (Mayer, 2014).

Nos deux études suivantes, portant sur les agents pédagogiques animés (Chapitres 4 et 5), ont mis en évidence des résultats plus nuancés. Dans l'étude 2, la présence d'un APA, qu'il soit expressif ou non, s'est avérée délétère pour les performances de mémorisation comparativement à une condition sans agent. Ce résultat peut s'interpréter à la lumière de la théorie de la charge cognitive (Sweller, 1994) : l'ajout d'un élément visuel supplémentaire tel qu'un APA, même peu expressif, a pu introduire une charge extrinsèque interférant avec les processus d'encodage et de rétention. Cette interprétation est corroborée par l'interaction observée entre les capacités attentionnelles des apprenants et leurs performances de mémorisation : les participants présentant de moindres capacités d'inhibition étaient particulièrement affectés par la présence de l'APA, suggérant une plus grande sensibilité à l'interférence générée.

En revanche, l'étude 2 a révélé un effet différentiel du degré d'expressivité de l'APA sur les performances de compréhension. Si la version peu expressive de l'agent a conduit à de moins bonnes performances que la condition sans APA, la version très expressive a au contraire amélioré la compréhension. Ce résultat peut s'expliquer par le rôle de guidage attentionnel et de soutien à l'intégration des informations joué par les comportements non-verbaux de l'agent (gestes, expressions faciales). En attirant l'attention des apprenants sur les éléments clés et en véhiculant des indices sur les relations entre les concepts, l'APA expressif a pu faciliter la construction d'un modèle mental cohérent (Mayer, 2014). Cet effet bénéfique ne s'est pas retrouvé pour l'APA peu expressif, suggérant un seuil minimal d'expressivité nécessaire pour soutenir les processus de compréhension.

Enfin, notre troisième étude (Chapitre 5) visait à étendre ces résultats en comparant un APA "traditionnel" (similaire à la version très expressive de l'étude 2) à un APA "nouvelle génération", capable d'interagir en langage naturel grâce à modèle d'intelligence artificielle. Nos hypothèses étaient que l'APA interactif, en permettant un dialogue personnalisé et un feedback adaptatif, favoriserait encore davantage la compréhension. Cependant, nous n'avons pas observé de différences significatives entre les deux types d'APA, tant pour la mémorisation que pour la compréhension. Ce résultat peut s'expliquer par le fait que les interventions de l'APA interactif, bien que plus flexibles, restaient limitées aux informations présentes dans le support. Ainsi, le contenu d'apprentissage étant identique dans les deux conditions, les gains liés à l'interactivité ont pu être trop ténus pour se traduire par une amélioration mesurable des performances.

Pris ensemble, ces résultats offrent plusieurs enseignements sur les relations entre design des supports d'apprentissage multimédia et performances d'apprentissage. Tout d'abord, ils suggèrent que les choix de design influencent différemment les processus de mémorisation et de compréhension. Si des éléments de surface tels que l'anthropomorphisme semblent pouvoir faciliter la mémorisation factuelle, potentiellement en augmentant la saillance et la distinctivité des informations, leur impact sur la compréhension apparaît plus limité. À l'inverse, des choix de design plus profonds, tels que le degré d'expressivité des APAs, semblent davantage influencer la compréhension, probablement en soutenant les processus d'intégration et de construction de modèles mentaux. Ensuite, nos résultats mettent en lumière l'importance de considérer les caractéristiques des apprenants, notamment leurs capacités attentionnelles, pour prédire l'effet des choix de design. Les apprenants présentant de moindres ressources attentionnelles semblent plus sensibles aux effets délétères d'éléments de design non optimaux, tels que des APA peu expressifs, particulièrement en termes de mémorisation. Ce constat souligne la nécessité d'une approche différentielle pour concevoir des supports d'apprentissage adaptés aux profils cognitifs des apprenants.

Enfin, nos travaux interrogent le degré optimal de richesse et d'interactivité des supports d'apprentissage multimédia. Si l'ajout d'éléments de design tels que l'anthropomorphisme ou les APA peut, dans certaines conditions, améliorer les performances d'apprentissage, nos résultats suggèrent également l'existence de seuils au-delà desquels les gains deviennent marginaux, voire s'inversent en raison d'une surcharge cognitive. La troisième étude illustre notamment ce phénomène : le supplément de complexité et d'interactivité apporté par l'APA nouvelle génération n'a pas conduit à de meilleures performances, potentiellement car le coût cognitif du traitement de ces informations supplémentaires a contrebalancé les bénéfices attendus.

D'un point de vue théorique, ces résultats enrichissent les modèles de l'apprentissage multimédia en soulignant l'influence différenciée des caractéristiques de design sur les processus mnésiques et de compréhension. Ils invitent à affiner ces modèles en y intégrant des variables relatives aux apprenants, telles que le contrôle attentionnel, pour prédire l'efficacité des choix de conception. Ils questionnent également l'hypothèse d'une relation linéaire entre la richesse des supports et les gains d'apprentissage, suggérant plutôt l'existence d'un niveau optimal au-delà duquel les bénéfices stagnent, voire déclinent.

En somme, nos études contribuent à une meilleure compréhension des mécanismes par lesquels les choix de design des supports d'apprentissage multimédia influencent les processus mnésiques et de compréhension. Elles ouvrent de nouvelles pistes pour optimiser ces choix en fonction des objectifs d'apprentissage et des caractéristiques des apprenants. Elles appellent également à poursuivre les recherches pour mieux cerner les conditions d'une ingénierie pédagogique alliant efficacement le design émotionnel et cognitif au service de l'apprentissage.

C. Implications pratiques

Sur le plan pratique, ces considérations ouvrent plusieurs pistes pour la conception d'environnements d'apprentissage multimédia engageants et cognitivement efficients. Nos résultats plaident pour un usage raisonné et ciblé des leviers d'engagement tels que l'anthropomorphisme ou les agents pédagogiques. S'ils peuvent indéniablement stimuler l'intérêt et l'implication des apprenants, leur implémentation doit être finement articulée aux objectifs pédagogiques et aux prérequis du contenu à apprendre. Ainsi, l'usage d'agents très expressifs ou interactifs pourra s'avérer pertinent pour des tâches de compréhension ou de résolution de problèmes complexes, où l'engagement joue un rôle clé. En revanche, pour des tâches de mémorisation d'informations factuelles, une présentation plus épurée et directive sera sans doute préférable pour focaliser l'attention.

D'un point de vue pratique, ces observations ouvrent des pistes intéressantes pour le design d'environnements d'apprentissage adaptés aux besoins attentionnels des apprenants. Il s'agirait de proposer des formats pédagogiques et des stratégies de guidage différenciés en fonction des capacités de contrôle attentionnel, afin d'optimiser les conditions d'engagement de chacun. Par exemple, les apprenants avec une forte propension à la focalisation attentionnelle pourraient bénéficier de séquences pédagogiques plus courtes et rythmées, limitant les risques de "sur-immersion" dans des détails. À l'inverse, les apprenants avec de plus fortes capacités d'attention divisée pourraient être orientés vers des formats plus riches et multidimensionnels, entretenant leur engagement par la variété des stimulations. Plus généralement, ces résultats plaident pour le développement d'outils de diagnostic des profils attentionnels qui pourraient être intégrés en amont ou au fil des parcours d'apprentissage afin d'en personnaliser l'expérience sur le plan de l'engagement. Des approches de plus en plus "attentionnellement informées", basées sur le suivi en temps réel des dynamiques d'engagement (par exemple par des mesures de suivi oculaire ou de temps de réponse), ouvrent à cet égard des perspectives prometteuses (D'Mello, 2021).

À terme, et ce fut notamment l'objectif de notre troisième étude en intégrant l'intelligence artificielle et l'adaptative learning, nous pourrions imaginer le développement d'environnements d'apprentissage "attentionnellement et motivationnellement intelligents", capables de détecter en temps réel les fluctuations de l'engagement et les signaux de surcharge cognitive pour adapter dynamiquement le type et l'intensité des supports pédagogiques. Les technologies d'IA conversationnelle, de suivi oculaire ou de reconnaissance émotionnelle ouvrent d'ores et déjà des possibilités prometteuses en ce sens (Dewan et al., 2019). Couplées à des modèles computationnels de plus en plus fins de l'engagement et de l'attention, nourris par l'exploitation des traces d'apprentissage (learning analytics), elles dessinent les contours d'une nouvelle génération d'environnements d'apprentissage numériques pleinement centrés sur l'apprenant et ses besoins.

Conclusion

Ce travail de thèse visait à mieux comprendre les déterminants de l'engagement et de l'attention dans l'apprentissage multimédia, en examinant l'influence des caractéristiques des environnements numériques et des différences individuelles des apprenants sur ces processus clés pour la réussite des apprentissages en ligne.

En proposant un modèle intégratif original articulant les cadres théoriques de la psychologie cognitive, des sciences de l'éducation et de l'ergonomie des interfaces, et en le mettant à l'épreuve à travers une série d'études empiriques, nous avons pu dégager plusieurs résultats novateurs. Ceux-ci concernent notamment l'impact différencié des choix de design (anthropomorphisme, présence et expressivité des agents pédagogiques, interactivité) sur différentes facettes de l'expérience d'apprentissage (engagement situationnel, charge cognitive, mémorisation, compréhension). Ils mettent aussi en lumière le rôle clé des capacités attentionnelles des apprenants et le lien avec leur engagement et la gestion de leurs ressources cognitives face à ces choix de design.

Sur le plan théorique, ces résultats enrichissent les modèles existants de l'apprentissage multimédia en y intégrant plus explicitement les notions d'engagement et d'attention, non comme de simples outputs mais comme des processus façonnés par l'interaction entre les caractéristiques des apprenants et celles des supports. Ils invitent à reconsidérer l'engagement non seulement comme un levier de motivation et de persévérance, mais aussi comme un état attentionnel et expérientiel spécifique, caractérisé notamment par une perception altérée du temps. Ce faisant, ils ouvrent de nouvelles pistes pour penser les liens entre émotions, cognition et expérience utilisateur dans les environnements numériques d'apprentissage.

Sur le plan méthodologique, nos études démontrent l'intérêt d'une approche intégrative et multidimensionnelle de l'évaluation de l'engagement et de l'attention, combinant mesures objectives (performances, données comportementales) et subjectives (échelles, entretiens). Elles soulignent en particulier la pertinence d'indicateurs peu coûteux comme la perception subjective du temps pour capturer les états d'engagement profond. Elles offrent ainsi des outils adaptés aux besoins croissants d'évaluation écologique et située de l'expérience d'apprentissage.

D'un point de vue appliqué, ce travail ouvre des perspectives prometteuses pour la conception d'environnements d'apprentissage numériques plus engageants et adaptés aux besoins des apprenants. Il suggère des principes de design finement articulés aux objectifs pédagogiques et aux prérequis attentionnels et motivationnels des contenus. Il invite au développement d'interfaces "attentives", capable d'identifier en temps réel les dynamiques d'engagement et d'adapter en conséquence le guidage pédagogique. Plus largement, il plaide pour une personnalisation croissante de l'expérience d'apprentissage prenant en compte la diversité des profils cognitifs et des trajectoires d'engagement des apprenants.

Les limites de ce travail, notamment en termes de validité écologique et de généralisation des résultats, constituent autant de pistes pour les recherches futures. Il importera de répliquer nos observations sur des temps plus longs, des contenus plus diversifiés et des publics plus larges, afin d'en éprouver la robustesse. L'examen des liens entre engagement situationnel et engagement à long terme, ou entre capacités attentionnelles et développement des compétences nécessitera des approches longitudinales. De même, l'exploitation croisée des traces d'interaction et des données d'expérience utilisateur permettra d'affiner notre compréhension des dynamiques d'engagement et d'attention à une granularité plus fine.

Enfin, les perspectives ouvertes par l'intelligence artificielle et l'analytique des apprentissages invitent à repenser en profondeur les contours d'une ingénierie de l'engagement et de l'attention en contexte éducatif. Le développement de systèmes d'apprentissage « affectivement et cognitivement intelligents », capables d'optimiser en temps réel les parcours de formation en fonction des états mentaux des apprenants est une piste à explorer. Il appelle un dialogue interdisciplinaire renouvelé entre sciences cognitives, sciences de l'éducation et informatique autour des enjeux de l'apprentissage.

En définitive, j'espère, dans cette thèse, avoir contribué à une meilleure compréhension des ressorts de l'engagement et de l'attention dans les environnements d'apprentissage numériques. Au-delà des réponses apportées, elle soulève de nouvelles questions sur la façon dont les technologies transforment en profondeur l'expérience d'apprendre.

Bibliographie

Abdullah, A., Adil, M., Rosenbaum, L., Clemmons, M., Shah, M., Abrahamson, D., & Neff, M. (2017). Pedagogical Agents to Support Embodied, Discovery-Based Learning. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Éds.), *Intelligent Virtual Agents* (p. 1-14). Springer International Publishing. https://doi.org/10.1007/978-3-319-67401-8_1

Agarwal, R., & Karahanna, E. (2000). Time Flies When You're Having Fun : Cognitive Absorption and Beliefs about Information Technology Usage. *MIS Quarterly*, *24*(4), 665-694. https://doi.org/10.2307/3250951

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory* (p. xiv, 524). V. H. Winston & Sons.

Ando, A., Smallwood, D., McMahon, M., Irving, L., Mazzone, S. B., & Farrell, M. J. (2016). Neural correlates of cough hypersensitivity in humans : Evidence for central sensitisation and dysfunctional inhibitory control. *Thorax*, *71*(4), 323-329. https://doi.org/10.1136/thoraxjnl-2015-207425

Angeli, C., Howard, S. K., Ma, J., Yang, J., & Kirschner, P. A. (2017). Data mining in educational technology classroom research : Can it make a contribution? *Computers & Education*, *113*, 226-242. https://doi.org/10.1016/j.compedu.2017.05.021

Appleton, J., Christenson, S., & Furlong, M. (2008). Student engagement with school : Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*, 369-386. https://doi.org/10.1002/pits.20303

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school : Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*(5), 369-386. https://doi.org/10.1002/pits.20303

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *13*(3), 214-228. https://doi.org/10.1177/1073858407299288

ASTLEITNER, H., & Wiesner, C. (2004). An Integrated Model of Multimedia Learning and Motivation. *Journal of Educational Multimedia and Hypermedia*, *13*.

Ayres, P., & Sweller, J. (2005). The Split-Attention Principle in Multimedia Learning. In *The Cambridge handbook of multimedia learning* (p. 135-146). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.009

Baddeley, A. (2003). Working memory : Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839. https://doi.org/10.1038/nrn1201

Baldauf, D., Burgard, E., & Wittmann, M. (2009). Time perception as a workload measure in simulated car driving. *Applied Ergonomics*, *40*(5), 929-935. https://doi.org/10.1016/j.apergo.2009.01.004

Banich, M., Milham, M., Atchley, R., Cohen, N., Webb, A., Wszalek, T., Kramer, A., Liang, Z.-P., Wright, A., Shenker, J., & Magin, R. (2000). fMRI Studies of Stroop Tasks Reveal Unique Roles of Anterior and Posterior Brain Systems in Attentional Selection. *Journal of cognitive neuroscience*, *12*, 988-1000. https://doi.org/10.1162/08989290051137521

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions : Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65-94. https://doi.org/10.1037/0033-2909.121.1.65

Beege, M., & Schneider, S. (2023). Emotional design of pedagogical agents : The influence of enthusiasm and model-observer similarity. *Educational technology research and development*, *71*. https://doi.org/10.1007/s11423-023-10213-4

Bempechat, J., & Shernoff, D. J. (2012). Parental Influences on Achievement Motivation and Student Engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Éds.), *Handbook of Research on Student Engagement* (p. 315-342). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_15

Ben-Eliyahu, A., & Linnenbrink-Garcia, L. (2013). Extending self-regulated learning to include self-regulated emotion strategies. *Motivation and Emotion*, *37*(3), 558-573. https://doi.org/10.1007/s11031-012-9332-3

Bingham, G. E., & Okagaki, L. (2012). Ethnicity and student engagement. In *Handbook of research on student engagement* (p. 65-95). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-2018-7_4

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments : A meta-analytic review. *Acta Psychologica*, *134*(3), 330-343. https://doi.org/10.1016/j.actpsy.2010.03.006

Bouvier, P., Lavoué, E., & Sehaba, K. (2014). Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming. *Simulation & Gaming*, *45*(4-5), 491-507. https://doi.org/10.1177/1046878114553571

Brault-Labbé, A., & Dubé, L. (2010). Engagement scolaire, bien-être personnel et autodétermination chez des étudiants à l'université. [School engagement, personal well-being and self-determination in university students.]. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, *42*(2), 80-92. https://doi.org/10.1037/a0017385

Brazzolotto, P., & Michael, G. A. (2019a). Interrupting an e-mail search : Influence of the complexity and the timing of the interruption. *Cognition, Brain, Behavior: An Interdisciplinary Journal*, *23*(2), 135-153. https://doi.org/10.24193/cbb.2019.23.08

Brazzolotto, P., & Michael, G. A. (2019b). Interrupting an e-mail search : Influence of the complexity and the timing of the interruption. *Cognition, Brain, Behavior: An Interdisciplinary Journal*, *23*(2), 135-153. https://doi.org/10.24193/cbb.2019.23.08

Brazzolotto, P., & Michael, G. A. (2020). Complexity of interruptions : Evidence supporting a non-interruption-based theory. *Scandinavian Journal of Psychology*, *61*(6), 723-730. https://doi.org/10.1111/sjop.12659

Broadbent, D. E. (2013). *Perception and Communication*. Elsevier.

Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review*, *25*, 100-119. https://doi.org/10.1016/j.edurev.2018.09.004

Burton, K. D., Lydon, J. E., D'Alessandro, D. U., & Koestner, R. (2006). The differential effects of intrinsic and identified motivation on well-being and performance : Prospective, experimental, and implicit approaches to self-determination theory. *Journal of Personality and Social Psychology*, *91*(4), 750-762. https://doi.org/10.1037/0022-3514.91.4.750

Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., & Paas, F. (2021). Effectiveness of multimedia pedagogical agents predicted by diverse theories : A meta-analysis. *Educational Psychology Review*, *33*(3), 989-1015. https://doi.org/10.1007/s10648-020-09587-1

Chalmers, P. A. (2000). User interface improvements in computer-assisted instruction, the challenge. *Computers in Human Behavior*, *16*, 507-517. https://doi.org/10.1016/S0747-5632(00)00022-4

Chen, A., Darst, P. W., & Pangrazi, R. P. (2001). An examination of situational interest and its sources. *British Journal of Educational Psychology*, *71*(3), 383-400. https://doi.org/10.1348/000709901158578

Cheon, S. H., & Reeve, J. (2015). A classroom-based intervention to help teachers decrease students' amotivation. *Contemporary Educational Psychology*, *40*, 99-111. https://doi.org/10.1016/j.cedpsych.2014.06.004

Chin, K.-Y., Hong, Z.-W., Huang, Y.-M., Shen, W.-W., & Lin, J.-M. (2016). Courseware development with animated pedagogical agents in learning system to improve learning motivation. *Interactive Learning Environments*, *24*(3), 360-381. https://doi.org/10.1080/10494820.2013.851089

Choi, W., Dyens, O., Chan, T., Schijven, M., Lajoie, S., Mancini, M., Dev, P., Felländer-Tsai, L., Ferland, M., Kato, P., Lau, J., Montanaro, M., Pineau, J., & Aggarwal, R. (2017). Engagement and learning in simulation : Recommendations of the Simnovate Engaged Learning Domain Group. *BMJ Simulation and Technology Enhanced Learning*, *3*, S23-S32. https://doi.org/10.1136/bmjstel-2016-000177

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement* (p. xxvii, 840). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-2018-7

Chi M, Glaser R, Rees E. Expertise in problem solving. In: Sternberg R, editor. Advances in the psychology of human intelligence. Hillsdale, NJ: Erlbaum; 1982. p. 7-75.

Cleary, T. J., & Zimmerman, B. J. (2001). Self-regulation differences during athletic practice by experts, non-experts, and novices. *Journal of Applied Sport Psychology*, *13*(2), 185-206. https://doi.org/10.1080/104132001753149883

Cleary, T. J., & Zimmerman, B. J. (2012). A Cyclical Self-Regulatory Account of Student Engagement : Theoretical Foundations and Applications. In S. L. Christenson, A. L. Reschly, & C. Wylie (Éds.), *Handbook of Research on Student Engagement* (p. 237-257). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_11

Conway, A., Cowan, N., & Bunting, M. (2001). The cocktail party phenomenon revisited : The importance of working memory capacity. *Psychonomic bulletin & review*, *8*, 331-335. https://doi.org/10.3758/BF03196169

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited : The importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*(2), 331-335. https://doi.org/10.3758/bf03196169

Couffe, C., & Michael, G. A. (2017a). Failures due to interruptions or distractions : A review and a new framework. *The American Journal of Psychology*, *130*(2), 163-181. https://doi.org/10.5406/amerjpsyc.130.2.0163

Couffe, C., & Michael, G. A. (2017b). Failures due to interruptions or distractions : A review and a new framework. *The American Journal of Psychology*, *130*(2), 163-181. https://doi.org/10.5406/amerjpsyc.130.2.0163

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational psychology review*, *26*(2), 197-223. https://doi.org/10.1007/s10648-013-9246-y

Csikszentmihalyi, M. (1990). *Flow : The Psychology of Optimal Experience*.

Csikszentmihalyi, M. (2000). *FLOW: The Psychology of Optimal Experience*.

Csikszentmihalyi, M., & Rathunde, K. (1993). The measurement of flow in everyday life : Toward a theory of emergent motivation. In *Nebraska Symposium on Motivation, 1992 : Developmental perspectives on motivation* (p. 57-97). University of Nebraska Press.

Davis, R., & Antonenko, P. (2017). Effects of pedagogical agent gestures on social acceptance and learning : Virtual real relationships in an elementary foreign language classroom. *Journal of Interactive Learning Research*, *28*(4), 459-480.

Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments : A meta-analysis. *Educational Research Review*, *24*, 193-209. https://doi.org/10.1016/j.edurev.2018.05.002

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, *111*(2), 225-236. https://doi.org/10.1037/0021-843X.111.2.225

Dewan, M. A. A., Murshed, M., & Lin, F. (2019). Engagement detection in online learning : A review. *Smart Learning Environments*, *6*(1), 1. https://doi.org/10.1186/s40561-018-0080-z

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285-290. https://doi.org/10.1037/h0033731

Dunsworth, Q., & Atkinson, R. K. (2007). Fostering multimedia learning of science : Exploring the role of an animated agent's image. *Computers & Education*, *49*(3), 677-690. https://doi.org/10.1016/j.compedu.2005.11.010

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning : The MASRL model. *Educational Psychologist*, *46*(1), 6-25. https://doi.org/10.1080/00461520.2011.538645

Engeser, S., & Rheinberg, F. (2008a). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, *32*(3), 158-172. https://doi.org/10.1007/s11031-008-9102-4

Engeser, S., & Rheinberg, F. (2008b). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, *32*(3), 158-172. https://doi.org/10.1007/s11031-008-9102-4

Farley, J., Risko, E., & Kingstone, A. (2013). Everyday attention and lecture retention : The effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00619

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting Clicks is Not Enough : Validating a Theorized Model of Engagement in Learning Analytics. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 501-510. https://doi.org/10.1145/3303772.3303775

Fincham, F. D., Rogge, R., & Beach, S. R. H. (2018). Relationship satisfaction. In *The Cambridge handbook of personal relationships, 2nd ed* (p. 422-436). Cambridge University Press. https://doi.org/10.1017/9781316417867.033

Fink, A., & Neubauer, A. C. (2001). Speed of information processing, psychometric intelligence : And time estimation as an index of cognitive load. *Personality and Individual Differences*, *30*(6), 1009-1021. https://doi.org/10.1016/S0191-8869(00)00089-1

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2019). Instructor presence in video lectures : The role of dynamic drawings, eye contact, and instructor visibility. *Journal of Educational Psychology*, *111*(7), 1162-1171. https://doi.org/10.1037/edu0000325

Folk, C. L., Ester, E. F., & Troemel, K. (2009). How to keep attention from straying : Get engaged! *Psychonomic Bulletin & Review*, *16*(1), 127-132. https://doi.org/10.3758/PBR.16.1.127

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement : Potential of the Concept, State of the Evidence. *Review of Educational Research*, *74*(1), 59-109. https://doi.org/10.3102/00346543074001059

Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science*, *22*(3), 361-368. https://doi.org/10.1177/0956797611398493

Gaytan, J. (2015). Comparing Faculty and Student Perceptions Regarding Factors That Affect Student Retention in Online Education. *American Journal of Distance Education*, *29*(1), 56-66.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation : Bridging selective attention and working memory. *Trends in cognitive sciences*, *16*(2), 129-135. https://doi.org/10.1016/j.tics.2011.11.014

Gazzaniga, M. S. (2005). Forty-five years of split-brain research and still going strong. *Nature Reviews. Neuroscience*, *6*(8), 653-659. https://doi.org/10.1038/nrn1723

GEARY, D. C. (2008). Whither Evolutionary Educational Psychology? *Educational Psychologist*, *43*(4), 217-226. https://doi.org/10.1080/00461520802392240

Ginns, P., Martin, A., & Marsh, H. (2013). Designing Instructional Text in a Conversational Style : A Meta-analysis. *Educational Psychology Review*, *25*, 445-472. https://doi.org/10.1007/s10648-013-9228-0

Grawemeyer, B., Mavrikis, M., Holmes, W., Gutiérrez-Santos, S., Wiedmann, M., & Rummel, N. (2017). Affective learning : Improving engagement and enhancing learning with affect-aware feedback. *User Modeling and User-Adapted Interaction*, *27*(1), 119-158. https://doi.org/10.1007/s11257-017-9188-z

Gray, P. (2016). *Libre pour apprendre*. Actes Sud Nature.

Guay, F., Vallerand, R., & Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation : The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*, 175-213. https://doi.org/10.1023/A:1005614228250

Hancock, P. A., Kaplan, A. D., Cruit, J. K., Hancock, G. M., MacArthur, K. R., & Szalma, J. L. (2019). A meta-analysis of flow effects and the perception of time. *Acta Psychologica*, *198*, 102836. https://doi.org/10.1016/j.actpsy.2019.04.007

Harnishfeger, K. K., & Bjorklund, D. F. (1994). A developmental perspective on individual differences in inhibition. *Learning and Individual Differences*, *6*(3), 331-355. https://doi.org/10.1016/1041-6080(94)90021-3

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations : On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, *89*(1), 92-102. https://doi.org/10.1037/0022-0663.89.1.92

Hegarty, M., & Kriz, S. (2008). Effects of knowledge and spatial ability on learning from animation. In *Learning with animation : Research implications for design* (p. 3-29). Cambridge University Press.

Heidig, S., & Clarebout, G. (2011a). Do pedagogical agents make a difference to student motivation and learning? A review of empirical research. *Educational Research Review*, *6*, 27-54. https://doi.org/10.1016/j.edurev.2010.07.004

Heidig, S., & Clarebout, G. (2011b). Do pedagogical agents make a difference to student motivation and learning? A review of empirical research. *Educational Research Review*, *6*, 27-54. https://doi.org/10.1016/j.edurev.2010.07.004

Heidig, S., Müller, J., & Reichelt, M. (2015). Emotional design in multimedia learning : Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior*, *44*. https://doi.org/10.1016/j.chb.2014.11.009

Heidig, S., Schwartz, R., & Plass, J. (2010). Interactivity in multimedia learning : An integrated model. *Computers in Human Behavior*, *26*, 1024-1033. https://doi.org/10.1016/j.chb.2010.03.003

Henrie, C., Halverson, L., & Graham, C. (2015). Measuring Student Engagement in Technology-Mediated Learning : A Review. *Computers & Education*, *90*, 36-53. https://doi.org/10.1016/j.compedu.2015.09.005

Hidi, S., & Baird, W. (1986). Interestingness—A neglected variable in discourse processing. *Cognitive Science*, *10*(2), 179-194.

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, *41*(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4

Höffler, T. N., Prechtl, H., & Nerdel, C. (2010). The influence of visual cognitive style when learning from instructional animations and static pictures. *Learning and Individual Differences*, *20*(5), 479-483. https://doi.org/10.1016/j.lindif.2010.03.001

Hove, M. C., & Corcoran, K. J. (2008). Educational Technologies : Impact on Learning and Frustration. *Teaching of Psychology*, *35*(2), 121-125. https://doi.org/10.1080/00986280802004578

Huang, S. Y. B., Huang, C.-H., & Chang, T.-W. (2022). A New Concept of Work Engagement Theory in Cognitive Engagement, Emotional Engagement, and Physical Engagement. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.663440

Ivala, E., Gachago, D., Condy, J., & Chigona, A. (2013). Enhancing Student Engagement with Their Studies : A Digital Storytelling Approach. *Creative Education*, *4*(10), Article 10. https://doi.org/10.4236/ce.2013.410A012

James, W. (1890). The perception of reality. In *The principles of psychology, Vol II* (p. 283-324). Henry Holt and Company. https://doi.org/10.1037/11059-005

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder : How experts and novices interpret dynamic stimuli. *Learning and Instruction*, *20*(2), 146-154. https://doi.org/10.1016/j.learninstruc.2009.02.019

Johnson, W. L., & Lester, J. C. (2016). Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later. *International Journal of Artificial Intelligence in Education*, *26*(1), 25-36. https://doi.org/10.1007/s40593-015-0065-9

Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A. C., & Brooks, C. (2018). How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. *Review of Educational Research*, *88*(1), 43-86. https://doi.org/10.3102/0034654317740335

Kahneman, D. (1973). *Attention and Effort*. Prentice Hall.

Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, *38*(5), 758-773. https://doi.org/10.1080/03075079.2011.598505

Kalyuga, S., & Liu, T.-C. (2015). Guest Editorial : Managing Cognitive Load in Technology-Based Learning Environments. *Journal of Educational Technology & Society*, *18*(4), 1-8.

Keller, J., & Suzuki, K. (2004). Learner motivation and E-learning design : A multinationally validated process. *Learning, Media and Technology*, *29*, 229-239. https://doi.org/10.1080/1358165042000283084

Kim, Y., & Baylor, A. L. (2006). A Social-Cognitive Framework for Pedagogical Agents as Learning Companions. *Educational Technology Research and Development*, *54*(6), 569-596.

Kintsch, W. (1988). The role of knowledge in discourse comprehension : A construction-integration model. *Psychological Review*, *95*(2), 163-182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294-303. https://doi.org/10.1037/0003-066X.49.4.294

Kirschner, F., Kester, L., & Corbalan, G. (2011). Cognitive load theory and multimedia learning, task characteristics and learning engagement : The Current State of the Art. *Computers in Human Behavior*, *27*(1), 1-4. https://doi.org/10.1016/j.chb.2010.05.003

Kizilcec, R., Pérez-Sanagustín, M., & Maldonado, J. (2016). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses (In press). *Computers & Education*, *104*. https://doi.org/10.1016/j.compedu.2016.10.001

Kok, A. (1999). Varieties of inhibition : Manifestations in cognition, event-related potentials and aging. *Acta Psychologica*, *101*(2-3), 129-158. https://doi.org/10.1016/s0001-6918(99)00003-7

Krapp, A., & Prenzel, M. (2011). Research on Interest in Science : Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27-50. https://doi.org/10.1080/09500693.2010.518645

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology. General*, *133*(3), 339-354. https://doi.org/10.1037/0096-3445.133.3.339

Lavoué, É., & Serna, A. (2021). Vers une ludification adaptative expressive des environnements numériques d'apprentissage. *STICEF (Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation)*, *28*(2), 153-176.

Lee, E. A.-L., & Wong, K. W. (2014). Learning with desktop virtual reality : Low spatial ability learners are more positively affected. *Computers & Education*, *79*, 49-58. https://doi.org/10.1016/j.compedu.2014.07.010

Lee, Y., & Choi, J. (2011). A review of online course dropout research : Implications for practice and future research. *Educational Technology Research and Development*, *59*(5), 593-618. https://doi.org/10.1007/s11423-010-9177-y

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect : Affective impact of animated pedagogical agents. *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, 359-366. https://doi.org/10.1145/258549.258797

L'état de l'École (2021). Ministère de l'Education Nationale et de la Jeunesse.

Liew, T. W., Mat Zin, N. A., & Sahari, N. (2017). Exploring the affective, motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment.

Human-Centric Computing and Information Sciences, *7*(1), 9. https://doi.org/10.1186/s13673-017-0089-2

Linnenbrink, E. A., & Pintrich, P. R. (2003). THE ROLE OF SELF-EFFICACY BELIEFS INSTUDENTENGAGEMENTANDLEARNING INTHECLASSROOM. *Reading & Writing Quarterly*, *19*(2), 119-137. https://doi.org/10.1080/10573560308223

Louwerse, M., Graesser, A., Lu, S., & Mitchell, H. (2005a). Social cues in animated conversational agents. *Applied Cognitive Psychology*, *19*, 693-704. https://doi.org/10.1002/acp.1117

Louwerse, M., Graesser, A., Lu, S., & Mitchell, H. (2005b). Social cues in animated conversational agents. *Applied Cognitive Psychology*, *19*, 693-704. https://doi.org/10.1002/acp.1117

Louwerse, M. M., Graesser, A. C., Lu, S., & Mitchell, H. H. (2005). Social cues in animated conversational agents. *Applied Cognitive Psychology*, *19*(6), 693-704. https://doi.org/10.1002/acp.1117

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents : Does their degree of embodiment impact learning from static or animated work examples? *Applied Cognitive Psychology*, *21*(6), 747-764. https://doi.org/10.1002/acp.1347

Luthman, S., Bliesener, T., & Staude-Müller, F. (2009). The Effect of Computer Gaming on Subsequent Time Perception. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *3*(1), Article 1. https://cyberpsychology.eu/article/view/4221

Martha, A. S. D., & Santoso, H. B. (2019). The Design and Impact of the Pedagogical Agent : A Systematic Literature Review. *Journal of Educators Online*, *16*(1). https://eric.ed.gov/?id=EJ1204376

Martin, A. J., Ginns, P., & Papworth, B. (2017). Motivation and engagement : Same or different? Does it matter? *Learning and Individual Differences*, *55*, 150-162. https://doi.org/10.1016/j.lindif.2017.03.013

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In *The Cambridge handbook of multimedia learning* (p. 31-48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E. (2009). *Multimedia learning, 2nd ed* (p. xiii, 304). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning, 2nd ed.* (p. 43-71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E. (2021). Cognitive Theory of Multimedia Learning. In L. Fiorella & R. E. Mayer (Éds.), *The Cambridge Handbook of Multimedia Learning* (3e éd., p. 57-72). Cambridge University Press. https://doi.org/10.1017/9781108894333.008

Mayer, R. E., & DaPra, C. S. (2012a). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied*, *18*(3), 239-252. https://doi.org/10.1037/a0028616

Mayer, R. E., & DaPra, C. S. (2012b). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied*, *18*(3), 239-252. https://doi.org/10.1037/a0028616

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, *33*, 12-18. https://doi.org/10.1016/j.learninstruc.2014.02.004

Mayer, R. E., & Jackson, J. (2005). The Case for Coherence in Scientific Explanations : Quantitative Details Can Hurt Qualitative Understanding. *Journal of Experimental Psychology: Applied*, *11*(1), 13-18. https://doi.org/10.1037/1076-898X.11.1.13

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning : Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312-320. https://doi.org/10.1037/0022-0663.90.2.312

Mayer, R. E., & Moreno, R. (2001). Animation as an Aid to Multimedia Learning. *Educational Psychology Review*, 13.

Mayer, R., & Fiorella, L. (2014). *Principles for Reducing Extraneous Processing in Multimedia Learning : Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles* (p. 279-315). https://doi.org/10.1017/CBO9781139547369.015

Mayer, R., & Pilegard, C. (2014). *Principles for managing essential processing in multimedia learning : Segmenting, pre-training, and modality principles* (p. 316-344). https://doi.org/10.1017/CBO9781139547369.016

Mekni, M. (2021). An Artificial Intelligence Based Virtual Assistant Using Conversational Agents. *Journal of Software Engineering and Applications*, *14*(09), 455-473. https://doi.org/10.4236/jsea.2021.149027

Michael, G. (2015). Disturbances of selective attention in traumatic brain injury and schizophrenia : What's common and what's different? *Psychologie Française*, *60*, 387-402.

Michael, G. A., & Couffe, C. (2018). Chapitre 21. Apprendre à être et à rester attentif. In *Psychologie cognitive des apprentissages scolaires* (p. 289-306). Dunod. https://doi.org/10.3917/dunod.ferra.2018.01.0289

Michael, G. A., Dorey, J.-M., Rey, R., D'Amato, T., Fabre, D., Brunet, S., & Padovan, C. (2020). Attention in schizophrenia : Impaired inhibitory control, faulty attentional resources, or both? *Psychiatry Research*, *290*, 113164. https://doi.org/10.1016/j.psychres.2020.113164

Michael, G. A., Garcia, S., Fernandez, D., Sellal, F., & Boucart, M. (2006). The ventral premotor cortex (vPM) and resistance to interference. *Behavioral Neuroscience*, *120*(2), 447-462. https://doi.org/10.1037/0735-7044.120.2.447

Michael, G. A., Kleitz, C., Sellal, F., Hirsch, E., & Marescaux, C. (2001). Controlling attentional priority by preventing changes in oculomotor programs : A job for the premotor cortex? *Neuropsychologia*, *39*(10), 1112-1120. https://doi.org/10.1016/S0028-3932(01)00029-X

Michael, G. A., Lété, B., & Ducrot, S. (2013). Trajectories of attentional development : An exploration with the master activation map model. *Developmental Psychology*, *49*(4), 615-631. https://doi.org/10.1037/a0028410

Michael, G. A., Mizzi, R., Couffe, C., & Gálvez-García, G. (2014). Dissociable yet tied inhibitory processes : The structure of inhibitory control. *Cognitive, Affective & Behavioral Neuroscience*, *14*(3), 1026-1040. https://doi.org/10.3758/s13415-013-0242-y

Molinari, G., Poellhuber, B., Heutte, J., Lavoué, E., Widmer, D. S., & Caron, P.-A. (2016). L'engagement et la persistance dans les dispositifs de formation en ligne : Regards croisés. *Distances et médiations des savoirs. Distance and Mediation of Knowledge*, *13*, Article 13. https://doi.org/10.4000/dms.1332

Moon, Y., & Nass, C. (1996). How « Real » are computer personalities? Psychological responses to personality types in human-computer interaction. *Communication Research*, *23*(6), 651-674. https://doi.org/10.1177/009365096023006002

Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, *22*(3), 149-158. https://doi.org/10.1111/j.1365-2729.2006.00170.x

Moreno, R., & Durán, R. (2004). Do Multiple Representations Need Explanations? The Role of Verbal Guidance and Individual Differences in Multimedia Mathematics Learning. *Journal of Educational Psychology*, *96*(3), 492-503. https://doi.org/10.1037/0022-0663.96.3.492

Moreno, R., & Mayer, R. (2007). Interactive Multimodal Learning Environments : Special Issue on Interactive Learning Environments: Contemporary Issues and Trends. *Educational Psychology Review*, *19*(3), 309-326. https://doi.org/10.1007/s10648-007-9047-2

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning : The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358-368. https://doi.org/10.1037/0022-0663.91.2.358

Moreno, R., Mayer, R., Spires, H., & Lester, J. (2001). The Case for Social Agency in Computer-Based Teaching : Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents? *Cognition and Instruction - COGNITION INSTRUCT*, *19*, 177-213. https://doi.org/10.1207/S1532690XCI1902_02

Motz, B., Quick, J., Schroeder, N., Zook, J., & Gunkel, M. (2019). The validity and utility of activity logs as a measure of student engagement. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 300-309. https://doi.org/10.1145/3303772.3303789

Muljana, P., & Luo, T. (2019). Factors Contributing to Student Retention in Online Learning and Recommended Strategies for Improvement : A Systematic Literature Review. *Journal of Information Technology Education: Research*, *18*. https://doi.org/10.28945/4182

Munassar, W., Yahaya, W., & Chong, T. (2010). When is the Redundancy Principle Effective in the Multimedia Learning of a Foreign Language? *Journal of Institutional Research South East Asia*, *8*, 16-36.

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In *Oxford handbook of positive psychology, 2nd ed* (p. 195-206). Oxford University Press.

Nass, C., Moon, Y., Fogg, B., Reeves, B., & Dryer, C. (1995). Can computer personalities be human personalities? *Int. J. Hum.-Comput. Stud.*, *43*, 223-239. https://doi.org/10.1145/223355.223538

Neumann, O. (1996). Chapter 10 Theories of attention. In O. Neumann & A. F. Sanders (Éds.), *Handbook of Perception and Action* (Vol. 3, p. 389-446). Academic Press. https://doi.org/10.1016/S1874-5822(96)80027-2

No authorship indicated. (2016). « Interruptions disrupt reading comprehension » : Correction to Foroughi et al. (2015). *Journal of Experimental Psychology: General*, *145*(7), 881-881. https://doi.org/10.1037/xge0000186

O'Brien, H., & Cairns, P. (Éds.). (2016). *Why Engagement Matters*. Springer International Publishing. https://doi.org/10.1007/978-3-319-27446-1

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938-955. https://doi.org/10.1002/asi.20801

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In *The Cambridge handbook of multimedia learning, 2nd ed* (p. 27-42). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.004

Paas, F., Tuovinen, J. E., Tabbers, H., & Gerven, P. W. M. V. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. In *Cognitive Load Theory*. Routledge.

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist - EDUC PSYCHOL*, *38*, 63-71. https://doi.org/10.1207/S15326985EP3801_8

Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning : An eyetracking study on the use of anthropomorphisms. *Computers & Education*, *86*, 30-42. https://doi.org/10.1016/j.compedu.2015.02.016

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In *Handbook of research on student engagement* (p. 259-282). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-2018-7_12

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, *31*(6), 459-470. https://doi.org/10.1016/S0883-0355(99)00015-4

Pintrich, P. R. (2000). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology*, *25*(1), 92-104. https://doi.org/10.1006/ceps.1999.1017

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change : The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, *63*(2), 167-199. https://doi.org/10.2307/1170472

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning : Effects of shape and color on affect and learning. *Learning and Instruction*, *29*, 128-140. https://doi.org/10.1016/j.learninstruc.2013.02.006

Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In *Emotions, technology, design, and learning* (p. 131-161). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-801856-9.00007-4

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25-42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Radovan, M. (2019). Should I Stay, or Should I Go? Revisiting Student Retention Models in Distance Education. *Turkish Online Journal of Distance Education*, *20*(3), Article 3. https://doi.org/10.17718/tojde.598211

Reeve, J., & Lee, W. (2014). Students' classroom engagement produces longitudinal changes in classroom motivation. *Journal of Educational Psychology*, *106*(2), 527-540. https://doi.org/10.1037/a0034934

Reeves, B., & Nass, C. I. (1996). *The media equation : How people treat computers, television, and new media like real people and places* (p. xiv, 305). Cambridge University Press.

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being : The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, *26*(4), 419-435. https://doi.org/10.1177/0146167200266002

Renninger, K. A., & Hidi, S. (2011). Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educational Psychologist*, *46*(3), 168-184. https://doi.org/10.1080/00461520.2011.587723

Renninger, K. A., & Su, S. (2012). Interest and its development. In *The Oxford handbook of human motivation* (p. 167-187). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399820.013.0011

Rhoades, B. L., Warren, H. K., Domitrovich, C. E., & Greenberg, M. T. (2011). Examining the link between preschool social–emotional competence and first grade academic achievement : The role of attention skills. *Early Childhood Research Quarterly*, *26*(2), 182-191. https://doi.org/10.1016/j.ecresq.2010.07.003

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction*, *21*(1), 58-67. https://doi.org/10.1016/j.learninstruc.2009.11.001

Rothes, A., Lemos, M. S., & Gonçalves, T. (2022). The Influence of Students' Self-Determination and Personal Achievement Goals in Learning and Engagement : A Mediation Model for Traditional and Nontraditional Students. *Education Sciences*, *12*(6), Article 6. https://doi.org/10.3390/educsci12060369

Rovai, A. P. (2002). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education*, *5*(4), 319-332. https://doi.org/10.1016/S1096-7516(02)00130-6

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Salar, H. C., Başarmak, U., & Sezgin, M. E. (2023). Educational Integration of the Metaverse Environment in the Context of Web 3.0 Technologies : A Critical Overview of Planning, Implementation, and Evaluation. In *Shaping the Future of Online Learning : Education in the Metaverse* (p. 154-173). IGI Global. https://doi.org/10.4018/978-1-6684-6513-4.ch009

Satar, H. M., & Akcan, S. (2018). *Pre-service EFL teachers' online participation, interaction, and social presence*. http://hdl.handle.net/10125/44586

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*(3-4), 299-323. https://doi.org/10.1207/s15326985ep2603&4_5

Schoenau-Fog, H. (2011). Hooked ! – Evaluating Engagement as Continuation Desire in Interactive Narratives. In M. Si, D. Thue, E. André, J. C. Lester, T. J. Tanenbaum, & V. Zammitto (Éds.), *Interactive Storytelling* (p. 219-230). Springer. https://doi.org/10.1007/978-3-642-25289-1_24

Schroeder, N., Adesope, O., & Barouch Gilbert, R. (2013a). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*, 1-39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N., Adesope, O., & Barouch Gilbert, R. (2013b). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*, 1-39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N. L., & Adesope, O. O. (2014). A Systematic Review of Pedagogical Agents' Persona, Motivation, and Cognitive Load Implications for Learners. *Journal of Research on Technology in Education*, *46*(3), 229-251. https://doi.org/10.1080/15391523.2014.888265

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013a). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*(1), 1-39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013b). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*(1), 1-39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013c). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*(1), 1-39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013d). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, *49*(1), 1-39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N. L., & Gotch, C. M. (2015). Persisting Issues in Pedagogical Agent Research. *Journal of Educational Computing Research*, *53*(2), 183-204. https://doi.org/10.1177/0735633115597625

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement.* (p. xvii, 368). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-7089-2

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection : Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, *69*(3), 493-525. https://doi.org/10.1177/0013164408323233

Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The Multiple Subfunctions of Attention : Differential Developmental Gateways to Literacy and Numeracy. *Child Development*, *83*(6), 2028-2041. https://doi.org/10.1111/j.1467-8624.2012.01809.x

Stevens, C., & Bavelier, D. (2012a). The role of selective attention on academic foundations : A cognitive neuroscience perspective. *Developmental Cognitive Neuroscience*, *2 Suppl 1*(Suppl 1), S30-48. https://doi.org/10.1016/j.dcn.2011.11.001

Stevens, C., & Bavelier, D. (2012b). The role of selective attention on academic foundations : A cognitive neuroscience perspective. *Developmental Cognitive Neuroscience*, *2 Suppl 1*(Suppl 1), S30-48. https://doi.org/10.1016/j.dcn.2011.11.001

Sweller, J. (1988). Cognitive load during problem solving : Effects on learning. *Cognitive science*, *12*(2), 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295-312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2011). CHAPTER TWO - Cognitive Load Theory. In J. P. Mestre & B. H. Ross (Éds.), *Psychology of Learning and Motivation* (Vol. 55, p. 37-76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, *68*(1), 1-16. https://doi.org/10.1007/s11423-019-09701-3

Sweller, J., Ayres, P., & Kalyuga, S. (2011a). Measuring Cognitive Load. In J. Sweller, P. Ayres, & S. Kalyuga (Éds.), *Cognitive Load Theory* (p. 71-85). Springer. https://doi.org/10.1007/978-1-4419-8126-4_6

Sweller, J., Ayres, P., & Kalyuga, S. (2011b). The Split-Attention Effect. In J. Sweller, P. Ayres, & S. Kalyuga (Éds.), *Cognitive Load Theory* (p. 111-128). Springer. https://doi.org/10.1007/978-1-4419-8126-4_9

Towns, S. G., Voerman, J. L., Callaway, C. B., & Lester, J. C. (1998). Coherent gestures, locomotion, and speech in life-like pedagogical agents. *Proceedings of the 3rd International Conference on Intelligent User Interfaces - IUI '98*, 13-20. https://doi.org/10.1145/268389.268393

Tractinsky, N., Katz, A., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, *13*, 127-145. https://doi.org/10.1016/S0953-5438(00)00031-X

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. L. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology*, *48*, 116-132. https://doi.org/10.1016/j.cedpsych.2016.10.001

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The Influences of Emotion on Learning and Memory. *Frontiers in Psychology*, *8*, 1454. https://doi.org/10.3389/fpsyg.2017.01454

Um, E. "Rachel", Plass, J. L., Hayward, E. O., & Homer, B. D. (2012a). Emotional design in multimedia learning. *Journal of Educational Psychology*, *104*(2), 485-498. https://doi.org/10.1037/a0026609

Um, E. "Rachel", Plass, J. L., Hayward, E. O., & Homer, B. D. (2012b). Emotional design in multimedia learning. *Journal of Educational Psychology*, *104*(2), 485-498. https://doi.org/10.1037/a0026609

Underwood, G., & Everatt, J. (1996). Automatic and controlled information processing : The role of attention in the processing of novelty. In *Handbook of perception and action, Vol. 3 : Attention* (p. 185-227). Academic Press.

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall : The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17*(2), 200-205. https://doi.org/10.3758/PBR.17.2.200

Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In *The Cambridge handbook of multimedia learning, 2nd ed* (p. 263-278). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.014

Veletsianos, G., & Russell, G. (2014). Pedagogical Agents. *Handbook of Research on Educational Communications and Technology*, 759-769. https://doi.org/10.1007/978-1-4614-3185-5_61

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning : Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology*, *110*(2), 250.

Wang, J., & Antonenko, P. D. (2017). Instructor presence in instructional video : Effects on visual attention, recall, and perceived learning. *Computers in Human Behavior*, *71*, 79-89. https://doi.org/10.1016/j.chb.2017.01.049

Wiebe, E., & Sharek, D. (2016). eLearning. In H. O'Brien & P. Cairns (Éds.), *Why Engagement Matters : Cross-Disciplinary Perspectives of User Engagement in Digital Media* (p. 53-79). Springer International Publishing. https://doi.org/10.1007/978-3-319-27446-1_3

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary educational psychology*, *25*(1), 68-81.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., Biocca, F., Sacau, A., Jäncke, L., Baumgartner, T., & Jäncke, P. (2007). A Process Model of the Formation of Spatial Presence Experiences. *Media Psychology*, *9*(3), 493-525. https://doi.org/10.1080/15213260701283079

Wolfson, S., & Case, G. (2000). The effects of sound and colour on responses to a computer game. *Interacting with Computers*, *13*(2), 183-192. https://doi.org/10.1016/S0953-5438(00)00037-0

Wolters, C. A., & Taylor, D. J. (2012). A Self-regulated Learning Perspective on Student Engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Éds.), *Handbook of Research* *on Student Engagement* (p. 635-651). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_30

Woo, H. l. (2009). Designing multimedia learning environments using animated pedagogical agents : Factors and issues. *Journal of Computer Assisted Learning*, *25*(3), 203-218. https://doi.org/10.1111/j.1365-2729.2008.00299.x

Xavier, M., & Meneses, J. (2020). *A literature review on the definitions of dropout in Online Higher Education*. https://doi.org/10.38069/edenconf-2020-ac0004

Yang, L., Kandasamy, K., & Hasher, L. (2022). Inhibition and Creativity in Aging : Does Distractibility Enhance Creativity? *Annual Review of Developmental Psychology*, *4*(Volume 4, 2022), 353-375. https://doi.org/10.1146/annurev-devpsych-121020-030705

Zaharias, P., & Poulymenakou, A. (2009). Developing a Usability Evaluation Method for e-Learning Applications : Beyond Functional Usability. *Int. J. Hum. Comput. Interaction*, *25*, 75-98. https://doi.org/10.1080/10447310802546716

Zaharias, P., & Poylymenakou, A. (2009). Developing a Usability Evaluation Method for e-Learning Applications : Beyond Functional Usability. *International Journal of Human–Computer Interaction*, *25*(1), 75-98. https://doi.org/10.1080/10447310802546716

Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of self-regulation of learning and performance* (p. xiv, 484). Routledge/Taylor & Francis Group.