



HAL
open science

L'apprentissage en environnement virtuel : rôle du guidage et du feedback sur la compréhension des écosystèmes forestiers

Laurie Porte

► **To cite this version:**

Laurie Porte. L'apprentissage en environnement virtuel : rôle du guidage et du feedback sur la compréhension des écosystèmes forestiers. Psychologie. Université Bourgogne Franche-Comté, 2021. Français. NNT : 2021UBFCH043 . tel-03597992

HAL Id: tel-03597992

<https://theses.hal.science/tel-03597992v1>

Submitted on 4 Mar 2022

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

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

THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE
Préparée au Laboratoire d'Etude de L'Apprentissage et du Développement – CNRS UMR5022

Ecole doctorale n°554
Environnements-Santé

Doctorat de psychologie

Par
Mme PORTE Laurie

**L'apprentissage en environnement virtuel : Rôle du guidage et
du feedback sur la compréhension des écosystèmes forestiers.**

Thèse présentée et soutenue à Dijon, le 03 décembre 2021

Composition du jury:

Pr. Jean-Pierre THIBAUT	LEAD, Université Bourgogne Franche-Comté, Dijon	Président
Pr. Mireille BETRANCOURT	TECFA, Université de Genève	rapporteuse
Pr. André TRICOT	EPSYLON, Université Paul Valéry Montpellier	rapporteur
Pr. Jean-Marie BURKHARDT	LaPEA, Université Gustave Eiffel, IFSTTAR & Université de Paris,	Examineur
Pr. Emérite Michel FAYOL	LAPSCO, Université Clermont Auvergne	Examineur
Pr. Jean-Michel BOUCHEIX	LEAD, Université Bourgogne Franche-Comté, Dijon	Directeur de thèse
M. Serge ABADIE	Directeur de l'EPL, Porteur du projet PIA Projet Silva Numerica	Invité
Mme Véronique WOZNIAK	Inspectrice du domaine pédagogique, Ministère de l'agriculture et de l'alimentation	Invitée

Opération soutenue par l'État dans le cadre du volet e-FRAN du Programme d'investissements d'avenir, opéré par la Caisse des Dépôts

Dédicaces . . .

À ma mie, Maryline, ma maman,

*À ma Mère,
À tonton G.*

Et . . . à mon Kimber . . .



Tout d'abord, je tiens à remercier mon directeur de thèse, Jean-Michel Boucheix, pour la confiance qu'il m'a accordée depuis mes années de master jusqu'à la fin de cette thèse. Merci pour sa patience, pour les nombreux conseils qui m'ont été donnés, pour toute l'aide qu'il m'a apportée et pour la richesse de nos échanges qui m'ont permis d'apprendre énormément.

Je remercie également les membres du jury, qui m'ont fait l'honneur d'accepter cette tâche. Merci pour leur bienveillance ainsi que pour les conseils avisés qu'ils ont pu me donner lors de rencontres ponctuelles.

Je remercie évidemment toute l'équipe SilvaNumerica, et tous les partenaires (visibles à la page 5) qui ont su faire de ce projet une aventure enrichissante. Merci à Michel Guyot, chef de projet. Merci à Serge Abadie, coordinateur, toujours rassurant, intéressé par nos travaux et motivant.

Je remercie également :

Les directeurs d'établissements et équipes pédagogiques qui nous ont reçu à l'occasion de la réalisation de nos expérimentations : le collège Jean Jaurès de Saint Vit, le collège Rameau de Fontaine d'Ouche, et le lycée Granvelle de Dannemarie sur Crète. Merci à Pascale Muin, Fabienne Martin, Jérôme Cuinet, Isabelle Brigand, Isabelle Petitjean, Claudie Gibey et Christelle Pavon qui se sont énormément investis dans ce projet.

Le lycée et CFPPA Velet, de m'avoir accueillie pour une expédition sur le terrain pendant laquelle j'ai beaucoup appris ; j'aurai aimé avoir l'occasion d'en faire d'autres. Merci à Jérémie, Daniel, Frédéric et toute l'équipe qui nous ont chaleureusement accueillis Thibault et moi.

Le CFPPA de Chateaufarine qui nous a toujours réservé un accueil chaleureux.

La Délégation Régionale du Numérique pour l'Education, plus particulièrement à Sarah Cornu et Nathalie Bécoulet qui ont veillé à ce que nous puissions disposer du matériel nécessaire à nos expérimentations et qui ont toujours été à notre écoute. Merci également pour votre sympathie !

Stéphane Argon, qui a développé nos deux premières expériences et qui est toujours disponible, dans la joie et la bonne humeur pour nous prêter main forte. Les stomates et tous ces animaux nous auront bien fait rire malgré un rythme de travail fort soutenu.

Je remercie également le PNR de l'Université de Dijon pour le prêt d'un casque de RV.

Jean-Luc Martinez, Louis Rapet et Rémy Malin de l'Institut Image de Chalon sur Saône, pour le développement de notre matériel expérimental en RVI.

Irène Dorget, Sandrine Pinson, Corinne Martin, Maëva Peter et Christelle Caillot, sans qui rien ne serait vraiment possible. Un grand merci pour tout mesdames ☺

Laurie Dunn, doctorante à l'INRA, qui m'a prêté les documents dont j'avais besoin et a su répondre à mes interrogations en matière de pédologie avec beaucoup de patience et de gentillesse.

Thibault Chiron, pour sa bienveillance, sa gentillesse, son humour et ses conseils. THE collègue que tout le monde rêverait d'avoir ^^

Benoit Haigre avec qui ce fut un plaisir de travailler et qui a été d'une grande aide lors de notre dernière expérimentation.

Tous les étudiants avec qui j'ai eu la chance de travailler sur ce projet : Clémence Rougeot, Camille Clerc, Yohan Grand, Romain Flaujaguet, Lénaïc Ferrero, Dimitri voisin...

Noa Jaonah, le nain, mon mini testeur de RV ☺

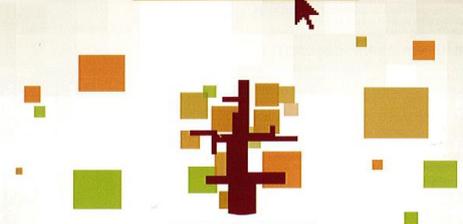
Tous mes collègues doctorants et chercheurs du LEAD, de Marey comme du pôle AAFE : Bénédite Poulin-Charronnat qui est toujours à notre écoute et d'un grand soutien ; Alessandro Carlini qui me fait toujours penser à « Three little birds » ♪... 'cause every little thing is gonna be alright ♪^^ ; Betty Laroche qui un jour a su me faire sortir de ma batcave, Aurélia Bugajska dont l'enthousiasme communicatif a contribué à ce que je postule pour cette thèse... Avec une pensée spéciale pour Laure-Hélène, Thomas, Joris, Guillaume et Camille qui ont supporté ma déco de bureau farfelue, mes logorrhées et mon rire des plus discret. Vous avez su enrichir ces années de thèse, les ponctuant d'excellents souvenirs !

Ma Mom', la meilleure, qui a été d'un soutien sans faille, qui a pris le temps de relire une grande partie de ma thèse, et qui m'a énormément encouragée.

Ma princesse à paillettes, Fabien Jaonah, qui a supporté mon stress, mes doutes perpétuels... et qui a été d'un grand soutien.

Je remercie également les irréductibles: Jérémy Cannet, Fanouille Coeur, Grat' Pommier, Justine Sabot, Nech, Matthieu, Jo' et Rocket.

APPRENDRE LA FORÊT PAR SIMULATION



SILVA NUMERICA

Silva Numerica : conception d'une plateforme numérique simulant un environnement forestier en réalité virtuelle.
 Cette plateforme pourra accueillir différents modules de formation scénarisés. Silva Numerica est un outil d'apprentissage destiné aux apprenants de l'enseignement général et professionnel en vue de développer leurs compétences métier.







Opération soutenue par l'État dans le cadre du volet e-FRAN du Programme d'investissements d'avenir, opéré par la Caisse des Dépôts

SILVA NUMERICA : AVEC QUI ?

Une expérimentation pédagogique conduite dans trois établissements de formation :

- > Collège Jean Jaures de Saint Vit, collèges de la 6^e à la 3^e
- > EPLEFPA (Établissement Public Local d'Enseignement et de Formation Professionnelle Agricole) Velet, Lycée & CFPPA de la Nature et de la Forêt (lycéens en BAC Pro Forêt et Gestion des Milieux Naturels et de la Faune, stagiaires en Brevet Professionnel Responsable de Chantiers Forestiers)
- > EPLEFPA de Besançon, Lycée Grammelet et Centre de Formation de Châteaufortine (lycéens en BAC Sciences et technologies de l'Agronomie et du Vivant, BAC Scientifique et apprenants en BAC Pro Forêt, Brevet Professionnel Responsable de Chantiers Forestiers et BTS Gestion Forestière)
- > Trois autres établissements seront aussi expérimentateurs : l'Établissement Forestier de Moymac, les EPLEFPA de Maymac, et de Bazas et le Centre de Formation Professionnelle Forestière de Mont-sur-Lausanne

SILVA NUMERICA : COMMENT ?

Il s'agit de mettre en place les conditions d'élaboration et d'évaluation d'un système d'enseignement et de formation en nous appuyant sur une plateforme de réalité virtuelle d'un environnement forestier.

SILVA NUMERICA : POUR QUOI ?

- > Enseigner l'éducation à l'environnement, la découverte des écosystèmes forestiers et des évolutions naturelles
- > Apprendre à gérer les peuplements forestiers
- > Apprendre à gérer les chantiers forestiers

UN PROJET PORTÉ PAR L'EPLFPA DE BESANÇON

2, rue des Charnats
 25410 Dammarie-sur-Creuse
 03 81 58 81 41

Chef de projet : Michel CUVOT
 michel.cuvot@educagri.fr
 06 84 60 71 30
<http://www.silvanumerica.fr>

Projet mené par neuf partenaires de l'enseignement, de la recherche, du développement durable et de l'évaluation : EPLFPA de Velet, collège Jean Jaures de Saint-Vit, Agrisup Dijon, Eduter Recherche, Université de Bourgogne LEAD, Université de Bourgogne REDU, AMVALOR, Shina Research, CFPPA Montmorot.

QUATRE ANS POUR CONCEVOIR UN ENVIRONNEMENT DE RÉALITÉ VIRTUELLE ET ÉTUDIER SA PLUS-VALUE PÉDAGOGIQUE



INSTITUT, LABORATOIRES DE RECHERCHE      	PARTENAIRES INSTITUTIONNELS ET FINANCIERS      	ÉTABLISSEMENTS D'ENSEIGNEMENT      	PROFESSEURS DE LA FILIÈRE FORÊT-BOIS BOURGOGNE-FRANCHE-COMTE        
--	--	--	---

Titre : L'apprentissage en environnement virtuel : Rôle du guidage et du feedback sur la compréhension des écosystèmes forestiers.

Mots clés : apprentissage, animations, écosystème, réalité virtuelle, cognition, guidage

Résumé : Dans le cadre du projet E-Fran Silva Numerica, nous avons étudié les processus cognitifs liés à l'apprentissage dans le but d'optimiser un simulateur de forêt en réalité virtuelle. Ces travaux de thèse ont donné lieu à quatre expérimentations portant sur un test des principes d'apprentissage multimédia (Mayer, 2009, 2014, 2021) en réalité virtuelle, ainsi que sur les mécanismes cognitifs sous-jacents aux acquisitions dans le domaine des écosystèmes forestiers. Notre première étude a porté sur l'apprentissage de la décomposition de la matière organique auprès des plus de 100 élèves de 5ème. Nous avons créé un environnement 3D respectant le principe de cohérence sémantique afin de comparer deux groupes d'élèves qui étudiaient soit la leçon selon une présentation sémantiquement cohérente fixe, soit la leçon selon une présentation libre, complète. Cette étude a permis de confirmer qu'un apprentissage sémantiquement cohérent fixe était préférable dans le cadre d'apprentissages complexes en RV auprès de jeunes élèves. Dans une deuxième étude sur le même thème et auprès d'une population de 230 élèves de 5ème, nous avons testé le principe de contiguïté temporelle et les conditions d'intégration de l'information verbale (commentaire audio) et de l'information visuelle. Le matériel proposé était identique à celui de la première expérience mais était présenté sous forme d'une vidéo non interactive dans laquelle le son était présenté soit 6 secondes, soit 2 secondes avant ou après l'information visuelle.

Un groupe étudiait cette vidéo avec une synchronisation (0 secondes) entre l'information verbale et l'information picturale. Cela nous a permis d'évaluer le degré de perturbation des apprentissages multimodaux dans un environnement complexe en 3D lorsque que les informations verbales et imagées ne sont pas présentées simultanément. Les résultats ont indiqué qu'un décalage temporel de 2 secondes seulement entre l'information verbale et l'information picturale perturbait significativement la compréhension et l'apprentissage. Cela nous a permis de mieux appréhender la conception de notre dernière expérimentation portant sur l'importance du guidage visuel par signalisation, et du feedback dans un environnement forestier en réalité virtuelle immersive, dans lequel il est possible d'extraire de l'information à 360° parmi une multitude d'éléments. La tâche expérimentale proposée à plus de 100 élèves de première et terminale d'un lycée agricole, impliquait un diagnostic de l'écosystème de zones forestières en RV en vue de l'implantation d'un lieu d'accueil du public. Les résultats ont montré un effet positif significatif de la signalisation visuelle et du feedback sur la performance de diagnostic et sur la récupération en mémoire de travail des informations pertinentes pour élaborer le diagnostic.

Title : The role of guidance and feedback on understanding forest ecosystems.

Keywords : learning, ecosystem, virtual reality, cognition, cueing, animation

Abstract : As part of the E-Fran Silva Numerica project, we studied cognitive processes related to learning in order to optimize a virtual reality forest simulator. This thesis work resulted in four experiments involving a test of multimedia learning principles (Mayer, 2009, 2014, 2021) in virtual reality, as well as the cognitive mechanisms underlying acquisitions in the forest ecosystem domain. Our first study focused on learning about the decomposition of organic matter with over 100 5th grade students. We created a 3D environment respecting the principle of semantic coherence in order to compare two groups of students who studied either the lesson according to a fixed semantically coherent presentation or the lesson according to a free, complete presentation. This study confirmed that fixed semantically consistent learning was preferable in complex VR learning with young students. In a second study on the same theme and with a population of 230 5th grade students, we tested the principle of temporal contiguity and the conditions for integrating verbal information (audio commentary) and visual information. The material proposed was identical to that of the first experiment but was presented in the form of a non-interactive video in which the audio was presented either 6 seconds or 2 seconds before or after the visual information.

One group studied this video with synchronization (0 seconds) between the verbal and pictorial information. This allowed us to assess the degree of disruption of multimodal learning in a complex 3D environment when verbal and pictorial information are not presented simultaneously. The results indicated that a time lag of only 2 seconds between verbal and pictorial information significantly harmed comprehension and learning. This result provided insight into the design of our latest experiment on the importance of visual guidance, signalling, and feedback in an immersive virtual reality forest environment in which 360° information can be extracted from a multitude of elements. The experimental task, proposed to more than 100 last year students of an agricultural high school, involved a diagnosis of the ecosystem of forest areas in VR with a goal of building a public reception area. The results showed a significant positive effect of visual signalling and feedback on the diagnostic performance and on the retrieval in working memory of the relevant information needed to elaborate correct diagnosis.

Table des matières

Introduction générale	1
Le projet SilvaNumerica	1
CHAPITRE 1 - LES APPRENTISSAGES MULTIMEDIAS.....	3
1.1 Le principe multimédia et les modèles cognitifs du traitement de l'information multimodale.....	4
1.2 La théorie de l'apprentissage multimédia.....	5
1.3 Un point sur la mémoire.....	6
1.4 La cohérence.....	8
1.5 La signalisation ou « cueing »	18
1.6 La contiguïté spatiale et temporelle.....	27
1.6.1 La contiguïté spatiale.....	27
1.6.2 La contiguïté temporelle.....	32
1.7 Le feedback.....	36
1.8 Théorie de la charge cognitive.....	38
CHAPITRE 2 - LES APPRENTISSAGES A PARTIR DE VISUALISATIONS DYNAMIQUES.....	43
2.1 Définition et caractéristiques des animations.....	44
2.2 Le rôle des habiletés de raisonnement spatial dynamique.....	45
2.3 Le traitement de l'animation : le Modèle APM, Lowe & Boucheix (2008).....	49
2.4 Les méta-analyses.....	51
2.4.1 Comparaison de présentations statiques et animées.....	51
2.4.2 Animé vs statique : la question de la fiabilité des comparaisons réalisées.....	52
2.4.3 Animations et habiletés visuo-spatiales	55
2.5 Caractéristiques des animations dans l'apprentissage.....	58
CHAPITRE 3 - APPRENDRE AVEC LA REALITE VIRTUELLE IMMERSIVE ET NON IMMERSIVE.....	60
3.1 Définition et caractéristiques de la réalité virtuelle.....	61
3.2 Le CAMIL, Makransky & Petersen (2021).....	62
3.3 Un gain de connaissances en RV ?.....	64
3.4 Comparaison de la RV à des supports traditionnels.....	66
3.5 La prise en compte du genre en RV.....	70
3.6 RVI et interactivité.....	72
3.7 L'effet de modalité en RV.....	73
3.7.1 Epurer l'environnement RV pour une meilleure allocation des ressources attentionnelles ?	75

3.8	L'effet du guidage par signalisation en réalité virtuelle.....	76
PARTIE EMPIRIQUE		79
Introduction.....		79
Chapitre 4 - expérience 1, Le principe de cohérence sémantique en réalité virtuelle non immersive.....		92
4.1	Introduction.....	93
4.2	Méthode.....	100
4.2.1	Participants.....	100
4.2.2	Matériel expérimental.....	101
4.2.3	Procédure.....	107
4.2.4	L'analyse des données.....	108
4.3	Résultats.....	109
Chapitre 5 - La contiguïté temporelle en réalité virtuelle : Effets des latences temporelles contrastées narration-animation et différences individuelles dans l'apprentissage d'un écosystème biologique au collège.....		115
5.1	Introduction	116
5.1.1	Quand et pourquoi l'effet de contiguïté temporelle fonctionne-t-il, et quels sont les processus cognitifs sous-jacents ?	117
5.1.2	Limites et résultats mitigés dans l'effet de contiguïté temporelle : quand et pourquoi ça ne fonctionne pas ?.....	118
5.1.3	Questions de recherche et hypothèses	120
5.2	Méthode.....	126
5.2.1	Participants.....	126
5.2.2	Matériel expérimental.....	127
5.2.3	Procédure.....	133
5.2.4	L'analyse des données	134
5.3	Résultats.....	135
5.3.1	Contrôle d'homogénéité des groupes.....	135
5.3.2	Connaissances factuelles et compréhension des scores pré et post-test.....	136
5.3.3	Mémoire de travail, capacité spatiale, contiguïté temporelle et performances d'apprentissage au post-test.....	139
5.3.4	Test de mémoire épisodique de correspondance texte-image	141
5.4	Discussion et conclusion	145
CHAPITRE 6 - APPRENTISSAGES EN REALITE VIRTUELLE IMMERSIVE : ROLE DE DEUX TYPES DE SIGNALISATION ET DU FEEDBACK.		152
6.1	Introduction.....	153
6.1.1	Apprendre avec la réalité virtuelle immersive : des résultats très contrastés...	153

6.1.2	Les mécanismes cognitifs liés aux apprentissages en réalité virtuelle immersive.	155
6.1.3	La signalisation ou « cueing » en réalité virtuelle immersive.....	156
6.1.4	Le rôle du feedback en réalité virtuelle immersive.	158
6.2	Méthode.....	163
6.2.1	Participants.....	163
6.2.2	Matériel expérimental.....	164
6.2.3	Procédure.....	173
6.3	Résultats.....	176
6.3.1	Contrôle d'homogénéité du groupe.....	176
6.4	Discussion et conclusion.	186
	CHAPITRE 7 : DISCUSSION GENERALE	190
7.1	Objectif de nos travaux.....	191
7.2	Les apprentissages en réalité virtuelle : les apports de nos travaux.	192
7.2.1	Le principe de cohérence sémantique en réalité virtuelle non immersive.	193
7.2.2	La contiguïté temporelle en réalité virtuelle.	194
7.2.3	Rôle de la signalisation et du feedback en RVI.....	195
7.2.4	Conclusion.....	197
7.3	Recommandations.	197
8	REFERENCES	200
9	ANNEXES.....	221

Introduction générale

Le projet SilvaNumerica

Cette thèse s'inscrit dans le projet E-FRAN **SILVA NUMERICA** qui vise à répondre aux besoins des établissements scolaires allant du collège à la formation professionnelle dans le cadre d'apprentissages complexes relatifs à l'environnement forestier. Ce projet d'une durée de quatre ans s'est déroulé en trois étapes : la conception, la modélisation et l'évaluation d'un environnement virtuel éducatif et évolutif répondant à tous les niveaux d'enseignement que nous venons d'évoquer. Le but était de reproduire une forêt virtuelle vivante conforme aux espaces forestiers que l'on peut rencontrer sur le territoire français, dans le but de la faire évoluer selon les besoins pédagogiques des différents enseignants et formateurs prenant part à ce projet. La modélisation de cette forêt permettra alors aux élèves de se représenter l'environnement forestier, sa diversité, sa complexité et d'évaluer l'impact de l'intervention de l'homme sur l'évolution, et le développement des écosystèmes qui la composent. Cet outil viendrait donc compléter les enseignements grâce à la simulation, rendant visible l'évolution de la forêt suite aux actions menées à plus ou moins long terme. En effet, l'environnement forestier est un système dynamique à long délai de réponse, il met du temps à se développer et à évoluer ; pour le comprendre il est nécessaire de pouvoir s'en créer une représentation dans le temps à l'échelle de plusieurs décennies. Si les professionnels sont capables de créer une représentation mentale des répercussions de l'action humaine sur l'environnement forestier sur plusieurs décennies, il n'en est pas de même pour les apprenants. Seule l'expérience permet une projection dans le temps, qui peut parfois se révéler insuffisamment précise pour une gestion optimale. Cette expérience acquise au fil des années par les professionnels forestiers donne lieu à l'application de connaissances explicites, académiques mais aussi à des connaissances plus instinctives c'est-à-dire implicites, automatiques liées entre autres à l'observation assidue du dit environnement dans de multiples situations pendant plusieurs années. De ce fait, leur expertise était indispensable à la création de cet environnement qu'il a fallu façonner et adapter de manière à ce qu'il soit le plus fidèle possible aux problématiques de terrain que les futurs professionnels pourraient rencontrer au cours de leur carrière.

Concernant la conception de l'outil, elle a été réalisée par deux équipes de développeurs : l'Institut Image AMVALOR de Chalon sur Saône et Shine research de Besançon. Ils ont travaillé en étroite collaboration avec les enseignants, formateurs et chercheurs de Bourgogne Franche-Comté. Le but était de répondre aux besoins d'enseignants de collège, de lycée

professionnel et de centres de formations professionnels, ce qui représente un large éventail de possibilités, de choix mais également de contraintes liées aux besoins de chacun. Quant aux chercheurs, deux équipes étaient également mobilisées : une équipe en sciences de l'éducation issue du laboratoire EDUTER d'Agrosup Dijon était chargée d'évaluer l'utilisabilité, et l'appropriation de l'outil pour les élèves et pour les enseignants/formateurs. Notre équipe, du Laboratoire d'Etude des Apprentissages et du Développement à l'Université de Dijon était chargée d'évaluer les processus cognitifs liés à l'utilisation de cet outil de manière à optimiser les apprentissages. Notre but était de nous assurer que l'information présente dans le simulateur soit organisée de manière à ce que l'apprentissage soit le plus efficace possible. En effet, si les outils multimédias semblent être une plus-value pour l'apprentissage, ils mettent en jeu une quantité d'informations considérable qui implique une charge perceptive, attentionnelle et cognitive élevée qui pourrait se révéler préjudiciable pour l'acquisition de connaissances bien circonscrites et intégrées. C'est sur ce point que nous allons nous centrer afin de préciser notre démarche et nos objectifs. Pour cela, au cours de cette introduction générale nous allons développer trois chapitres : le premier portera sur les apprentissages multimédias. Nous ferons le point sur les connaissances acquises et nous présenterons les modèles qui sont à l'origine des principes ergonomiques de conception des environnements multimédias qui en résultent. Le second chapitre sera dédié aux apprentissages à l'aide de visualisations dynamiques et d'animations qui sont très fréquents dans les environnements virtuels, en particulier immersifs. Pour terminer cette partie théorique, nous nous centrerons sur l'apprentissage en réalité virtuelle. Nous aborderons les différentes études encore peu nombreuses et toutes réalisées récemment. Nous nous intéresserons en particulier aux conditions de présentation de l'information virtuelle, riche et multiple, pour l'optimisation des apprentissages : « qu'est-ce qui fonctionne, qu'est-ce qui ne fonctionne pas ? ».

CHAPITRE 1 - LES APPRENTISSAGES MULTIMEDIAS

1.1 Le principe multimédia et les modèles cognitifs du traitement de l'information multimodale.

Les technologies prennent une place grandissante dans notre quotidien et sont de plus en plus utilisées dans les milieux scolaire et professionnel. Que notre but soit le divertissement ou l'apprentissage, nous utilisons quotidiennement ordinateurs, tablettes, smartphones ou encore consoles de jeux qui nous demandent de réaliser un traitement complexe d'informations multiples. C'est ce à quoi renvoie la notion de multimédia : « *on se réfère à la capacité des ordinateurs à fournir des présentations audio et visuelles à la fois réalistes et construites, dans une combinaison souhaitée de texte, de son, d'images fixes, de vidéos ou d'animations.* » Clark & Feldon (2014) *The Cambridge handbook of multimedia Learning*, Richard E. Mayer, p152. Concernant l'apprentissage multimédia, Mayer & Moreno (2003) le définissent comme le fait d'apprendre ou de comprendre, à partir de mots et d'images et dont les enseignements tendent à présenter les mots et les images de manière à favoriser l'apprentissage. Plus spécifiquement, il s'agit de la construction d'une représentation mentale à partir d'informations verbales (orale ou écrite) et d'informations visuelles sous forme d'images, d'animations ou de vidéos, Mayer (2014) *The Cambridge handbook of multimedia Learning*, Richard E. Mayer, p3. Si les supports multimédias sont de plus en plus utilisés dans l'enseignement, on peut se demander dans quelles mesures ils sont ou non bénéfiques pour les apprentissages. Si aujourd'hui de nombreuses études ont été menées à ce sujet, il n'en reste pas moins que beaucoup de supports multimédia éducatifs sont élaborés « *presque entièrement sur la base de l'opinion d'experts plutôt que sur les résultats empiriques de la recherche* » Najjar (1998). Il est donc possible que certains outils multimédias éducatifs ne soient pas conformes aux différents modèles et principes mis en exergue par les recherches réalisées jusqu'alors. Cela expliquerait en partie le fait que dans certains cas, les supports traditionnels montrent une efficacité semblable voire meilleure que les supports multimédias. Il semble donc important d'indiquer que si l'expertise est indispensable à la conception d'outils multimédias éducatifs, elle se doit d'être étroitement liée à la recherche de manière à ce que les apprentissages qui en résultent soient les plus efficaces possible. En effet « *Le multimédia offre des avantages considérables pour l'éducation quand il est utilisé pour fournir un accès rentable à un enseignement de haute qualité, fondé sur des preuves.* » Clark & Feldon (2014) *The Cambridge handbook of multimedia Learning*, Richard E. Mayer, p151. Les travaux réalisés dans cette thèse visent à tester les différents principes et modèles élaborés par la recherche, en réalité virtuelle. Le but de ce premier chapitre est donc

d'exposer ces principes et modèles de manière à pouvoir expliquer précisément comment nos expérimentations ont été pensées et conçues. Pour commencer, nous allons détailler le cœur de notre champ de recherche : la théorie de l'apprentissage multimédia.

1.2 La théorie de l'apprentissage multimédia.

La théorie de l'apprentissage multimédia part de faits expérimentaux montrant que l'on apprend mieux à partir d'informations multiples plutôt qu'à partir d'une seule source d'information, et cela sous certaines conditions limitantes (Mayer, 2009). On apprendrait donc mieux un texte lorsqu'il est illustré que lorsqu'il est présenté seul et inversement. Cette théorie propose trois types de processus. Le premier renvoie au fait que l'être humain posséderait plusieurs canaux sensoriels pour traiter l'information visuelle et auditive (Paivio, 1986, 2006). Le deuxième est lié à la capacité de traitement de l'information limitée dans chaque canal perceptif (Baddeley, 1992). Le troisième propose que l'apprenant dirigerait son attention sur les informations pertinentes (sélection) pour ensuite les organiser (organisation) de manière à créer une représentation mentale cohérente (intégration) tout en y intégrant des connaissances déjà acquises (Mayer, 2009 ; Wittrock, 1989). En somme, l'apprenant perçoit d'abord les informations visuelles et verbales puis sélectionne les éléments pertinents qui s'y rapportent. Il organise ensuite mentalement les informations sélectionnées en modèles logiques verbaux et en représentations visuelles. Enfin, ces deux modèles sont intégrés l'un avec l'autre, en relation avec les connaissances antérieures de l'apprenant.

Cela nous renvoie alors au modèle de Mayer (Figure 1.1), décrivant les étapes de l'apprentissage selon les paramètres suivants :

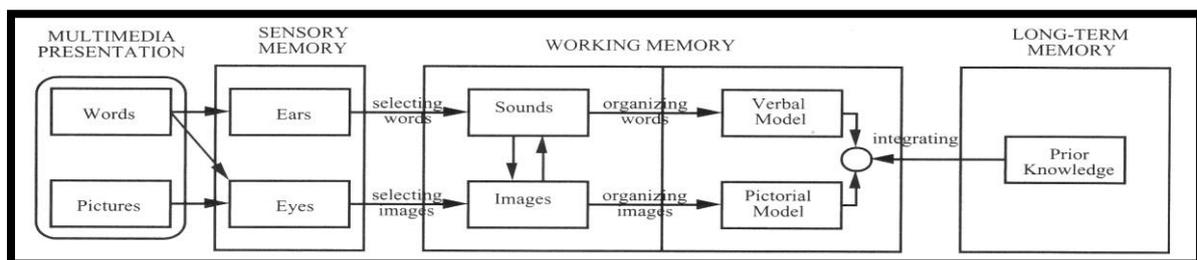


Figure 1.1. Schématisation du modèle de Mayer (2009).

Les mots empruntent ici l'entrée sensorielle auditive ou visuelle selon qu'il sont présentés oralement ou visuellement. Logiquement, les images empruntent l'entrée sensorielle visuelle et

les mots, selon qu'ils sont entendus ou lus empruntent d'abord l'entrée sensorielle visuelle et/ou auditive. Ces informations sont alors codées puis sélectionnées pour être manipulées en mémoire de travail puis transférées et donc intégrées en mémoire à long terme. Selon le modèle de Baddeley (2000), cela voudrait dire que les informations verbales associées aux informations visuelles, seraient distribuées par l'administrateur central sur les canaux auditif et visuel ce qui se traduirait donc par un partage de la charge attentionnelle. Selon le modèle plus précis de Schnotz (1999,2014), le système auditif va traiter les informations verbales orales afin de créer un modèle mental verbal alors que le système visuel va traiter les informations picturales afin de créer un modèle mental imagé. Ces informations sont intégrées dans le but de créer un modèle mental cohérent. Cela dit, si l'un ou les deux canaux sensoriels sont trop sollicités, alors cela va créer une charge cognitive trop importante entraînant un échec partiel ou total de l'intégration de l'information et donc de l'apprentissage. Pour favoriser un apprentissage optimal, il est donc impératif de présenter l'information de manière à ce qu'elle soit suffisamment complète pour entraîner l'acquisition de connaissances sans pour autant surcharger les canaux sensoriels ni la mémoire de travail. Sur la base de cette théorie, bon nombre d'études ont permis de faire émerger 22 principes (voir Mayer (2014), *The Cambridge handbook of multimedia Learning*, Richard E. Mayer, p117-119) ayant permis de détailler les caractéristiques de présentation et de traitement de l'information, selon le type de matériel éducatif utilisé et ce, en fonction du type de population visée. Dans les paragraphes suivants, nous présenterons cinq de ces principes multimédias. Ils font référence à ceux que nous avons testés en réalité virtuelle au cours des travaux réalisés au sein de cette thèse. Afin de mieux les comprendre, nous allons commencer par rappeler brièvement les caractéristiques de la mémoire.

1.3 Un point sur la mémoire.

Selon Clarys (2001), la mémoire désigne « *l'encodage, la rétention et la restitution d'une information.* ». On distingue deux composantes principales que sont la mémoire à court terme et la mémoire à long terme. La mémoire de travail est une mémoire à court terme que Baddeley (1992) décrit comme étant « *un système de capacité limitée, qui permet le stockage temporaire et la manipulation des informations qui sont nécessaires pour la réalisation de tâches complexes, telles que la compréhension, l'apprentissage et le raisonnement* ». Cette mémoire est divisée en trois composantes au sein du modèle de Baddeley & Hitch (1974) :

- L'administrateur central est décrit comme la composante principale dédiée au contrôle attentionnel. Il permettrait de coordonner le calepin visuo-spatial et la boucle phonologique alors considérés comme deux sous-systèmes.
- Le calepin visuo-spatial permettrait de maintenir les informations visuelles et spatiales en mémoire. Il permet également la création et la manipulation de représentations mentales.
- La boucle phonologique permettrait de stocker et de traiter les informations relatives au langage qu'elles soient verbales ou auditives.

En 2000, Baddeley modifie son modèle et insère la notion de Buffer épisodique (Figure 1.2). Ce dernier sous système à capacité limitée permettrait de regrouper les informations provenant de sources différentes. Il pourrait par exemple combiner les informations visuelles aux informations sonores et/ou verbale. On appelle ce phénomène le "binding". Celui-ci serait aussi en interaction avec la mémoire à long terme (MLT) de laquelle il extrairait des informations permettant de créer une représentation unitaire cohérente.

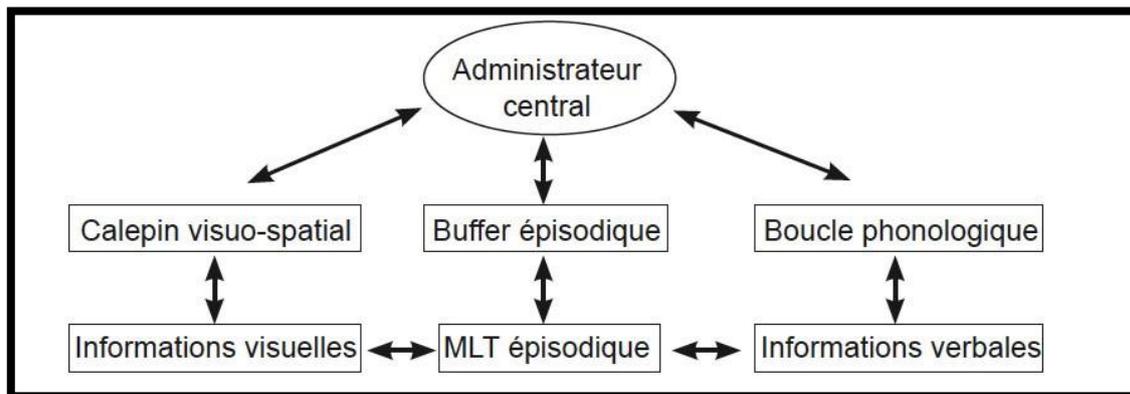


Figure 1.2. Modèle de Baddeley (2000).

Comme nous l'avons mentionné, la mémoire de travail a une capacité limitée. Cette limite a été déterminée par Miller (1956) indiquant que notre empan mnésique en mémoire à court terme serait de 7 plus ou moins 2 éléments.

Concernant la mémoire à long terme, il s'agit d'une « vaste réserve de connaissances et d'un registre d'évènements antérieurs » Cowan (2008). Il s'agit d'une mémoire « à capacité illimitée, possédant entre autres un nombre important de schémas automatisés. » Tricot (1998). Ces deux types de mémoire fonctionnent de concert : l'information est traitée en mémoire de travail avant d'être transférée et stockée en mémoire à long terme, Sweller, Ayres & Kalyuga

(2011). Cela dit, la qualité du traitement de l'information en mémoire de travail dépend du type de tâche demandée : si la tâche requiert des ressources mnésiques et attentionnelles trop importantes alors elle sera moins bien réalisée. De ce fait, on dira que la tâche en question requiert une charge cognitive trop élevée pour être réalisée de manière optimale. On peut alors voir la relation entre l'apprentissage et la mémoire comme une balance à plateaux. Si le poids de la tâche demandée est trop élevé par rapport à celui des capacités mnésiques, alors il sera plus difficile de rééquilibrer les deux plateaux. Lors de la création d'une tâche, il faut trouver le juste milieu de manière à ce que celle-ci soit suffisamment exigeante pour permettre un gain de connaissances et assez accessible pour être réalisée correctement. Cela dépend également du type d'informations que l'on propose à l'apprenant. La charge cognitive sera par exemple plus élevée dans le cas de la réalité virtuelle qui propose des informations visuelles dynamiques complexes, associées à des informations sonores, verbales et à de l'interactivité plutôt que dans le cas d'une association d'images statiques à du texte. La mémoire est évidemment omniprésente lorsque nous traitons l'information, quelle que soit sa nature, elle est donc présente dans chacun des principes et modèles que nous allons aborder. L'un des principes centraux sur lequel nous nous sommes appuyés dans la conception de nos dispositifs expérimentaux est celui de la cohérence, c'est donc celui que nous allons aborder plus en détail dans le paragraphe suivant.

1.4 La cohérence

Les supports multimédias sont composés d'un nombre important d'informations à traiter avant de pouvoir en créer une représentation cohérente. Il convient donc d'organiser la présentation de ces informations multiples de manière à ce qu'elles puissent être intégrées de manière optimale, sous forme de schéma, par l'apprenant. « Le principe de cohérence » renvoie à l'idée que l'apprentissage multimédia serait plus efficace lorsque les informations non pertinentes et/ou superflues sont exclues du dispositif éducatif, et lorsque les informations sont sémantiquement organisées. L'apprentissage est entravé quand des images, des sons, des musiques ou encore des mots intéressants (voir motivants et plaisants émotionnellement) mais incohérents sont ajoutés dans une présentation multimédia (Mayer, 2002). En d'autres termes, l'apprenant doit pouvoir traiter au mieux les informations essentielles à la compréhension sans que son apprentissage ne soit entravé par des éléments superflus accroissant la charge cognitive extrinsèque. Sont considérées comme essentielles les informations multiples qui pour un apprenant particulier, ne peuvent être intelligibles lorsqu'elles sont isolées (Sweller, Ayres,

Kalyuga & Chandler ; 2003). Aussi, avant de mieux préciser le principe de cohérence il est indispensable de rappeler la théorie du double codage (Figure 1.3). Selon cette théorie, la cognition comprend deux systèmes distincts : un système verbal spécifique au traitement direct du langage et un système non verbal spécifique au traitement non linguistique d'objets et d'évènements. Ces deux systèmes seraient composés d'unités de représentation internes qui s'activeraient lorsqu'on reconnaît, qu'on manipule ou qu'on pense à des mots ou à des objets.

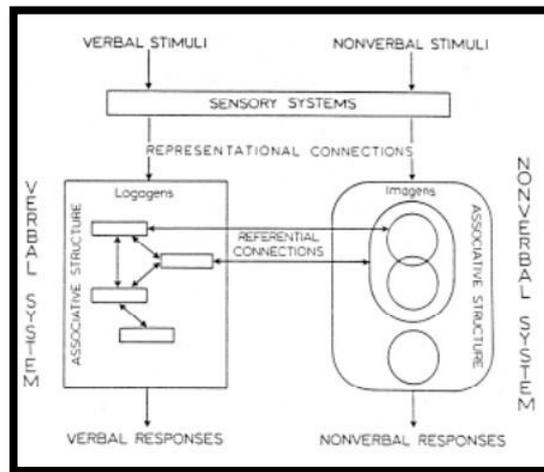


Figure 1.3. Modèle du double codage de Paivio (2006).

Les représentations seraient alors connectées à une entrée sensorielle (input) et à un système de réponse en sortie (output). Ainsi, chaque système peut fonctionner de manière indépendante ou complémentaire dans le but de médiatiser les comportements, qu'ils soient verbaux ou non-verbaux (Paivio,1986,2006).

Ce modèle est étroitement lié au modèle de compréhension texte-image de Schnotz & Bannert (2003) qui est composé de deux branches à la fois distinctes et interconnectées (voir Figure 1.4). Elles ne renvoient pas à seulement à deux canaux sensoriels, mais à deux systèmes de représentations conjoints. A gauche se trouve la branche de représentations descriptives basée sur le traitement symbolique (le texte) et à droite la branche de représentations décriptives basée sur la mise en correspondance des structures analogiques (images et schémas). Concernant le texte, le lecteur va construire une représentation mentale de la structure de surface pour ensuite générer une représentation propositionnelle du contenu sémantique. Pour finir, il va construire un modèle mental de ce qui est décrit dans le texte. En bref, il va sélectionner l'information pertinente, l'organiser et créer un modèle mental cohérent. En ce qui concerne les images et schémas, pour les comprendre l'apprenant va tout d'abord créer une représentation mentale

visuelle de ce qu'il voit. Il va ensuite créer un modèle mental ainsi qu'une représentation propositionnelle de ce que représente l'image par un traitement sémantique. Selon les auteurs, dans le processus perceptif, les informations pertinentes sont sélectionnées pour ensuite être organisées visuellement par des routines visuelles automatisées.

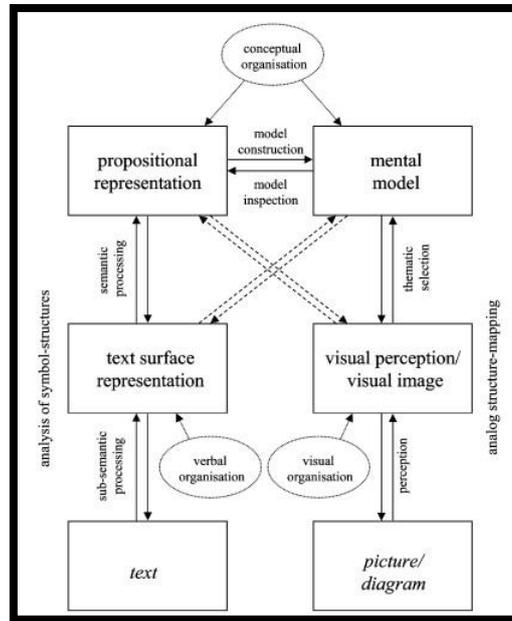


Figure 1.4. Modèle de compréhension texte image (Schnitz & Bannert, 2003).

Les auteurs partent du principe que pour construire un modèle mental, une transition est nécessaire entre la représentation descriptive et la représentation dépicative. Une fois que le modèle mental verbal ou imagé est construit, alors les informations extraites peuvent être recodées sous forme de représentations propositionnelles et analogiques interagissant continuellement entre elles.

C'est sur cette base que Mayer & Moreno (1998) aborderont la notion de double canal sensoriel. Selon eux, le système de traitement de l'information humaine est composé de deux canaux distincts : un canal destiné à traiter les informations auditives et les représentations verbales, et un canal destiné à traiter les entrées et les représentations visuelles. Ces canaux traiteraient donc l'information visuelle et auditive séparément et auraient tous deux des capacités limitées.

Pour illustrer l'impact du type d'information visuelle verbale et imagée sur l'apprentissage, on peut évoquer l'étude de Mayer, Bove, Bryman, Mars & Tapangco (1996) qui ont réalisé trois expériences sur le thème de la météorologie, plus précisément sur la formation de la foudre (Figure 1.5). Dans la première expérimentation, les auteurs présentaient une suite d'images

illustrant ce phénomène, accompagnées de différents types d'instructions. Les chercheurs ont créé quatre groupes expérimentaux qui étudiaient les mêmes images, mais dont les instructions variaient. Un groupe avait accès à un extrait complet et à un résumé du phénomène, le deuxième étudiait l'extrait complet uniquement, le troisième étudiait uniquement le résumé et le quatrième ne disposait d'aucune information verbale pour accompagner les images.

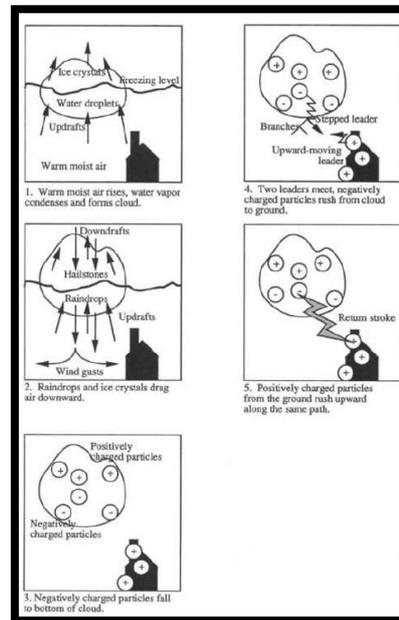


Figure 1.5. Illustration du phénomène de formation de la foudre utilisé par Mayer & al. (1996).

Il s'avère que pour la tâche de rappel, les apprenants n'ayant reçu aucune instruction verbale avaient de moins bonnes performances que les autres groupes. Les performances étaient plus élevées lorsque les images étaient accompagnées d'un résumé, et aucune différence n'était observée entre les conditions extrait + résumé et extrait seul. Concernant le transfert d'informations, les résultats ont montré que les illustrations annotées guidaient l'apprenant vers l'information explicative. Lorsqu'il n'y avait pas d'information verbale distractive, alors l'apprenant pouvait se focaliser sur les éléments clés et comprenait mieux le phénomène présenté. Dans l'extrait de texte complet, l'apprenant pouvait s'intéresser à des informations non explicatives qui venaient surcharger ses ressources attentionnelles, il avait donc plus de difficultés à encoder les informations clés.

Afin d'analyser plus finement les traitements visuels et verbaux, ces mêmes auteurs ont réalisé une seconde expérience dans laquelle les élèves devaient lire le même extrait de texte et le même résumé que dans la première expérience, mais cette fois-ci, ils étaient accompagnés ou

non d'illustrations. Les chercheurs ont donc constitué quatre groupes : les deux premiers groupes étudiaient soit l'extrait de texte accompagné du résumé, soit le résumé seul, mais ne bénéficiaient pas des illustrations. Les deux autres groupes étudiaient soit les images accompagnées de l'extrait de texte et du résumé, soit les images accompagnées uniquement du résumé. Ici, le résumé le plus efficace était celui qui regroupait les aspects verbaux et visuels de l'information plutôt que celui qui résumait l'explication seulement en mots ou en images. L'intégration de ces deux systèmes de représentation (verbal et visuel) aidait les sujets novices à construire une représentation cohérente de l'explication scientifique.

Dans une troisième expérience, Mayer, Bove, Bryman, Mars & Tapangco (1996) ont évalué l'impact du nombre de mots contenus dans une explication verbale écrite sur la qualité de l'apprentissage. Les auteurs ont créé trois groupes auxquels ils ont proposé soit le même résumé que dans les deux premières expériences, soit le résumé comprenant une description plus approfondie (50 mots supplémentaires), soit le résumé accompagné de légendes supplémentaires, ces légendes étant des extraits du texte complet (550 mots supplémentaires). Chaque extrait était présenté conjointement à l'image correspondante. Ici, les élèves ont montré de meilleures performances de compréhension dans la condition « résumé » initial, que dans les autres conditions. Plus le nombre de mots ajouté était important, plus l'élève avait de difficultés à comprendre les explications initiales. Les mots supplémentaires surchargeaient la mémoire de travail verbale ce qui limitait les ressources attentionnelles au détriment de l'explication verbale pertinente. Concernant les performances de transfert, les élèves ont obtenu les mêmes résultats pour le résumé initial que pour le résumé auquel 50 mots avaient été ajoutés. Cependant, ces deux groupes ont été plus performants que celui ayant étudié le résumé étendu de 550 mots. On peut donc conclure que quand le nombre de mots était trop important, il était plus difficile pour les apprenants d'abstraire les informations verbales principales, pour les mettre en relation avec les informations visuelles. Pour favoriser un bon apprentissage, il est plus judicieux de ne pas surcharger les ressources attentionnelles des apprenants en ne présentant que les informations verbales cohérentes, nécessaires à la compréhension.

Ce même thème de formation de la foudre a été repris par Harp & Mayer (1997) qui ont étudié le caractère attractif du texte et des illustrations sur l'apprentissage. Les auteurs ont présenté un matériel principal, identique à celui de leur première expérience à un groupe. Trois autres groupes étudiaient soit le matériel principal accompagné d'un texte attrayant, soit accompagné d'illustrations attrayantes, soit accompagné à la fois de texte et d'images attrayants. Ici, le groupe ayant reçu la présentation principale obtenait de meilleurs résultats de compréhension

que les trois autres groupes. Les groupes ayant étudié soit le texte soit les illustrations attrayantes obtenaient les mêmes résultats, ils réalisaient cependant de meilleures performances d'apprentissage que le groupe ayant vu à la fois le texte et l'illustration attrayante. Concernant les aptitudes de transfert, le groupe ayant reçu le matériel principal obtenait les mêmes résultats que ceux ayant reçu des illustrations attrayantes. Leurs performances étaient cependant plus élevées que celles du groupe ayant vu un texte attrayant, qui eux-mêmes obtenaient des performances plus élevées que le groupe ayant étudié à la fois le texte et les illustrations attrayantes. Ici encore, on remarque que l'ajout d'informations attrayantes mais inutiles surchargeait les ressources attentionnelles de l'apprenant et perturbait donc l'apprentissage.

Sung & Mayer (2012) ont étendu ces travaux en comparant des images de différentes natures dans une leçon web sur l'histoire et la définition de l'éducation à distance. Les auteurs ont proposé des images : instructives, décoratives, attrayantes ou encore une modalité sans image (Figure 1.6).



Figure 1.6. Leçon web proposée par Sung & Mayer (2012).

Les résultats ont montré une corrélation positive en faveur de la pertinence des images proposées en tâche de rappel. Les performances des participants étaient plus élevées lorsque les illustrations proposées étaient instructives plutôt que décoratives ou attrayantes. D'ailleurs, aucune différence n'a été constatée entre la condition dans laquelle aucune image n'était présentée et la condition dans laquelle des images attractives étaient présentées. S'il n'y a pas de corrélation entre les performances de rappel et la présence d'illustrations, on peut tout de même

noter que lorsqu'on ajoute des images à un support éducatif, il est préférable que celles-ci soient instructives et donc cohérentes de manière à permettre un apprentissage plus efficace sans surcharger les ressources cognitives.

La question que l'on peut également se poser est la suivante : qu'en est-il du canal verbal auditif ? est-ce que les informations verbales entendues sont susceptibles d'altérer ou d'améliorer l'apprentissage selon leur caractère utile ou simplement attractif dans un dispositif éducatif ? C'est le point sur lequel Moreno & Mayer (2000) se sont centrés en abordant de nouveau le thème de la foudre. Les auteurs ont comparé quatre groupes d'apprenants auxquels ils présentaient les images du processus et les explications correspondantes selon 4 modalités : narration, narration + son environnemental, narration + musique et pour finir, narration + son environnemental + musique. Les résultats ont montré que les apprenants mémorisaient moins bien lorsqu'on ajoutait de la musique plutôt que des sons environnementaux. Aucune différence n'était à noter entre les groupes qui avaient entendu ou non des sons environnementaux. Le même résultat a été obtenu dans les épreuves de transfert. Si les sons environnementaux ne contribuaient pas à rendre la leçon plus intelligible, on peut dire que lorsqu'ils étaient présentés simultanément à de la musique ou lorsque l'on présentait simplement de la musique, la compréhension des élèves s'en trouvait perturbée. La musique représenterait alors une charge extrinsèque préjudiciable pour l'apprentissage que l'on peut expliquer en partie par le fait qu'elle n'avait pas de rapport direct avec les informations visuelles et verbales présentées. De ce fait, elle surchargeait l'attention de l'apprenant qui n'aurait alors plus assez de ressources cognitives pour traiter correctement les informations essentielles à la compréhension du processus.

Quant à la signification de l'information verbale, est-il plus judicieux de la présenter de manière orale plutôt qu'écrite ou de combiner ces deux formes ? c'est la question à laquelle Mayer, Heiser & Lonn (2001) ont tenté de répondre. Les auteurs ont réalisé quatre expériences sur le thème du déroulement du cycle de l'eau (Figure 1.7). Dans leur première expérience, les chercheurs sont partis du principe qu'il valait mieux présenter l'information verbale de manière orale plutôt que de la présenter de manière à la fois orale et écrite. Les auteurs pensaient également que l'ajout d'informations visuelles attrayantes mais inutiles perturberait l'apprentissage. Quatre groupes expérimentaux ont été constitués selon les caractéristiques suivantes : les participants étudiaient des images avec détails attrayants ou sans détails attrayants. Dans chacune de ces conditions, ils entendaient soit une explication orale seule, soit

une explication orale accompagnée de sous-titres. A noter que les sous-titres correspondaient exactement à la narration orale.

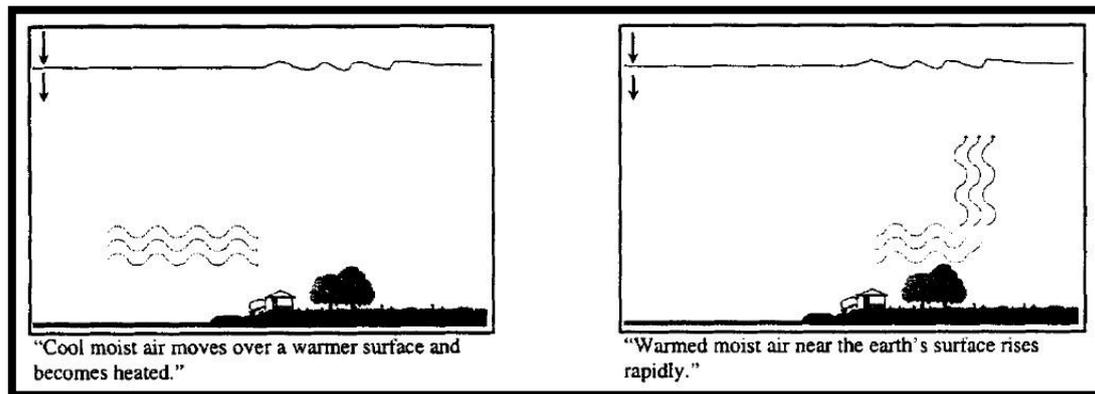


Figure 1.7. Exemple du matériel proposé par Mayer sur le cycle de l'eau.

Les résultats de l'étude ont montré que, concernant la rétention immédiate de l'information comme son transfert, les sous-titres et les détails attrayants des images conduisaient à de moins bonnes performances qu'un design plus simple avec une explication orale seule. Au-delà de l'effet de redondance, les auteurs ont évoqué l'effet de cohérence : les apprenants semblaient s'impliquer plus fortement dans la compréhension du matériel quand celui-ci était présenté de façon à mettre en valeur les structures conceptuelles du contenu du document plutôt que l'aspect attractif. Il vaut donc mieux éviter de proposer des informations superflues qui vont perturber la mise en place et l'assimilation de schémas mentaux appropriés. En fait, proposer des informations superflues si attractives soient-elles, va provoquer une surcharge des processus extrinsèques, ce qui va empêcher les processus essentiels de fonctionner efficacement, c'est à dire : l'élaboration et l'intégration de schémas cohérents.

Dans une troisième expérience, Mayer & al (2001) ont proposé une tâche d'apprentissage sur la formation de la foudre. Ils ont mis en œuvre le même type de matériel que dans leur première expérience à savoir : une succession d'images sans détails attrayants. Ce matériel était comparé à une même leçon mais cette fois-ci, de courtes animations pertinentes à propos du processus étaient intercalées entre les images. Les résultats ont montré que le rappel des participants n'était pas affecté par le type de leçon proposé, en revanche, les auteurs ont observé de moins bonnes performances à la tâche de transfert lorsque de courtes animations étaient ajoutées à la leçon. Ces résultats sont en adéquation avec ce qui a été observé précédemment, c'est-à-dire que même s'ils sont pertinents les éléments ajoutés perturbent l'apprentissage.

En 2007, Lehman & al. ont également étudié l'effet de l'ajout d'informations attrayantes sur l'apprentissage en abordant le thème de la formation de la foudre, reprenant et adaptant le matériel utilisé par Harp & Mayer (1998). Les auteurs ont comparé les performances d'apprenants selon qu'on leur présente le texte principal décrivant la formation de la foudre et les conditions les plus favorables à leur apparition, ou des phrases contenant des détails attractifs répartis dans le texte initial. Les auteurs ont montré que les participants qui étudiaient le texte principal pourvu de détails attrayants, lisaient plus rapidement les phrases du texte principal que ceux qui n'avaient lu que ce texte principal. Aussi, la présence de phrases attrayantes dans le texte semblait interférer avec le rappel des éléments importants du texte principal. En effet, les participants ayant lu un texte pourvu de phrases attrayantes rappelaient moins d'informations sur la formation des éclairs que ceux qui avaient lu le texte principal. Les auteurs ont donc conclu que les détails attrayants interrompaient la construction de la chaîne causale interne du processus présenté, ce qui réduisait la compréhension holistique du texte principal. Les détails réduisaient alors la capacité des lecteurs à mémoriser puis intégrer les aspects importants relatifs à la formation des éclairs. En somme, on peut dire que les informations attrayantes augmenteraient la charge extrinsèque du fait de leur caractère non pertinent, ce qui appauvrirait l'apprentissage.

Ces résultats ont été nuancés par les travaux de Park, Moreno, Seufert & Brünken (2011) qui ont étudié la même problématique. Les auteurs ont indiqué que le traitement cognitif consistant à sélectionner les informations pertinentes puis à organiser ces informations en un modèle mental cohérent pouvait être affecté par des détails attrayants mais que cela n'avait pas toujours un effet négatif. Selon les auteurs, cela peut être positif si les apprenants ont assez de ressources disponibles pour utiliser des informations non redondantes, intéressantes, même si elles ne sont pas pertinentes. Quand la charge extrinsèque reste faible malgré la présence de détails non pertinents alors les activités de traitement essentielles ne sont pas affectées.

Si les éléments attrayants mais inutiles peuvent perturber l'apprentissage, il est tout de même indispensable de prendre en compte les différences interindividuelles des apprenants afin de comprendre si celle-ci peuvent permettre de nuancer les résultats obtenus dans les précédentes études. C'est le point sur lequel Sanchez & Wiley (2006) se sont penchés afin de relier les capacités de mémoire de travail à la compréhension de documents multimédias. Ils ont proposé une leçon sur le thème de la formation des glaciers. Les élèves devaient lire un texte explicatif à ce sujet selon trois modalités : soit ils voyaient le texte seul, soit le texte était accompagné d'illustrations conceptuellement pertinentes, soit il était accompagné d'illustrations attrayantes.

Pour chaque condition, les auteurs ont comparé les performances des participants selon leur niveau de capacité de mémoire de travail. Les résultats de l'étude ont montré que le caractère attractif des illustrations n'était préjudiciable que pour les élèves ayant une faible capacité de mémoire de travail. On peut penser que ces élèves avaient plus de difficultés à « contrôler » leur attention pour se concentrer sur les informations pertinentes lorsqu'on leur présentait des informations supplémentaires qu'il serait préférable d'ignorer ou de traiter d'une autre manière. Cela a été vérifié par l'étude du mouvement des yeux, réalisée lors de cette expérience. En effet, les apprenants ayant de bonnes capacités de mémoire de travail consultaient moins d'informations incohérentes que les apprenants ayant de faibles capacités de mémoire de travail. Ces derniers auraient alors plus de difficultés à ignorer les informations incohérentes.

Une étude similaire a été menée par Doolittle & Altsaedter (2009) sur le thème du processus de formation de la foudre au moyen d'animations multimédia (Figure 1.8). Les auteurs ont proposé un tutoriel au format flash basé sur le matériel de Mayer & Chandler (2001).

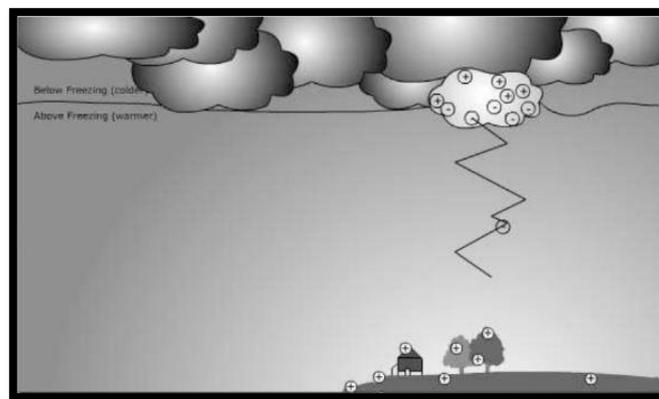


Figure 1.8. Illustration du phénomène de foudre de l'étude de Doolittle & Altsaedter (2009).

Il s'agissait d'une animation décrivant le processus dynamique de la formation des éclairs. Les auteurs ont constitué deux groupes d'apprenants selon leurs capacités de mémoire de travail : haute ou faible. Chaque groupe d'élèves étudiait soit une animation accompagnée d'une narration explicative, soit une animation accompagnée d'une narration et de détails attrayants. Les détails attrayants étaient des sons environnementaux dit extrinsèques, mais très réalistes, comme le bruit du vent, de la pluie ou du tonnerre. Des détails visuels étaient également présents dans cette condition : le ciel s'obscurcissait, les nuages devenaient plus nombreux et un flash lumineux accompagnait la foudre. Les résultats obtenus ont montré que les performances de rappel et de transfert étaient plus élevées pour les participants ayant des capacités de mémoire de travail élevées, que pour ceux qui avaient des capacités plus faibles. De plus contrairement

à ce que suggère le principe de cohérence, aucune différence n'a émergé entre les deux conditions de présentation de l'information multimédia, c'est-à-dire que les détails attrayants ajoutés n'ont pas amoindri les performances des apprenants. Ici, contrairement à ce que les précédentes études ont pu démontrer, rien n'indique que la condition multimédia affecte les élèves quel que soit leur niveau de mémoire de travail. Ce seraient donc les capacités de mémoire de travail elles-mêmes qui détermineraient les performances de rétention et de transfert des apprenants. Il faut donc veiller à ne pas surcharger la mémoire de travail des apprenants pour favoriser un apprentissage efficace. Pour cela, il est malgré tout indispensable de respecter les résultats mis en évidence dans le cadre du principe de cohérence comme par exemple en évitant le surplus d'informations verbales écrites. En bref, pour favoriser l'apprentissage multimédia mieux vaut préserver les ressources cognitives extrinsèques de manière à ce que les informations pertinentes soient traitées et intégrées correctement.

En résumé, il est important de proposer des environnements d'apprentissage comportant le moins d'informations inutiles possible tout en prenant en compte les capacités limitées de chaque canal sensoriel. Malgré tout, certains processus sont très complexes à intégrer même lorsqu'on ne présente que les informations essentielles qui les composent. C'est le cas du matériel de réalité virtuelle que nous allons présenter par la suite. Nous avons pris soin de ne modéliser que les informations pertinentes cependant, l'environnement virtuel et le processus décrit restent très complexes et donc potentiellement difficiles à traiter et à intégrer. Dans ce cas, il peut être judicieux d'accompagner l'apprenant dans la création d'une chaîne causale cohérente en le guidant, ce qui nous renvoie au principe de signalisation/guidage.

Les travaux menés sur la cohérence se sont principalement centrés sur le rôle des informations attrayantes ou redondantes mais inutiles ou ajoutées et non reliées aux informations principales des documents présentés. Très peu de travaux se sont penchés sur le rôle de la cohérence sémantique entre les éléments d'information délivrés dans la leçon. Cette dimension paraît cruciale pour l'apprentissage de notions complexes comme le fonctionnement d'un écosystème dans lequel de nombreux éléments d'informations sont en interaction les uns avec les autres. Cette question fera l'objet de l'expérience 1 de la partie empirique, dans le contexte de la réalité virtuelle.

1.5 La signalisation ou « cueing »

Le principe de signalisation consiste à insérer des indices visuels ou auditifs (verbaux) dans le dispositif éducatif ou le document multimédia de manière à attirer l'attention de

l'apprenant directement et sélectivement, sur l'information pertinente. De cette manière on peut tenter de le guider, pas à pas, de sorte à lui permettre d'extraire une représentation cohérente du processus étudié.

Scheiter & Eitel (2010) ont réalisé une étude dont l'objectif était d'analyser l'intégration texte/image en introduisant une signalisation au sein des deux modalités sensorielles (visuelle et verbale) dans le but d'aider la mise en relation du texte avec l'image et donc, l'intégration conjointe de ces deux modalités. Dans cette étude, les auteurs ont proposé un texte accompagné d'une illustration sur le thème du fonctionnement cardiaque (Figure 1.9). Un groupe étudiait le texte et l'illustration présentés conjointement dans un format conventionnel. Le deuxième groupe étudiait le même matériel à la différence que l'illustration comportait des légendes sous forme de texte de différentes couleurs. Les mots correspondants présents dans le texte étaient indiqués par la même couleur de police d'écriture.

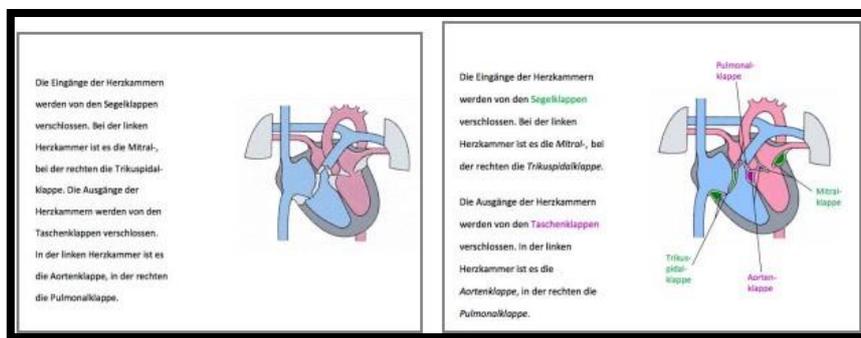


Figure 1.9. Le fonctionnement cardiaque par Scheiter & Eitel (2010)

Les résultats ont montré un effet positif du guidage ; selon les auteurs, les signaux amélioraient l'intégration cross modale ou intermodale de l'apprentissage des textes et des schémas. L'analyse des mouvements oculaires a permis aux chercheurs de conclure que ces résultats sont dus au fait que ce type de signalisation amenait les apprenants à se référer plus rapidement au schéma en accroissant l'attention qui leur était accordée.

Ce type de signalisation a également été étudié par Désiron, Bétrancourt & de Vries (2021). L'objectif des chercheurs était d'évaluer l'effet de la signalisation sur l'apprentissage, en fonction du degré de cohérence du texte présenté sur le thème de la navigation fluviale (Figure 1.10).

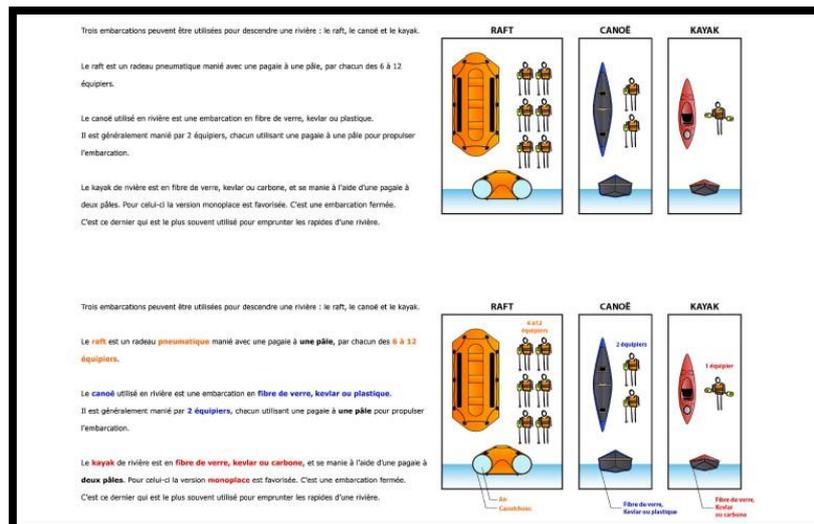


Figure 1.10. Exemple de matériel proposé par Désiron, Bétrancourt & de Vries (2021), respectivement : support multimédia sans, et avec signalisation.

Quatre groupes expérimentaux ont été constitués : les apprenants étudiaient le support avec ou sans signalisation, et le texte était proposé soit dans une version de basse cohérence, soit dans une version cohérente. Si la cohérence du texte n'a globalement pas montré d'effet sur la compréhension, les apprenants qui ont bénéficié de la signalisation ont obtenu des résultats plus élevés que ceux dont le document ne comportait pas de signalisation. La signalisation aurait donc bien un effet positif sur les apprentissages multimédias.

Naumann, Richter, Flender, Christmann & Groeben (2007) ont apporté une autre perspective au guidage réalisé au sein de textes à travers deux expériences sur le thème de la perception visuelle. Leur but était de comparer le rôle des signaux dans la compréhension d'un hypertexte en fonction des compétences en lecture des apprenants. Dans la première expérience, ils ont comparé deux groupes d'étudiants : un groupe devait apprendre à l'aide d'un texte linéaire et l'autre à partir d'un hypertexte. Dans la seconde expérience, les auteurs ont fait varier le niveau de complexité de l'hypertexte présenté. Les auteurs ont comparé le matériel hypertexte initial à une version plus restreinte dans laquelle les principaux signaux ont été retirés, ne laissant que certaines fonctionnalités principales accessibles. Les chercheurs ont étudié ces conditions auprès d'élèves ayant de fortes ou faibles compétences de lecture. Les résultats ont montré que l'hypertexte n'avait pas d'effet sur l'apprentissage des apprenants ayant de bonnes compétences de lecture par rapport à l'apprentissage à l'aide d'un texte linéaire typique. Cependant, les élèves dont les compétences de lecture étaient plus faibles ont bénéficié de l'hypertexte pourvu de signaux si bien que leurs résultats se sont équilibrés avec ceux des élèves dont les compétences

de lecture étaient meilleures. Lorsque l'hypertexte ne contenait pas de signaux, les stratégies employées par les élèves étaient moins efficaces. On peut donc dire que la signalisation facilite le processus de compréhension visant à extraire la structure générale d'un texte en favorisant l'organisation, la sélection et l'intégration des informations qui doivent être apprises.

Mautone & Mayer (2001) ont testé le guidage à l'aide de différents dispositifs sur le thème de l'aviation. Ils ont réalisé trois expériences ayant pour thème le décollage d'avion (Figure 1.11).

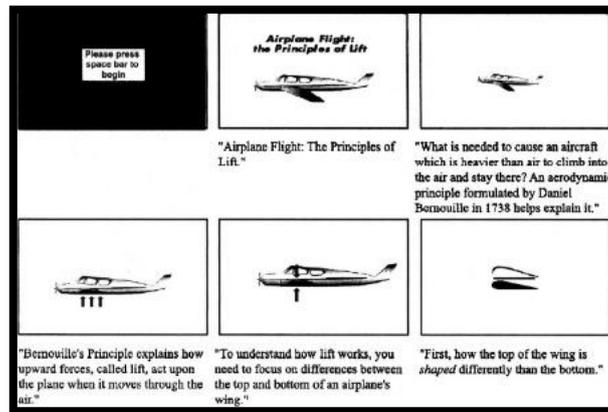


Figure 1.11. Leçon d'aviation par Mautone & Mayer (2001).

Dans la première étude, les chercheurs proposaient un texte brut qu'ils comparaient au même texte accompagné d'un résumé, dont les principales parties étaient indiquées par un titre et dont les mots clés étaient indiqués en gras et en italique. Ici, les participants ayant bénéficié du guidage avaient des performances plus élevées en tâche de transfert que ceux ayant lu le texte original. Dans la deuxième expérience, les auteurs ont repris ce même matériel à la différence que dans cette deuxième étude, le texte n'était pas écrit mais écouté. Pour le mode guidé, les mots auparavant indiqués en gras et italique étaient énoncés de manière plus lente et accentuée. Dans cette version, les élèves du groupe guidé ont obtenu des résultats plus élevés aussi bien en tâche de rétention que pour la tâche de transfert. Pour terminer, les auteurs ont réalisé une troisième étude reprenant le même matériel, dans laquelle ils ont proposé une animation multimédia à quatre groupes expérimentaux. L'animation comme la narration étaient guidées ou non. Les résultats ont montré une interaction en tâche de rétention, entre la signalisation de la narration et celle de l'animation. Les résultats obtenus ont révélé que la narration guidée permettait une meilleure rétention des informations lorsque l'animation n'était pas guidée. Par contre, la signalisation présente dans la narration nuisait aux performances de rétention lorsque l'animation comportait également des signaux. Pour la tâche de transfert, de meilleures

performances ont été observées quand l'animation et la narration étaient guidées que lorsqu'il n'y avait aucun guidage. Aussi, si aucun effet n'a été relevé, que l'animation comporte ou non des signaux, il est à noter que les performances de transfert étaient plus élevées lorsque la narration était signalée plutôt que lorsqu'elle ne l'était pas.

On retrouve également un effet bénéfique du guidage dans les travaux de de Koning & al. (2010) qui ont comparé un mode signalé visuellement à un mode non guidé à partir d'une animation sur le thème du système cardio vasculaire. Les performances relatives au mode signalé étaient supérieures à celles du mode non guidé pour les tests de rétention, d'inférence ainsi que de transfert.

Jamet, Gavota & Quaireau, (2008) ont utilisé un type de guidage différent, alliant saillance perceptuelle et séquentialité de la présentation d'images sur le thème de l'anatomie cérébrale responsable de la production du langage (Figure 1.12).

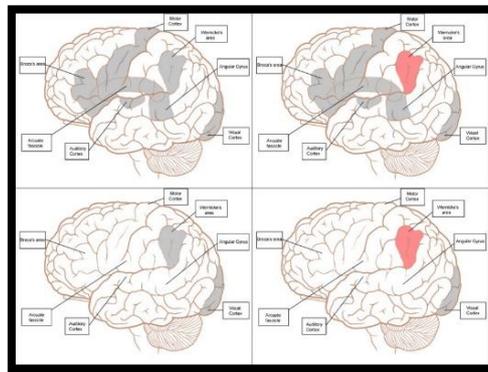


Figure 1.12. Illustration cérébrale de la production de langage (Jamet & al., 2008).

Les auteurs ont montré que la saillance perceptuelle moyennant une coloration rouge des aires cérébrales offrait de meilleurs résultats de rétention de la fonction et du processus présenté que quand il n'y avait pas de saillance perceptuelle. Aussi, les performances de rétention étaient plus élevées lorsque l'information était présentée de manière séquentielle plutôt que de manière statique et donc complète. Par contre, aucun effet n'a été souligné en tâche de transfert. La présentation séquentielle améliorerait l'organisation des représentations en mémoire (Bétrancourt, Bisseret & Faure, 2001).

Cependant, Ozcelik, Arslan-Ari & Cagiltay, (2010) ont montré que lors d'une tâche dans laquelle ils présentaient une illustration de moteur à réaction légendée (Figure 1.13), le fait de changer temporairement la couleur (coloration rouge) des éléments de la légende entraînait de

meilleures performances en tâche de transfert que lorsqu'aucune saillance perceptible n'était mise en place. Ils n'ont cependant pas relevé de différences de performances au test de rétention.

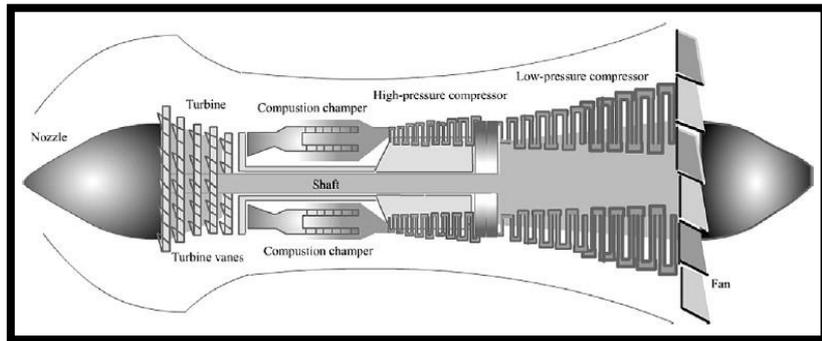


Figure 1.13. Illustration de moteur à réaction (Ozcelik & al., 2010)

Doolittle & Altstaedter (2009) avaient pour objectif de valider l'effet de signalement selon lequel l'apprentissage serait facilité par la présence de signaux guidant l'attention des apprenants en mettant en évidence la structure de l'information du mécanisme présenté. Le matériel qu'ils ont proposé illustre un mécanisme de freinage automobile (Figure 1.14)

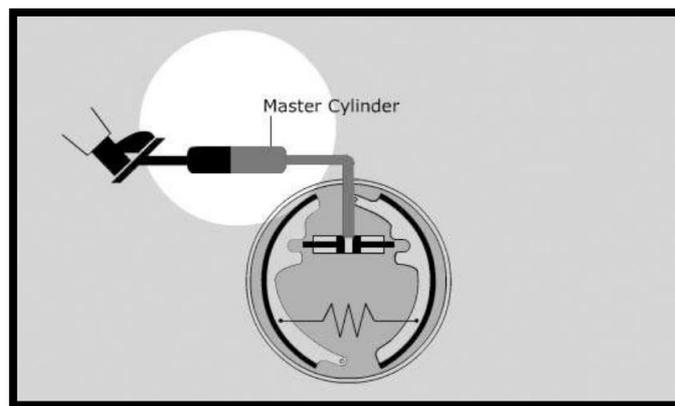


Figure 1.14. Mécanisme de freinage de voiture proposé par Doolittle & Altstaedter (2009).

Un groupe étudiait une animation flash (basée sur le matériel utilisé par Mayer & Anderson, 1992) accompagnée d'une narration explicative sur le fonctionnement du mécanisme. Un autre groupe étudiait ce même matériel accompagné de mots clés présentés spatialement proche des éléments énoncés, ainsi que d'une mise en lumière visuelle dont le rôle était d'attirer l'attention des apprenants sur les éléments pertinents de l'animation (voir le rond blanc, « spotlight » de la Figure 1.14). Les résultats n'ont montré aucun effet du guidage proposé. En effet, le type de

matériel utilisé, sa simplicité ou sa complexité, peut jouer un rôle considérable dans l'apprentissage.

Comme l'ont montré Boucheix & Lowe (2010), il est important de créer un guidage qui soit en adéquation avec la présentation proposée. Leur étude portait sur l'apprentissage du fonctionnement d'un mécanisme de piano, et leur objectif était de comparer deux types de signalisation : une signalisation dynamique par diffusion progressive d'une flèche de couleur sur les sections appropriées du mécanisme, et une signalisation fléchée (flèches externes au mécanisme). Pour cela, les auteurs ont réalisé deux expériences sur la compréhension du fonctionnement d'un mécanisme animé et interactif de piano (Figure 1.15).

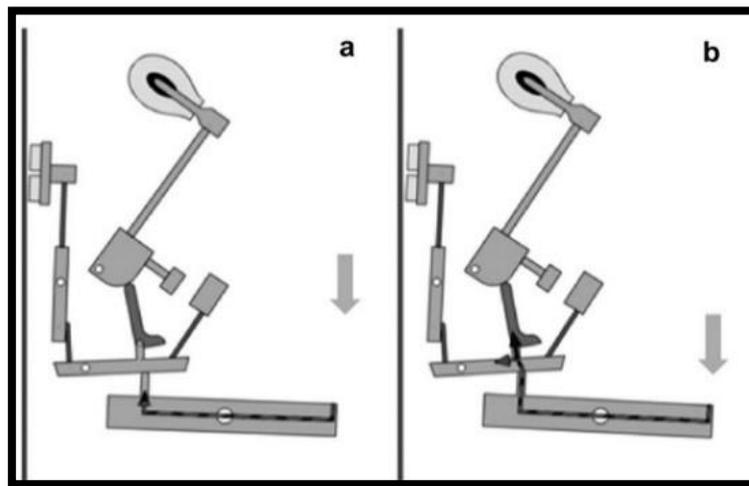


Figure 1.15. Aperçu du mécanisme de piano proposé par Boucheix & Lowe (2010).

Dans la première étude, les auteurs ont comparé la version initiale de l'animation interactive du mécanisme à deux versions guidées. Pour les animations guidées, un groupe étudiait un mécanisme dont les mouvements étaient indiqués par des flèches ; un autre groupe étudiait un mécanisme dont les mouvements étaient signalés par le biais de la diffusion progressive dynamique d'une flèche colorée sur les éléments pertinents du mécanisme. Les auteurs ont montré que les performances de compréhension étaient supérieures lorsque les mouvements étaient signalés par la diffusion progressive d'une flèche de couleur, plutôt que lorsque les mouvements étaient indiqués par des flèches externes au mécanisme. Les résultats n'ont pas montré de différences de performances d'apprentissage entre la signalisation par présentation de flèches externes et l'absence de signalisation. Les auteurs ont répliqué cette expérience en testant la synchronisation entre la signalisation et le mouvement du mécanisme actionné par l'élève. Ils ont comparé chacun des types de signalisation avec un mode où la signalisation

apparaissait immédiatement, dès que le participant actionnait le mécanisme, et un mode ou la signalisation suivait le mouvement réalisé par le mécanisme de manière synchronisée, suite à l'action de l'apprenant. Les résultats ont montré que les deux versions mettant en œuvre une signalisation synchronisée offraient de meilleures performances de compréhension. Selon les auteurs, l'efficacité du guidage dépendait de la coordination spatio-temporelle entre le début de l'évènement animé pertinent du mécanisme et l'attention de l'apprenant.

Boucheix, Lowe, Putri & Groff, (2013) ont de nouveau étudié le guidage par le biais du mécanisme de piano animé. Les auteurs ont constitué trois groupes expérimentaux : un groupe étudiait le mécanisme dont différentes entités étaient présentées successivement. Dans ce cas, un groupe de pièces du mécanisme était clairement visible pendant que le reste du mécanisme était affadi (Figure 1.16).

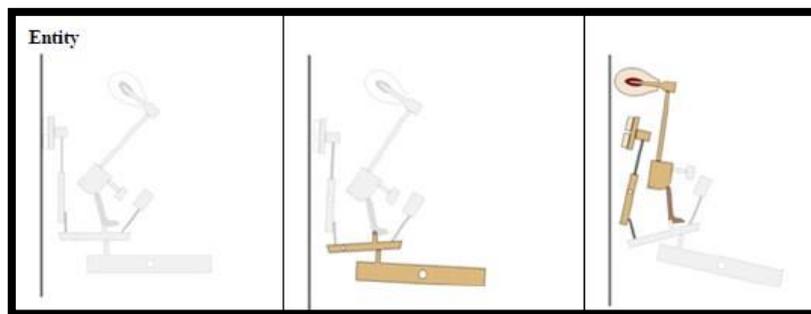


Figure 1.16. Signalisation par groupe d'éléments du mécanisme de piano.

Un autre groupe étudiait le mécanisme sous sa forme complète, avec un signalement progressif synchronisé des chaînes causales représenté par des flèches animées parcourant le mécanisme (Figure 1.17).

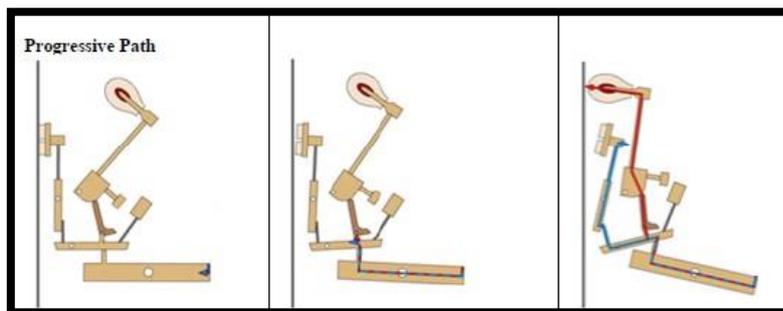


Figure 1.17. Signalisation progressive du mécanisme de piano

Le troisième groupe voyait également le mécanisme entièrement, mais dans cette version la signalisation était faite de repères colorés localisés sur des parties précises du mécanisme, et coordonnées avec le mouvement (Figure 1.18).

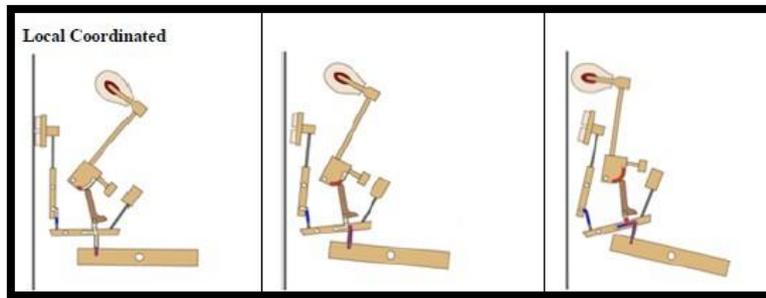


Figure 1.18. Signalisation locale coordonnée du mécanisme de piano.

Les résultats ont montré des performances plus élevées pour la signalisation progressive et la signalisation locale coordonnée que pour la signalisation par groupes d'éléments isolés concernant la tâche de compréhension et la mesure de modèle mental. L'analyse du mouvement oculaire a montré que les apprenants ayant bénéficié d'une signalisation d'unités d'évènements coordonnés avaient plus volontiers reporté leur attention sur les éléments très pertinents mais peu saillants que les apprenants ayant étudié à partir d'entités isolées. Les auteurs ont conclu qu'un guidage signalant des chaînes causales était plus efficace pour la création d'un modèle mental cohérent qu'un guidage isolant des combinaisons d'éléments.

Un autre type de guidage a été expérimenté par Jarodzka, Van Gog, Dorr, Scheiter & Gerjets (2013) sur le thème de l'apprentissage des catégories de mouvements de locomotion des poissons avec la technique EMME (Eye Movement Modeling Expert). Les auteurs ont proposé une vidéo illustrant le mouvement du poisson à trois groupes expérimentaux. Le premier groupe étudiait la vidéo initiale (groupe témoin), le deuxième groupe étudiait une vidéo dans laquelle le mouvement des yeux d'un expert était superposé à l'image sous forme de points (points de fixation) et le troisième groupe étudiait une vidéo dans laquelle le mouvement des yeux d'un expert était superposé à l'image sous forme de zones lumineuses. Les trois groupes bénéficiaient d'une narration explicative par l'expert dont le mouvement des yeux avait été préalablement enregistré. Les résultats ont montré que, comparé au groupe témoin, les deux types de signalisation (EMME) ont permis d'améliorer significativement la recherche visuelle d'informations pertinentes ainsi que leur interprétation lors de la présentation de nouveaux stimuli. Les auteurs ont conclu que l'utilisation de l'EMME permettait de favoriser

l'apprentissage et d'améliorer les performances lors de la présentation de nouveaux stimuli perceptifs.

Cependant, le guidage peut ne pas avoir d'incidence sur l'apprentissage. C'est ce qu'ont montré Harp & Mayer (1998) dans leur étude dont le but était de comparer différents types de signalisation. Ils ont proposé à des apprenants de lire un texte sur la formation de la foudre en faisant varier le type de présentation. Le texte était présenté seul pour l'un des groupes. Pour les trois autres groupes, il pouvait être agrémenté de détails de signalisation, de détails séduisants ou de ces deux éléments combinés. La signalisation consistait à rendre saillants les éléments importants du texte proposé. Les résultats obtenus ont montré d'une part que le fait de signaler les éléments importants du texte n'avait pas d'effet sur les performances des apprenants lorsque le texte comportait des détails séduisants. D'autre part, les résultats n'ont pas montré d'effet pour le guidage par la mise en lumière des éléments pertinents seuls. Le guidage ne semblait donc pas aider les élèves à se focaliser sur les informations pertinentes. Les auteurs ont indiqué que le texte initial était probablement suffisamment bien rédigé pour que les élèves le comprennent sans avoir besoin d'être guidés dans leur lecture. Dans le cas présent, la signalisation ne semble pas indispensable à un bon apprentissage.

Comme on a pu le voir, le guidage peut aider les apprenants à intégrer les informations et à créer des modèles mentaux cohérents. Cependant, il est indispensable d'adapter le type de guidage à l'information à laquelle on donne accès. On ne signalera pas de la même manière des éléments verbaux, picturaux, statiques ou animés. Il convient par exemple de respecter les exigences spatio-temporelles entre l'information initiale et les éléments de guidage ajoutés (Boucheix & Lowe., 2010). Dans le cas de la réalité virtuelle, l'apprenant est immergé dans un environnement visuellement très riche. Compte tenu de la richesse des environnements virtuels, nous nous sommes demandé si un dispositif de signalisation était susceptible d'améliorer ou d'optimiser la sélection et l'organisation par l'apprenant des informations pertinentes pour la compréhension et l'apprentissage. Nous avons donc jugé indispensable d'étudier l'impact d'un décalage temporel entre l'information verbale et visuelle sur l'apprentissage. Cette question fera l'objet de l'expérience 3 dans la partie empirique.

1.6 La contiguïté spatiale et temporelle.

1.6.1 La contiguïté spatiale.

Le principe de contiguïté spatiale et temporelle renvoie au fait qu'il serait plus facile d'apprendre des informations lorsque celles-ci sont présentées simultanément. C'est le fait de

présenter l'information verbale physiquement proche de l'information visuelle à laquelle elle correspond dans le cas où le texte est écrit. Lorsque l'information verbale est présentée de manière orale, cela renvoie au fait de présenter la narration simultanément à l'information visuelle à laquelle elle correspond

Mayer (1989) a démontré que pour intégrer correctement les informations à travers l'apprentissage d'un mécanisme de freinage, il faut que celles-ci soient présentées de façon contigüe (Figure 1.19). Il a proposé un matériel composé d'un texte explicatif accompagné soit d'un schéma comportant des mots clés, soit d'un schéma sans mots clés associés, soit de mots clés sans schémas.

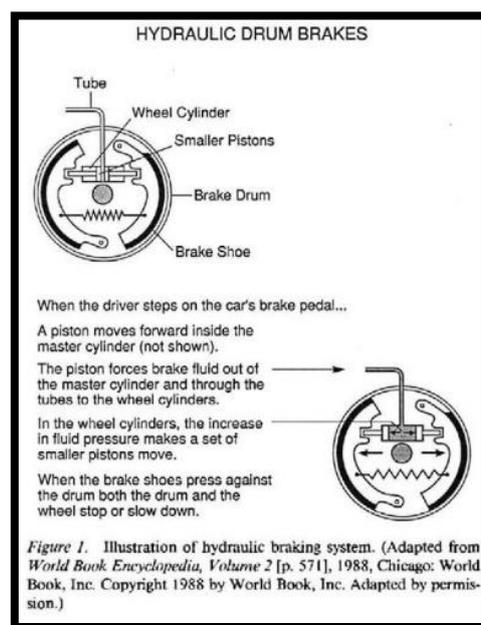


Figure 1.19. Exemple du système de freinage proposé par Mayer (1989).

Les résultats ont montré que l'illustration dans laquelle les mots clés étaient présentés de manière contigüe aux éléments correspondants était plus favorable à la création de connexions entre les différents éléments du mécanisme en tâche de transfert. Selon l'auteur, sans le schéma qui intègre les informations clés, les performances de résolution de problèmes sont amoindries. Dans le cas présent, il convient donc de présenter les informations de manière spatialement proches pour que la mise en relation des informations soit effective. Cela renvoie à la conclusion suivante : « *Le matériel pédagogique qui oblige à intégrer mentalement des sources disparates d'informations se référant mutuellement (par exemple, du texte et des diagrammes) interfère également avec l'apprentissage en détournant l'attention et en imposant une lourde charge cognitive.* » Sweller, Chandler, Tierney & Cooper, (1990).

Ceci a été mis en lumière dans les travaux réalisés par Chandler & Sweller (1991) dont l'objectif était de mettre en évidence les bénéfices d'un apprentissage à partir d'un support dont les informations étaient intégrées comparé à un support traditionnel. La tâche d'apprentissage portait sur le thème de tests d'installations électriques. Ils ont réalisé deux groupes expérimentaux auxquels ils ont présenté soit des notes d'instructions conventionnelles, soit des notes d'instructions pédagogiques sous un format intégré. Les auteurs ont remarqué que malgré une évolution des deux groupes, même après douze semaines d'accès aux notes et au matériel, les performances du groupe ayant eu accès aux instructions initiales ne dépassaient pas les performances de départ du groupe ayant eu accès à un matériel intégré. Ils ont également réalisé une expérience dont le but était similaire, sur le thème de la circulation sanguine et ont obtenu des résultats similaires. Les auteurs ont de nouveau observé le même profil de résultats en comparant une présentation conventionnelle à une présentation intégrée sur le thème de la programmation d'une fraiseuse. Un livret était proposé, présentant soit un schéma indépendamment du texte, soit des extraits de texte disposés directement sur le schéma et reliés aux éléments auxquels ils correspondent (Chandler & Sweller, 1992). Les auteurs ont expliqué les résultats obtenus par le fait que dans le format intégré, les apprenants avaient moins besoin de reformuler le matériel textuel pour le comprendre. De ce fait, les ressources cognitives pouvaient être allouées à l'apprentissage et à la révision du matériel imagé. A contrario, les allés retours entre le texte conventionnel et le schéma se faisaient sur une durée plus longue ce qui augmentait la charge cognitive ; on parle ici du phénomène d'attention divisée. Il est donc important de présenter les informations de manière à pouvoir associer les différents éléments clés qui les composent en un temps raisonnable.

Le phénomène d'attention divisée a été étudié par Cierniak Scheiter & Gerjet. (2009) dans le cadre d'une tâche d'apprentissage d'un phénomène biologique lié au fonctionnement des reins. Les auteurs ont comparé une présentation intégrée où le texte et les schémas étaient présentés de manière conjointe, à une présentation divisée du processus dont le texte et les images étaient spatialement éloignés. Les résultats obtenus ont révélé que les élèves ayant appris avec le format divisé avaient de moins bonnes performances d'apprentissage que ceux ayant bénéficié d'un format intégré.

Kester, Kirschner & van Merriënboer (2005) ont réalisé une expérience à l'aide d'un simulateur de circuits électriques selon deux modalités. Un groupe voyait l'information de manière fractionnée, dans ce cas le schéma était présenté à gauche de l'écran et les explications correspondantes étaient présentées à droite de l'écran. Le deuxième groupe voyait l'information

de manière intégrée, c'est-à-dire que les explications étaient inscrites directement sur le circuit. Les auteurs ont observé des performances supérieures en tâche de transfert pour la condition intégrée comparé à la condition fractionnée. Selon les chercheurs, ce type de présentation serait bénéfique pour la création de schémas généraux à condition de ne pas surcharger les ressources cognitives.

Cet aspect a également été relevé dans les travaux d'Hegarty (1992) qui a étudié la compréhension de documents techniques via un mécanisme de poulies en s'appuyant sur le mouvement oculaire (Figure 1.20).

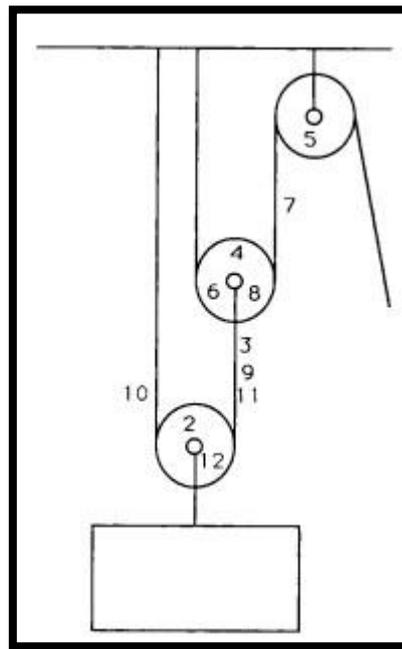


Figure 1.20. Mécanisme de poulies proposé par Hegarty (1992)

Ici, le mécanisme était présenté de manière statique par le biais d'images successives représentant les différentes étapes du mouvement. Les résultats obtenus ont révélé que le sujet analysait séquentiellement les images pour s'en faire une représentation mentale. Ensuite, l'étude d'Hegarty et Just (1993) a associé le mécanisme de poulies à un texte explicatif. Une fois encore, le mouvement des yeux des participants était enregistré. Cette expérience a pu montrer que la compréhension était meilleure lorsque le texte était présenté simultanément et conjointement à l'image. En effet, le fait de réduire le temps de transition visuel entre l'image et le texte permettait une meilleure mise en relation, ce qui conduisait à un coût attentionnel moindre et donc à une meilleure compréhension.

Cette dimension de proximité spatiale des informations a également été abordée par Mayer, Steinhoff, Bower & Mars, (1995) qui ont comparé deux groupes de participants lors d'une tâche d'apprentissage portant sur le thème de la formation de la foudre. Les auteurs ont proposé un format intégré au premier groupe qui disposait d'un texte dont les différentes étapes étaient illustrées par un schéma annoté, répétant les informations verbales relatives aux relations de cause à effet présentes dans le texte. Le même texte et les mêmes illustrations étaient présentées au deuxième groupe à la différence que les illustrations ne pouvaient être consultées qu'après la lecture du texte et qu'elles n'étaient pourvues d'aucune annotation. Ici, les résultats ont montré un bénéfice du format intégré concernant les performances en tâche de transfert. Le fait d'avoir une proximité spatiale entre les informations visuelles verbales et imagées permettait un meilleur apprentissage. En effet, le lien entre les éléments pouvait s'opérer directement, de manière à créer une représentation mentale cohérente du processus. Lorsque les informations étaient séparées, cela nécessitait d'avoir compris, intégré et mémorisé les informations clés au préalable, ce qui était cognitivement beaucoup plus coûteux. Les auteurs ont conclu qu'un bon apprentissage nécessitait la construction de liens entre les représentations verbales et visuelles d'un système.

Chung (2007) a également observé cela dans une tâche de traduction linguistique à l'aide de cartes. Les élèves devaient visualiser une carte sur laquelle un mot anglais était écrit pour comprendre sa traduction à l'aide d'une carte sur laquelle on présentait le signe chinois correspondant. Les cartes étaient présentées soit de manière adjacente soit de manière très éloignée. À l'instar des résultats déjà présentés, les performances des élèves se sont révélées supérieures dans la condition où les cartes étaient proches, comparé à la condition dans laquelle les cartes étaient très éloignées. Ces études portant sur du matériel statique, il est également important de s'intéresser à la contiguïté spatio-temporelle dans les animations.

Cela nous renvoie à l'étude de Moreno & Mayer (1999) qui ont proposé une animation sur le thème du processus de formation de la foudre. L'animation était accompagnée soit d'explications orales, soit d'explications écrites. Concernant les informations verbales écrites, elles étaient disposées soit conjointement à l'élément pertinent sur l'image soit sous forme de sous-titres. Les auteurs parlent alors d'une condition intégrée comparée à une condition séparée en rapport avec la position de l'information verbale écrite sur l'animation. Dans cette étude, les élèves qui apprenaient avec la narration obtenaient de meilleurs résultats que les deux autres groupes. Ils rappelaient plus d'informations, étaient plus performants en tâche de résolution de problèmes et parvenaient mieux à relier les éléments visuels aux informations verbales que ceux

qui apprenaient à partir d'un texte. Concernant la contiguïté spatiale du texte, il est à noter que les apprenants avaient des résultats de rappel supérieurs et étaient plus performants en tâche de résolution de problèmes lorsqu'ils étudiaient une animation dans laquelle le texte était intégré, proche des éléments clés auxquels il renvoyait, plutôt que quand ils en étaient éloignés. La narration permettait aux élèves de focaliser leur attention au bon endroit, au bon moment sur l'animation, ce qui réduisait la charge cognitive et favorisait donc l'intégration des informations. Par contre, lorsque l'information visuelle imagée était associée à de l'information visuelle verbale, conformément aux résultats obtenus par d'Hegarty & Just (1993), les élèves devaient déplacer leur attention et leur regard, allant de l'information verbale à l'information imagée. Ici, comme l'information était animée, il est possible qu'une fois le texte lu, l'élève ne dispose pas de suffisamment de temps pour se référer à l'information visuelle correspondante qui pourrait même avoir disparu à ce moment-là. Cela explique le fait que la proximité du texte avec l'image à laquelle il correspond joue un rôle dans la qualité de l'apprentissage qui en résulte. Bien évidemment, la contiguïté spatiale peut être fortement liée à la contiguïté temporelle. En effet, pour passer d'une information à une autre, cela nécessite plus ou moins de temps. La contiguïté temporelle peut s'appliquer à toutes les modalités sensorielles, c'est pourquoi il est important de comprendre en quoi elle peut représenter une aide ou au contraire un obstacle à l'apprentissage.

1.6.2 La contiguïté temporelle.

La contiguïté temporelle renvoie au fait que les individus apprendraient mieux lorsque l'information verbale (orale ou écrite) est présentée simultanément à l'information visuelle à laquelle elle correspond. Mayer & Anderson (1991) ont étudié ce sujet au moyen de deux études réalisées sur le thème du fonctionnement d'une pompe à vélo (Figure 1.21).

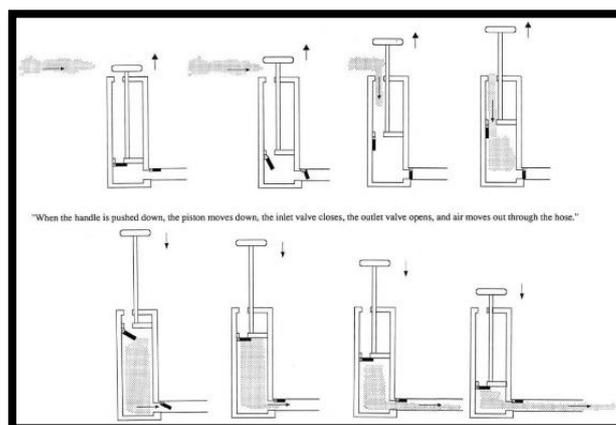


Figure 1.21. Illustrations du fonctionnement de la pompe à vélo (Mayer & Anderson, 1991).

Ce mécanisme était présenté sous forme d'une animation accompagnée d'explications orales. Un premier groupe visualisait l'animation et entendait la narration simultanément. Le deuxième groupe entendait d'abord la narration puis visualisait l'animation. Les résultats indiquent que lorsque la présentation des informations verbales et imagées était simultanée, les performances en tâche de résolution de problèmes étaient plus élevées que lorsque l'information verbale était entendue avant l'animation. Les auteurs ont répliqué cette expérience en ajoutant un test de rappel. Les résultats obtenus étaient les mêmes que dans la première étude concernant le test de résolution de problèmes et le test de rappel : la présentation simultanée des informations offrait de meilleures performances que la présentation successive. Selon les auteurs, cela indique qu'un double codage sollicitant simultanément les deux canaux sensoriels (visuel et auditifs) permet une intégration plus efficace des deux systèmes de représentation.

En 1992 Mayer & Anderson ont de nouveau testé le principe de contiguïté temporelle à travers deux expériences, la première portait sur le mécanisme de pompe à vélo et la deuxième sur le système de freinage de voiture. Ces deux expériences étaient similaires à l'exception du mécanisme étudié, le but étant de confirmer la validité des résultats de la première étude. Les auteurs ont présenté une animation et une narration relative au système à étudier dans le but de comparer les performances en tâche de rétention et en tâche de résolution de problèmes, selon que la narration était présentée simultanément ou successivement à l'animation. Pour un premier groupe, la narration était présentée simultanément à l'animation et était diffusée trois fois. Pour d'autres groupes, la narration était présentée soit avant, soit après l'animation et le tout était diffusé trois fois. Pour deux autres groupes, la narration était diffusée trois fois avant ou après l'animation qui était elle aussi diffusée trois fois. Deux autres groupes ont également visualisé soit l'animation seule, soit la narration seule, à trois reprises également. Pour finir, un groupe contrôle ne recevait aucune instruction. Les auteurs ont observé que pour le test de rétention, le groupe contrôle obtenait de moins bons scores que les autres groupes. Dans l'expérience 1 les autres groupes ne différaient pas entre eux, à l'exception du groupe dont l'animation était diffusée avant l'animation, qui obtenait des scores supérieurs au groupe visualisant uniquement l'animation. Dans l'expérience 2, tous les groupes obtenaient de meilleures performances que le groupe contrôle à l'exception du groupe qui ne visualisait que l'animation. Concernant la résolution de problème, les auteurs ont observé que la diffusion simultanée de l'animation et de la narration entraînait des résultats supérieurs aux autres groupes qui soit ne disposaient d'aucune instruction soit visualisaient une diffusion successive.

Ces résultats sont en accord avec le phénomène de contiguïté temporelle qui implique que les éléments verbaux et visuels soient présentés simultanément.

Le même type d'expérience a été réalisé par Mayer & Sims (1994) à la différence que le niveau d'habiletés spatiales était pris en compte. A travers deux expériences, les auteurs ont comparé le niveau d'habiletés spatiales des élèves au caractère simultané ou successif de la présentation des animations et de la narration. La première expérience portait sur le thème du fonctionnement d'une pompe à vélo. Les auteurs ont comparé quatre groupes expérimentaux comprenant un groupe contrôle, un groupe visualisant l'animation et la narration simultanément, un groupe visualisant la narration avant l'animation et un groupe visualisant la narration après l'animation. Les résultats obtenus ont montré qu'en tâche de transfert, les performances étaient supérieures pour les apprenants ayant étudié la présentation simultanée comparé aux autres groupes. Par contre, il n'y avait pas de différence de performances entre la présentation successive et la présentation simultanée pour les apprenants ayant de moins bonnes habiletés spatiales. La présentation simultanée semblait donc profiter aux élèves ayant de bonnes habiletés spatiales uniquement. Les chercheurs ont répliqué cette expérience, mais cette fois ils ont abordé le thème du fonctionnement du système respiratoire. Ils ont établi un groupe contrôle, un groupe simultané et un groupe successif qui voyait la narration avant l'image. Ici les performances en tâche de transfert étaient meilleures pour la présentation simultanée comparativement à la présentation successive dont les résultats étaient supérieurs à ceux du groupe contrôle. Aussi, les élèves ayant de faibles habiletés spatiales obtenaient les mêmes résultats, qu'ils aient étudié à l'aide du mode simultané ou successif. Si les résultats des apprenants ayant de bonnes habiletés spatiales semblaient plus élevés que ceux des apprenants ayant de faibles habiletés visuospatiales, ceux-ci n'étaient pas significatifs dans cette deuxième expérience. Concernant les deux expériences, les auteurs ont conclu que les élèves ayant de bonnes habiletés spatiales étaient plus à même de créer des relations entre les informations visuelles et verbales, ce qui expliquerait le fait qu'ils aient bénéficié plus fortement de la présentation simultanée.

Mayer, Moreno, Boire & Vagge (1999) ont fait varier le type de présentation au niveau de la successivité des informations narratives et animées dans deux expériences. La première portait sur la formation de la foudre et faisait appel à 5 groupes expérimentaux, faisant varier la longueur des extraits narratifs et animés. Le premier groupe entendait la narration simultanément à l'animation. Deux autres groupes visualisaient soit un grand morceau d'animation suivi de la narration (AN), soit un grand morceau d'animation diffusé après la

narration (NA). Pour finir, deux autres groupes étudiaient de petits morceaux d'animation alternés avec la narration (ANAN), soit de petits morceaux de narration alternés avec l'animation (NANA). Ici, que ce soit pour les tests de rétention, de transfert ou de matching, les performances étaient équivalentes entre AN et NA qui obtenaient des résultats supérieurs à ANAN et NANA, qui étaient également équivalents entre eux. La deuxième expérience était la même à la différence que le thème abordé était celui du fonctionnement d'un système de freinage de voiture. Les résultats obtenus aux tests de rétention et de transfert étaient les mêmes que dans la première expérience. Par contre concernant le test de matching, les auteurs ont observé que la version simultanée montrait des résultats supérieurs aux autres conditions qui elles, étaient toutes équivalentes entre elles. Les auteurs ont montré que les représentations multimédia étaient clairement plus efficaces quand la présentation des informations visuelles et narratives était simultanée. Il est donc important que ces informations soient traitées en même temps en mémoire de travail de manière à être intégrées correctement. Cependant, lorsqu'il y a un décalage temporel entre le son et l'image, on peut se demander à quel degré les performances seraient affectées selon temps de latence entre les deux informations.

C'est ce qu'a étudié Baggett (1984) en présentant une vidéo de 30 minutes sur le thème de Legos de construction techniques dans laquelle le temps de présentation de la narration était manipulé. La narration était soit synchronisée avec le film, soit décalée de 7, 14 ou 21 secondes avant ou après la présentation visuelle. Pour la moitié des participants était proposé un rappel immédiat, l'autre moitié devait réaliser la tâche de rappel 7 jours après l'expérimentation. Concernant le rappel immédiat, les résultats ont révélé qu'il n'y avait pas de différence significative entre la condition synchronisée et la condition où la narration était présentée après l'information visuelle. Les groupes de ces deux modalités ont obtenu des résultats significativement supérieurs aux autres groupes qui eux ne différaient pas entre eux. Une fois le rappel réalisé, 7 jours plus tard, le même profil de performances était observé pour les conditions synchronisées et +7, qui montraient des performances supérieures aux autres groupes. Cependant, parmi les autres groupes, les résultats obtenus lorsque la narration était présentée 14 ou 17 secondes après l'image montraient de meilleures performances que les trois groupes dont la narration était présentée avant l'information visuelle. L'auteur a montré une asymétrie du décalage temporel et a conclu que quand l'information visuelle précédait l'information verbale jusqu'à 7 secondes, le rappel était d'aussi bonne qualité que lorsque les deux types d'informations étaient synchronisés. A contrario, quand la narration précédait

l'information visuelle de 7 secondes ou plus, alors il y avait une perte considérable d'informations, d'autant plus après un délai d'une semaine.

Si un décalage de 7 secondes semblait raisonnable, Xie, Mayer, Wang & Zhou, (2018) montrent que l'apprentissage peut être perturbé dans un délai beaucoup plus court. Dans cette étude portant sur le thème de la transmission neurologique, les auteurs ont travaillé à partir de brèves animations desquelles ils ont fait varier la présentation d'indices visuels et auditifs. Ils ont montré qu'un décalage de 3 secondes était préjudiciable à l'apprentissage.

A l'inverse, Meyerhoff & Huff (2016) qui ont proposé des extraits de films en faisant varier le temps de latence relatif à la présentation d'une piste audio de 3, 5 ou 4 secondes n'ont pas observé d'effet. Conjointement aux temps de latences proposés entre le son et l'image, le type de matériel proposé aux apprenants peut jouer un rôle important dans la compréhension qu'ils peuvent en avoir. C'est pourquoi, nous nous sommes inspirés de la méthodologie de Baggett (1984) qui se rapprochait le plus de ce que nous souhaitons proposer aux élèves dans le but de tester la contiguïté temporelle en réalité virtuelle non immersive. Nous avons pensé que, dans une leçon complète en réalité virtuelle non immersive, il serait plus difficile d'intégrer la narration et l'information animée correspondante lorsque celles-ci ne sont pas présentées de manière contiguës, du fait de la complexité des images et du processus présenté. Les résultats que nous avons obtenus et que nous présenterons dans un chapitre dédié nous ont conduit à penser que dans le cadre d'un apprentissage en réalité virtuelle immersive, il était tout à fait possible que les informations pertinentes soient difficiles à extraire et à mettre en relation les unes avec les autres. Nous avons donc pensé que l'élaboration d'un modèle mental cohérent pourrait être facilité par l'ajout d'un feedback explicatif, permettant à l'élève de se remémorer et de synthétiser de manière plus efficace l'environnement exploré. Il est donc important que nous abordions la notion de feedback plus précisément.

1.7 Le feedback.

Si beaucoup d'études intègrent le feedback dans leur méthodologie, peu d'entre elles ont étudié ses effets en tant que tels dans les apprentissages. Concernant la réalité virtuelle, qu'elle soit immersive ou non, aucune étude n'a encore présenté de résultats sur le bénéfice potentiel que pourrait présenter la diffusion d'un feedback sur l'apprentissage.

Le feedback consiste à donner un retour informatif à l'apprenant après la réalisation d'une tâche. Il peut se manifester sous des formes multiples comme un son, ou une couleur rouge ou

verte, pour indiquer si la réponse donnée par l'apprenant est correcte ou non. Il peut également être présenté sous forme d'explications pour que l'apprenant puisse comprendre et corriger son erreur. Bien évidemment, cette définition est réductrice, il convient donc de développer ce principe en abordant les différentes études réalisées à ce sujet. Narciss & Huth (2002) évoquent deux composantes du feedback : une composante évaluative dite de vérification (Kulhavy & Stock, 1989) et une composante informationnelle. La composante évaluative donne le niveau de performance de l'apprenant et la composante informationnelle donne des informations supplémentaires sur la tâche réalisée, notamment quand l'apprenant fait des erreurs.

Mayer & Johnson (2010) ont par exemple étudié le rôle de deux types de feedback sur l'apprentissage. Les chercheurs ont proposé une sorte de jeu d'arcade composé de 10 niveaux à leurs étudiants, leur exposant des circuits électriques. Les circuits électriques proposés étaient composés d'une diversité de batteries et de résistances connectées soit en série, soit en parallèle. Les apprenants devaient indiquer lequel des circuits proposés montrait la circulation électrique la plus rapide. Deux groupes expérimentaux étaient établis : certains apprenants recevaient un feedback sonore dont la tonalité aigüe ou grave indiquait s'ils avaient choisi la bonne ou la mauvaise réponse. Le deuxième groupe recevait un feedback explicatif en plus du feedback auditif. Les résultats ont révélé que les élèves ayant reçu un feedback explicatif montraient des performances plus élevées pendant l'apprentissage et en tâche de rappel que les élèves qui n'avaient reçu qu'un feedback sonore. Le feedback explicatif permettrait donc une compréhension plus complète du dispositif étudié.

Un feedback est également proposé dans l'expérience de Moreno & Mayer (2005). Les auteurs ont fait varier la production de réflexion et le guidage. Les chercheurs ont proposé un jeu multimédia interactif sur un thème botanique dans lequel tous les élèves recevaient un feedback lorsqu'ils donnaient une bonne réponse. Les chercheurs ont proposé ou non aux élèves d'expliquer leur choix de réponse. Aussi, les élèves recevaient ou non une explication après avoir été informés que leur réponse était correcte. Les résultats ont montré que les élèves qui recevaient un feedback explicatif obtenaient des scores plus élevés en tâche de transfert, donnaient un plus grand nombre d'explications correctes dans leur choix de réponse et arrivaient mieux à corriger leurs réponses lorsqu'ils se trompaient que les autres élèves. Par contre, demander aux sujets d'expliquer leurs réponses ne semblait pas permettre un apprentissage plus efficace.

Dans l'étude de Johnson, Perlow & Pieper (1993), deux types de feedback ont été comparés dans une tâche de simulation de télémanipulation d'un bras robotique de navette spatiale. Les

auteurs ont constitué deux groupes expérimentaux : le premier groupe bénéficiait d'un feedback axé sur la performance ; dans ce cas, on donnait aux participants le résultat de leur action sans explication particulière. Le deuxième groupe bénéficiait d'un feedback axé sur l'apprentissage qui leur indiquait leurs résultats et leur fournissait des commentaires ainsi que des informations pédagogiques sur la façon d'éviter de reproduire les erreurs commises dans le futur. Dans l'ensemble, les résultats indiquaient que quel que soit le type de feedback, les performances dans la réalisation de la tâche étaient similaires pour les deux groupes. Cependant, le groupe ayant bénéficié d'explications faisait significativement moins d'erreurs que le groupe n'en ayant pas bénéficié.

Tous les facteurs dont nous venons de faire l'analyse et que nous allons évoquer dans nos travaux font appel à une quantité plus ou moins importante de ressources cognitives. C'est pourquoi, pour clore ce chapitre, nous allons aborder la théorie de la charge cognitive.

1.8 Théorie de la charge cognitive

La théorie de la charge cognitive (Sweller 1988,2010, 2011, 2012 ; Chanquoy, Tricot & Sweller, 1992 ; Sweller et al. 1998) relate les mécanismes cognitifs en jeu dans les apprentissages et la résolution de problèmes ainsi que la manière avec laquelle ils s'articulent. Le but est de comprendre ce qui augmente la charge cognitive et de ce fait perturbe l'apprentissage, pour ensuite tenter d'optimiser la présentation de l'information et subséquemment les performances des apprenants. Le modèle de Sweller propose trois sources de charge cognitive. Ce modèle distingue la charge intrinsèque, la charge extrinsèque et enfin la charge dite « germane ». La charge intrinsèque repose sur la complexité de la tâche à réaliser et ses exigences, par exemple : plus la tâche comporte d'interactions entre des éléments complémentaires (qui peuvent donc être traités simultanément pour que l'apprenant puisse réaliser la tâche correctement), plus la charge intrinsèque augmente. Le fait de devoir manipuler une multiplicité d'éléments simultanément en mémoire de travail peut alors provoquer une surcharge cognitive, rendant ainsi la tâche plus difficile. La charge intrinsèque peut être amoindrie soit en changeant la nature de la tâche soit lorsque l'apprenant a acquis plus de connaissances, Paas & Sweller (2012). Ainsi, les connaissances préalables des apprenants sont importantes à prendre en compte dans la conception du matériel. Vient ensuite la charge extrinsèque. Elle est liée à la présentation des informations et à leur organisation dans le document. Ici il ne s'agit pas de comprendre le lien étroit entre différents éléments pour réussir la tâche mais de comprendre comment les éléments sont présentés et organisés au sein même

du dispositif ou du document multimédia. La difficulté vient alors du fait que l'apprenant doit résoudre un problème de dispositif inapproprié avant de pouvoir traiter l'information nécessaire à la réalisation correcte de la tâche. Pour finir, il y a la charge « germane » ou essentielle, qui est « *la charge consacrée à l'acquisition et à l'automatisation des schémas ou, en d'autres termes, à l'apprentissage* » Debie & van de Leemput (2014) ainsi qu'aux mécanismes de traitement : de sélection, d'organisation et d'intégration des informations (Mayer, 2014). Elle fait référence aux ressources de la mémoire de travail qui traite la charge intrinsèque plutôt que la charge extrinsèque (Paas & Sweller, 2012). Le coût de la charge extrinsèque va influencer celui de la charge intrinsèque. Si on allège le coût de la charge extrinsèque, alors l'apprenant aura plus de ressources en mémoire de travail pour traiter l'information requise en charge intrinsèque. De cette manière, il sera centré davantage sur les informations liées directement à la tâche et pourra plus facilement créer et automatiser des schémas corrects.

Bien évidemment, il convient d'apporter quelques nuances à ces trois types de charges cognitives. Comme nous l'avons souligné, il est indispensable de considérer la place des connaissances préalables des individus lors de l'apprentissage. S'il semble évident qu'un niveau de connaissances préalables élevé allège la charge cognitive intrinsèque et permette un meilleur apprentissage, il est important de noter que cela peut également présenter un désavantage dans certains cas de Figure. Par exemple, Kalyuga, Chandler & Sweller (2000) ont montré que cet effet pouvait s'inverser grâce à deux expériences réalisées auprès de participants novices. Dans leur première expérience, les auteurs ont présenté à des participants novices un schéma selon 4 conditions expérimentales. Les expérimentateurs présentaient un schéma accompagné soit d'un texte (voir Figure 1.22), soit d'une explication orale, soit d'un texte écrit et d'une explication orale ou le schéma seul. La tâche des élèves était de déterminer le nombre correct de tours par minute pour percer une pièce de bronze à l'aide des explications fournies.

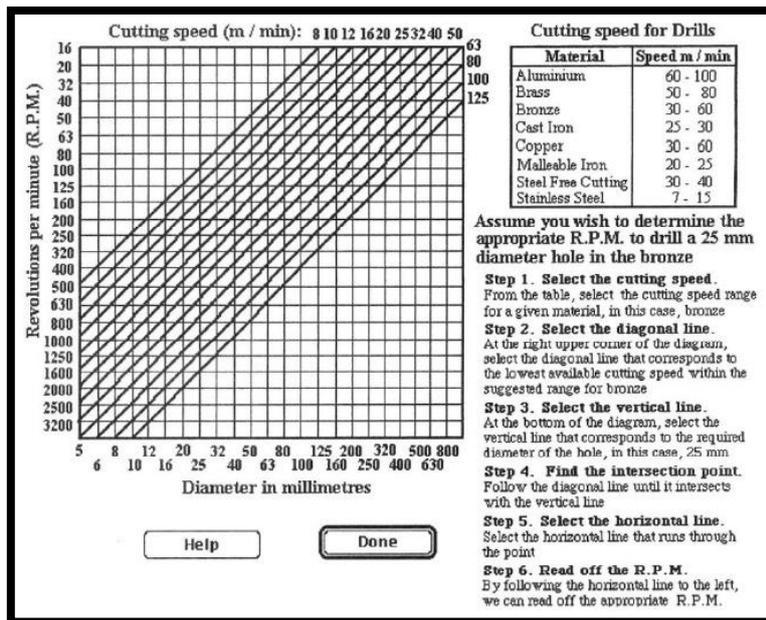


Figure 1.22. Exemple du matériel proposé par Kalyuga & al. (2000).

Dans ce cas, ils obtiennent de meilleures performances lorsque l'information verbale est présentée oralement par rapport à l'image seule. Cependant, les auteurs n'ont pas relevé de bénéfices d'apprentissage lorsque le texte était présenté à la fois visuellement et oralement, et encore moins lorsque le schéma était présenté seul. Par contre, après deux sessions d'entraînement, les résultats ne montraient plus d'effet bénéfique de la présentation orale sur la tâche. La deuxième session avait lieu auprès des mêmes élèves deux mois et demi plus tard, durée pendant laquelle ils ont pu pratiquer en formation la tâche qu'ils devaient réaliser lors de l'expérimentation. Les auteurs présentaient alors deux conditions après une session d'entraînement de 25 minutes : présentation d'un schéma seul ou présentation d'un schéma accompagné de l'explication orale. Cette fois ci, les résultats ont révélé des performances plus élevées lorsque le schéma était présenté seul plutôt qu'avec des explications orales. Les auteurs ont donc conclu que le type d'instructions donné était bénéfique uniquement dans certaines circonstances bien définies selon le niveau d'expertise des apprenants. Ce qui était bénéfique pour des apprenants novices pouvait à l'inverse être un désavantage pour les experts puisque leurs besoins n'étaient pas les mêmes dans la réalisation de la tâche. En réalité, imposer des informations cohérentes mais dont on n'a pas besoin interférerait avec les schémas mentaux stockés en mémoire à long terme et créeraient donc une charge cognitive plus importante. Schnotz & Rasch (2005) ont également observé de tels effets. Ils ont réalisé deux

expériences dont le support d'apprentissage portait sur les différents fuseaux horaires ainsi que sur les déplacements autour de la terre (voir Figure 1.23).

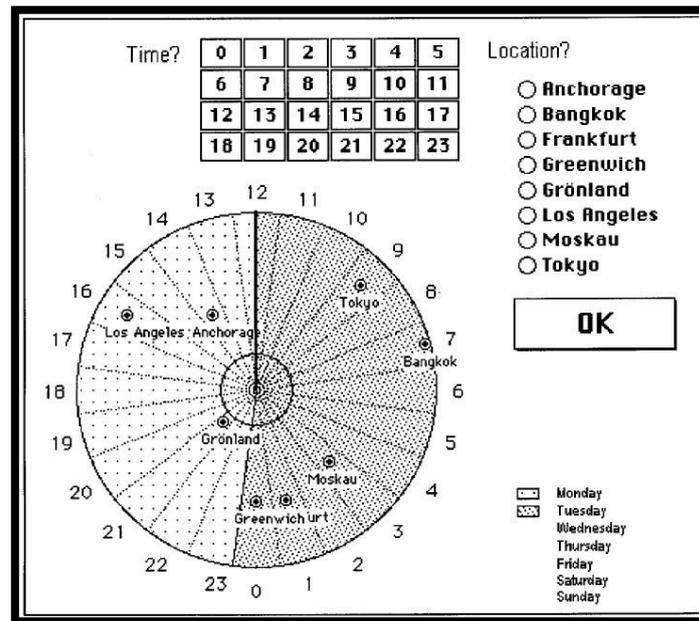


Figure 1.23. Exemple de matériel utilisé par Schnotz & Rasch (2005).

Dans une première expérience ils ont comparé l'apprentissage à l'aide d'une version statique et d'une version animée de ce matériel. Les auteurs ont montré que les animations avaient un effet d'habilitation pour les apprenants ayant de bonnes connaissances préalables et un effet de facilitation pour les apprenants ayant peu de connaissances préalables. Les chercheurs ont répliqué cette expérience en comparant cette fois-ci un apprentissage par manipulation ou par simulation du processus. Les résultats obtenus ont fait émerger le fait que les différents types d'animation avaient des fonctions différentes pour l'apprentissage. Le fait de manipuler le dispositif aurait un effet d'habilitation alors que la simulation aurait un effet plutôt facilitateur. Dans les deux cas, une diminution de la charge cognitive était relevée. Cependant, ces fonctions ne représentaient un atout que pour un certain type d'apprenants : la fonction facilitante serait inappropriée pour les apprenants sachant déjà simuler mentalement le processus puisque cela les empêcheraient de réaliser un traitement cognitif pertinent. Cet effet facilitateur, au lieu de diminuer la charge extrinsèque diminuerait dans ce cas la charge germane, indispensable à l'apprentissage. Comme la représentation mentale était déjà acquise, une sorte de redondance était alors créée, amenant l'apprenant à traiter des informations supplémentaires inutiles, ce qui était considéré comme une augmentation de la charge extrinsèque (Kalyuga & al, 1998 ; Kalyuga & Sweller, 2005).

Schnotz & Rasch (2005) ont souligné le fait que la charge cognitive était bien plus complexe à évaluer qu'il n'y paraissait et qu'il ne suffisait pas d'utiliser cette théorie en tant que cadre de recherche mais plutôt de la mesurer à l'aide d'échelles objectives afin de mieux définir ses différents aspects. Cependant cette théorie, lorsqu'on ne l'étudie pas spécifiquement reste indispensable lors de la conception de dispositifs d'apprentissage. C'est d'ailleurs de cette manière que certains principes ont pu émerger au fil du temps et des recherches. Les cinq principes centraux que nous avons décrits précédemment avaient pour but de réduire la charge extrinsèque lors de l'apprentissage multimédia.

Dans le cadre de nos travaux, nous allons présenter des dispositifs d'apprentissage en réalité virtuelle, immersive et non immersive, dont le contenu est riche et dynamique. Pour mieux comprendre comment nous apprenons à l'aide d'animations, le chapitre suivant portera spécifiquement sur ce type d'apprentissages.

CHAPITRE 2 - LES APPRENTISSAGES A PARTIR DE VISUALISATIONS DYNAMIQUES.

Les dispositifs de réalité virtuelle permettent de présenter un nombre conséquent d'informations de sources multiples, plus particulièrement des informations visuelles animées. Dans le cas de la réalité virtuelle non immersive, les animations proviennent du système. Par contre, en réalité virtuelle immersive (RVI), les visualisations proviennent également de l'apprenant, qui par ses déplacements ainsi que par les mouvements des yeux et de la tête, contrôle le défilement des images. Afin de mieux comprendre les mécanismes d'apprentissage en réalité virtuelle, il est nécessaire de développer plus précisément les mécanismes cognitifs impliqués dans l'apprentissage à partir d'animations. L'objectif de ce chapitre est donc de développer les caractéristiques relatives à l'apprentissage à partir de dispositifs multimédias dynamiques. Pour commencer, nous allons définir les caractéristiques relatives aux animations, puis nous détaillerons plus précisément l'importance de la perception dynamique dans la compréhension de documents multimédias. Nous développerons ensuite les étapes du modèle de traitement des animations (APM) et pour finir, nous aborderons 4 méta-analyses relatives aux apprentissages multimédias à partir d'animations.

2.1 Définition et caractéristiques des animations.

Comparativement aux présentations statiques conventionnelles, les animations mettent en jeu différentes dimensions qui sont traitées de manière différente par l'apprenant. Pour commencer, il est important de définir ce qu'est une animation. Selon Lowe (2017), ce qui distingue l'animation de la présentation d'images statiques, ce sont les changements opérés dans le temps. Il s'agit de « *toute application qui génère une série d'images, de manière que chaque image apparaisse comme une altération de l'image précédente, et où la succession des images est déterminée soit par le concepteur, soit par l'utilisateur* » Bétrancourt (2005). Elle ajoute que « *dans la plupart des cas, l'animation implique un mélange d'éléments représentationnels qui ressemblent à l'objet réel, des signes et des symboles communs ou spécifiques à un domaine (comme par exemple des flèches) et surtout, des informations verbales* » Bétrancourt (2005). En effet, les animations peuvent intégrer bon nombre d'informations qui de surcroît sont en mouvement et peuvent apparaître, disparaître ou encore se transformer. Cela renvoie à la notion de « transience » c'est-à-dire au caractère fugace, transitoire des présentations animées. Lors d'une présentation conventionnelle statique incluant un texte accompagné d'images, l'apprenant peut explorer le matériel à son rythme et dans l'ordre qu'il souhaite. A l'inverse, lorsque la présentation est animée, l'apprenant doit rapidement détecter l'information pertinente et inhiber les potentielles distractions pour finalement pouvoir créer un modèle mental cohérent. Il doit alors stocker, traiter et relier les

informations présentes et antérieures de manière simultanée, ce qui peut solliciter une charge cognitive importante (Boucheix & Forestier, 2017) tout dépendant des caractéristiques individuelles des apprenants, comme entre autre, leurs connaissances préalables. Sachant cela, on peut se demander quel type de présentation, statique ou animée, peut susciter un meilleur apprentissage et dans quelles conditions ? Avant de détailler les mécanismes d'apprentissages à partir de documents multimédias animés, il est important de comprendre comment nous percevons et traitons l'information animée.

2.2 Le rôle des habiletés de raisonnement spatial dynamique.

Bien que peu nombreux, quelques travaux se sont intéressés à l'étude du traitement des caractéristiques spécifiques aux animations en relation avec la compréhension de dispositifs éducatifs animés. Ils se basent principalement sur la discrimination de vitesse entre plusieurs éléments simples en mouvement. La première étude effectuée à ce sujet est celle de Gibson (1947) sur le thème de l'aviation. Son étude regroupait de nombreuses épreuves, dont trois tâches portant sur la perception de la vitesse. La première portait sur l'estimation de la vitesse, la seconde épreuve était une tâche de vitesse relative et la troisième portait sur le classement d'éléments présentés les uns indépendamment des autres selon leur vitesse. Trois facteurs ont été mis en évidence ici, à savoir la perception de la vitesse, la détection du mouvement ainsi que le facteur d'attention divisée. Aucun facteur n'a été mentionné à propos des différences interindividuelles, cependant les résultats obtenus ont montré que l'estimation de la vitesse était corrélée avec le raisonnement abstrait et le rappel de position, selon l'article de Fischer & al. (1994). Ces chercheurs ont d'ailleurs repris ces tâches en y intégrant d'autres dimensions. Ils mettaient en compétition un « 1 » et un « 0 » qui se déplaçaient en direction d'une cible, sachant que les stimuli disparaissaient avant de l'atteindre. Les auteurs ont fait varier les trois facteurs suivants : la différence de temps d'arrivée entre les deux objets, le rapport de vitesse qu'ils entretenaient ainsi que la configuration de la trajectoire des stimuli. Trois trajectoires différentes étaient déterminées : les stimuli étaient alors soit parallèles, perpendiculaires ou obliques. Les participants devaient dire lequel des deux stimuli atteindrait la cible en premier. Pour tenter de comprendre le raisonnement des participants, les auteurs ont divisé leur expérience en quatre blocs, sachant que pour trois d'entre eux, les participants avaient pour consigne de penser tout haut et pour l'autre ils avaient pour consigne de ne formuler aucune réponse orale. Leur but était ici de voir si les participants pouvaient exprimer explicitement leur conscience de la vitesse et/ou de la notion de distance et de déterminer quelles stratégies ils avaient employées. Les résultats obtenus ont montré qu'il existait une relation entre les

performances dans les tâches dynamiques spatiales et l'usage de stratégies. Les auteurs en ont déduit que les participants ayant de faibles scores globaux n'avaient pas précisément intégré les caractéristiques de vitesse et de distance dans leur processus de décision.

Hunt & al. (1988) ont également étudié la capacité de raisonner sur le mouvement dans la modalité visuelle. Ils ont fait varier les épreuves statiques et dynamiques en postulant que le raisonnement visuel sur les modalités dynamiques était différent du raisonnement visuel sur les modalités statiques. Les participants de leur expérimentation ont effectué 6 épreuves écrites d'habiletés visuospatiales, 5 tâches statiques sur ordinateur, et 6 tâches dynamiques elles aussi sur ordinateur. Concernant les habiletés spatiales dynamiques, les auteurs pensaient que le participant devrait coordonner son jugement temporel avec la direction et le mouvement moteur. Selon eux, faire un jugement coordonné pouvait être une source significative de différences individuelles. Leurs tâches dynamiques tendaient donc à évaluer la mémoire de trajectoire, le jugement de rapidité du mouvement, le jugement de vitesse relative du mouvement, la capacité à traiter un mouvement relatif dans différentes conditions ainsi que la capacité à combiner l'extrapolation de la vitesse et de la trajectoire. Concernant la mémoire de trajectoire, le participant voyait un petit carré suivre une trajectoire parabolique à l'écran à trois reprises, sachant qu'une trajectoire était différente des deux autres. Il devait alors indiquer laquelle des trois trajectoires différait des deux autres. Pour le jugement de vitesse, la tâche proposée ressemblait à celle de Gibson (1947). Un carré se déplaçait linéairement jusqu'à une ligne, disparaissant avant de la toucher. Le sujet devait appuyer sur un bouton lorsqu'il pensait que le carré toucherait la ligne en supposant que le mouvement du stimulus continuait. La tâche de jugement de la vitesse relative du mouvement mettait deux objets en compétition. Chaque objet prenait une vitesse constante différente et une trajectoire soit perpendiculaire, soit parallèle soit prenant des angles différents. Le sujet devait dire lequel des deux stimuli arriverait en premier à destination. L'avant dernière tâche mesurait l'habileté à traiter un mouvement relatif. Ici le sujet voyait 4 « digits » prenant leur départ à une distance différente de la ligne d'arrivée. La tâche était de trouver lequel de ces quatre digits toucherait la ligne en premier. Pour finir, les auteurs ont mesuré la capacité des sujets à combiner l'extrapolation de la vitesse et de la trajectoire avec la tâche « intercept ». Ici, un petit triangle partait de la gauche de l'écran en suivant la trajectoire soit d'une ligne horizontale, soit d'une onde sinusoïdale, soit d'une parabole. En bas à droite de l'écran, un missile attendait d'être lancé par le sujet. La tâche du participant était donc de lancer le missile de manière à ce que celui-ci entre en collision avec la cible. Les chercheurs ont comparé ces tests dynamiques aux autres tests statiques et en ont déduit que les habiletés requises pour le traitement d'objets statiques étaient bien différentes de

celles requises par le traitement d'objets dynamiques. Ceci dit selon les auteurs, bien que les épreuves soient jugées comme fiables, elles ne permettaient pas de définir de facteurs précis, ils ont donc évoqué la nécessité de mener d'autres études à ce sujet.

Cela nous renvoie à l'étude de Sanchez et Wiley (2014), dont l'objectif était d'évaluer si les performances à une tâche impliquant un traitement d'éléments dynamiques étaient liées à la compréhension que les apprenants pouvaient avoir d'un texte de géoscience sur les plaques tectoniques et l'activité d'un volcan. Les chercheurs ont fait varier le type de présentation pour évaluer la compréhension des sujets. Trois types de présentations étaient possibles : texte seul, texte accompagné d'un schéma approprié et texte accompagné d'une version animée de l'illustration. Pour la tâche « intercept », on retrouvait la même configuration que celle proposée par Hunt & al. (1988) sauf qu'ici la cible gardait une trajectoire linéaire (Figure 2.1). Le missile gardait une vitesse constante alors que la cible elle, prenait trois vitesses différentes. Ici aussi, le but était de lancer le missile à l'aide de la barre espace du clavier d'ordinateur de manière à ce qu'il entre en collision avec la cible. Afin de rendre la tâche un peu plus complexe, les auteurs ont fait varier le temps d'apparition de la cible.

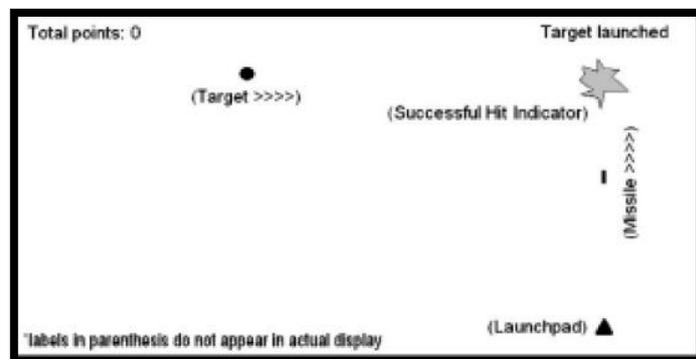


Figure 2.1. Tâche intercept de Sanchez & Wiley (2014)

En comparant leurs différentes tâches, les auteurs ont remarqué que les participants ayant de bonnes performances dans la tâche intercept comprenaient mieux le texte, sauf lorsque celui-ci n'était accompagné d'aucune illustration. Les chercheurs ont également montré que l'illustration était en interaction avec la tâche intercept d'autant plus lorsque celle-ci n'était pas animée. Ils ont alors conclu que la pensée spatiale était critique pour les activités cognitives en science.

La perception d'éléments dynamiques en relation avec la compréhension d'animations a ensuite été étudiée par Porte, Boucheix & al. (2016) à travers 4 tâches. Les deux premières étaient des tâches de discrimination de vitesse dans lesquelles on présentait deux stimuli ayant des

vitesses relatives différentes (Figure 2.2). La première tâche était nommée « course parallèle » (CPA) et comportait des couples de stimuli présentés selon des vitesses globales différentes allant de 100 pixels par seconde à 650 pixels par seconde. La vitesse relative entre les deux stimuli variait de 0.5px/s à 3 px/s.

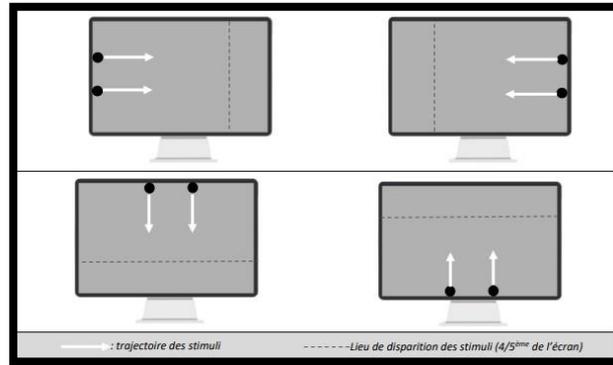


Figure 2.2. Tâche de course parallèle.

La tâche du participant était de cliquer sur le stimulus le plus rapide des deux et ce, le plus rapidement possible. La consigne était la même pour la seconde tâche nommée « course à angles différents » (CAN) (Figure 2.3). Comme son nom l'indique, les stimuli prenaient cette fois-ci des trajectoires différentes de 90° ou de 180° mais gardaient une vitesse globale constante. Seule leur vitesse relative variait de 0.5px/s à 7px/s.

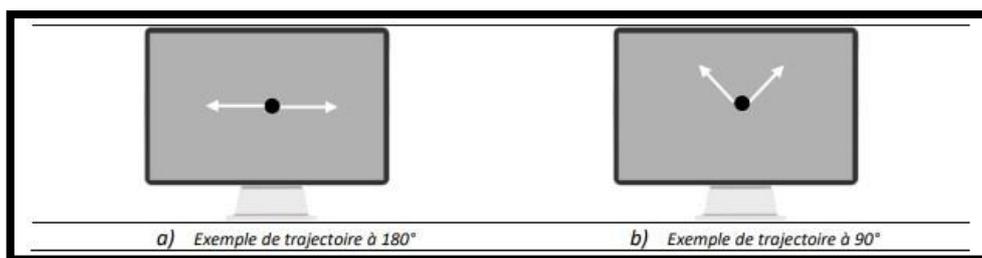


Figure 2.3. Exemples de stimuli de la tâche CAN.

Les deux dernières tâches reprenaient celle de Sanchez & Wiley 2014 (voir Figures 2.1 et 2.4). Il y avait une tâche autocontrôle dans laquelle les stimuli amorçaient leur trajectoire pour disparaître avant de se toucher. La tâche des participants était alors d'indiquer si oui ou non, les deux stimuli allaient entrer en collision.

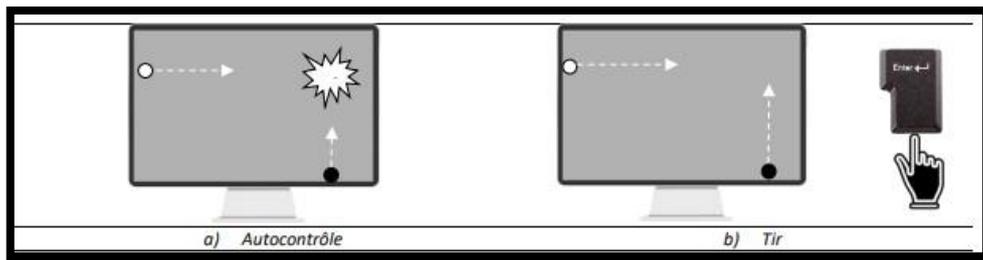


Figure 2.4. Tâches autocontrôle et tir, Porte, Boucheix & al. (2016).

Concernant la tâche de tir, l'apprenant devait lancer le missile (en noir sur la Figure 2.4) de manière à ce qu'il entre en collision avec la cible (en blanc sur la Figure 2.4). Les performances à ces tâches étaient comparées aux performances à des tâches de compréhension d'animation à savoir : la compréhension du mécanisme de piano et le mouvement de locomotion du kangourou. Les habiletés visuospatiales ont également été mesurées à l'aide du test DAT. Les résultats obtenus ont montré que les performances aux tâches dynamiques n'étaient pas corrélées à celles du DAT, ce qui indiquait que les processus engagés pour traiter l'information statique et dynamique étaient différents. Les auteurs ont également noté une corrélation significative (tendancielle) entre les performances de compréhension d'animation et les tâches d'interception (autocontrôle et tir) ainsi qu'entre le DAT et la compréhension d'animation. Ceci suggère que deux composantes seraient impliquées dans le traitement de visualisations dynamiques : une composante de mouvement, et une composante de relations spatiales.

Ces études expliqueraient donc en partie, pourquoi les performances seraient différentes selon que la présentation proposée est animée ou statique. Traiter les composantes du mouvement nécessite d'anticiper, de traiter visuellement des vitesses, des trajectoires ou encore des transformations. Ceci n'est pas le cas des présentations statiques qui elles, peuvent nécessiter la représentation interne d'un mouvement, mais ne demandent pas son traitement en tant que tel. On peut donc se demander quelles seraient les caractéristiques à appliquer préférentiellement dans les dispositifs d'apprentissage à partir d'animations, pour cela, il est important de comprendre comment les apprenants traitent et intègrent l'information animée dans le cadre d'un apprentissage à partir d'animations.

2.3 Le traitement de l'animation : le Modèle APM, Lowe & Boucheix (2008).

Les étapes de traitement de l'information visuelle animée diffèrent de celles de l'information statique. Il est donc important d'évoquer le modèle APM (Animation Processing

Model). Lowe et Boucheix (2008) ont élaboré le premier modèle de traitement de l'animation, composé de 5 étapes successives (Figure 2.5). La première étape renvoie à un processus bottom-up (ascendant) qui consiste à explorer l'information dynamique afin de la segmenter en unités individuelles. Etant donné le caractère continu de la diffusion des informations dynamiques, l'apprenant doit alors observer l'animation pour pouvoir anticiper le flux d'informations. Ainsi, il lui est possible de segmenter les différents événements pour en dégager des unités individuelles.

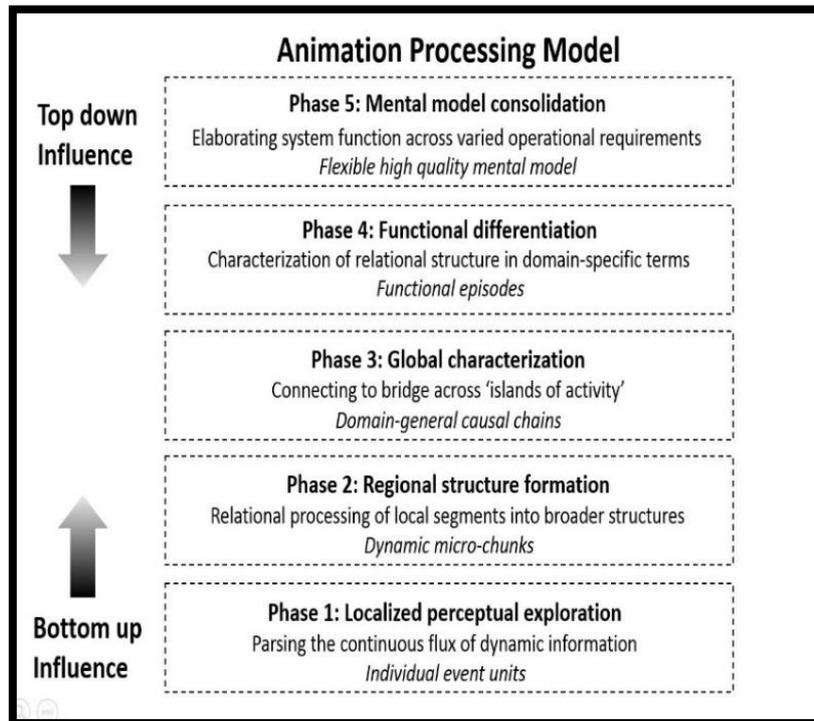


Figure 2.5. Le Modèle APM, Lowe & Boucheix (2008).

Dans la deuxième étape, les unités individuelles sont mises en relation les unes avec les autres pour créer des ensembles plus larges appelés « dynamic micro chunk ». Ces mises en relations peuvent être liées à la proximité spatiale entre deux unités et/ou à la cohérence des relations qu'elles entretiennent. Dans la troisième étape ces micro chunks sont connectés, rassemblés, pour créer une chaîne causale générale. Cette chaîne causale résulte de la compréhension que l'apprenant a de l'animation proposée dans l'espace et également dans le temps. Si cette étape est réalisée avec succès, les auteurs postulent que l'apprenant a pu dresser « *un tableau spatio-temporel bien structuré spécifiant le comportement de toutes les entités opérationnelles essentielles représentées dans l'animation* » Lowe & Boucheix (2008). Dans la quatrième phase on passe de la perception, à la compréhension fonctionnelle plus spécifique des relations

entre les différents éléments. L'apprenant ne s'attarde donc plus sur les unités spécifiques de l'animation mais sur la représentation globale qu'il se fait de l'image. Pour finir, dans la cinquième phase le sujet consolide le modèle mental établi, on parle alors d'un modèle mental flexible de haute qualité. Ce modèle nous permet de dire que chaque élément et les relations qu'il entretient avec les autres éléments d'une animation doivent être traités pour être compris puis intégrés en un modèle mental cohérent. La question que nous pouvons désormais nous poser est la suivante : est-ce que les animations permettent un meilleur apprentissage que les présentations multimédias statiques ?

2.4 Les méta-analyses.

Apprendre à partir de textes et d'images semble plus complexe lorsque la présentation fournie est animée plutôt que statique. En effet, dans le cas d'animations le rythme de visualisation est imposé, alors que dans le cas de présentations statiques, l'apprenant est libre d'explorer les informations au rythme qui lui convient. En un sens, la présentation statique peut aussi laisser place à la divagation quand l'animation peut forcer la concentration lorsqu'on souhaite réellement comprendre l'information présentée. Au-delà du caractère attractif et possiblement motivant des animations, se pose la question de l'efficacité qu'elles engendrent en termes d'apprentissages. Nous nous appuyons ici sur les méta-analyses réalisées afin de déterminer en quoi les animations offrent ou non un meilleur apprentissage et dans quelles conditions.

2.4.1 Comparaison de présentations statiques et animées.

Dans leur revue de questions, Berney et Bétrancourt (2009) ont comparé 50 études dont de nombreux paramètres spécifiques ont été extraits, comme par exemple : les différences interindividuelles des apprenants, les spécificités du domaine d'apprentissage proposé, ou encore la possibilité de pouvoir contrôler l'animation. Les résultats préliminaires obtenus ont montré un bénéfice de la présence d'animations pour l'apprentissage de phénomènes dynamiques, comparé aux présentations statiques. Les auteures ont stipulé que les habiletés visuospatiales des apprenants étaient positivement corrélées avec leur réussite aux tests dynamiques. Ceci dit, les auteures sont restées prudentes quant à l'interprétation des résultats obtenus en précisant que l'éducation et les connaissances des sujets étaient aussi des facteurs importants à prendre en compte. Les chercheuses ont finalement souligné la multiplicité des tests d'évaluation employés et le fait que les caractéristiques diverses des animations proposées dans les études sélectionnées variaient largement. Il n'était donc pas aisé d'établir une

conclusion précise sur les éléments responsables du bénéfice qu'offre l'animation sur l'apprentissage.

En 2016, Berney & Bétrancourt ont comparé 50 publications pour un total de 61 études afin de comparer l'impact du type de présentation statique versus dynamique, sur l'acquisition de connaissances. Leurs résultats ont de nouveau montré un bénéfice de la présentation animée comparativement aux présentations statiques pour l'apprentissage. Les auteures apportent également d'autres éléments, notamment sur l'aspect interactif du matériel : l'effet positif en faveur des animations n'était notable que lorsque les présentations animées n'étaient pas contrôlées ou contrôlables par le sujet. Les chercheuses notent tout de même que certaines études (ex : Boucheix & Guignard, 2005) montraient un effet positif de la présence du contrôle des animations contrairement à d'autres pour lesquelles aucun effet n'était ressorti, de ce fait il n'était pas possible de conclure à ce sujet de manière tranchée. Les résultats n'ont pas non plus montré d'effet de la présence d'éléments de signalisation, qui ne semblaient donc pas favoriser les performances liées à la présentation animée. Les résultats de cette méta-analyse allaient donc à l'encontre des conclusions de certaines études (Boucheix & Lowe, 2010 ; De Koning & al., 2010) mais ils pouvaient également s'expliquer par les résultats obtenus dans les travaux de Boucheix, Lowe & Putri (2013) que nous avons évoqués dans le chapitre précédent. Les auteurs ont expliqué cette absence d'effet par le fait que les éléments de signalisation pouvaient représenter un ajout d'informations visuelles étrangères relativement aux processus présentés, qui pouvaient entraver l'exploration spontanée des apprenant. On peut alors supposer que ces éléments ajoutés augmentaient la charge extrinsèque, ne permettant donc pas un apprentissage optimal. On peut également ajouter que ces éléments pouvaient détourner l'attention de l'information pertinente au profit du signal lui-même qui en soit, n'était pas pertinent à lui seul pour la compréhension d'un dispositif. Les auteures évoquent alors des facteurs importants à prendre en compte : le style cognitif des sujets, leurs connaissances préalables, l'objectif de la tâche et le mode de présentation des informations complémentaires. Les auteures ont malgré tout signalé un point crucial : la description du matériel utilisé n'est pas toujours suffisamment riche pour que des conclusions précises puissent être formulées.

2.4.2 Animé vs statique : la question de la fiabilité des comparaisons réalisées.

Comme nous avons pu le voir, une multitude d'études offrent une comparaison de versions animées versus statiques de dispositifs d'apprentissage. En abordant la question des différences inter-études, Berney & Bétrancourt (2009, 2016) ont relevé des points cruciaux menant à une

interprétation très prudente des résultats obtenus au sein de leurs méta-analyses. D'une étude à l'autre, on comparait certes les performances d'apprentissage entre statique et animé, mais une question restait en suspens : les animations, les illustrations statiques, les tests administrés, les tâches réalisées... ces éléments sont-ils vraiment comparables entre eux ? Peut-on « *comparer des pommes et des oranges* ? ». C'est la question que se sont posée Castro-Alonso, Ayres & Paas (2016) à travers une méta-analyse mettant en lumière les points d'ombre issus du matériel expérimental d'études comparatives statique/animé dans leur individualité. Ici, ils se sont consacrés au matériel expérimental lui-même, excluant la question des pré-tests et post-tests. S'il est tout à fait normal que d'une étude à l'autre le matériel soit différent, il est important que celui-ci soit strictement contrôlé et de ce fait, comparable au sein d'une même étude. « *Lorsque des facteurs tels que l'attrait, le support, le réalisme, la taille et l'interaction ne sont pas étroitement contrôlés entre les [images] statiques et les animations, les chercheurs peuvent souvent comparer des pommes avec des oranges.* » Les auteurs ont donc relevé différents biais suscitant le doute quant à l'interprétation des résultats qui en sont faits (Figure 2.6). Ils ont relevé sept biais principaux dans la littérature :

- Le biais d'attrait, qui renvoie aux caractéristiques attrayantes concernant la qualité des éléments visuels présentés, comme par exemple la couleur qui pourrait être différente d'une présentation à l'autre (noir et blanc versus couleur).
- Le biais de variété, quand le nombre d'éléments visuellement perceptible est plus élevé dans l'une des deux présentations.
- Le biais de média, lorsque qu'un même matériel est présenté sur des supports différents, par exemple une présentation papier comparée à une présentation sur écran d'ordinateur.
- Le biais de réalisme, qui se produit lorsqu'on compare une animation classique à une image statique photo-réaliste ou inversement.
- Le biais de nombre, qui renvoie au nombre d'images qui peut être plus élevé d'une présentation à l'autre.
- Le biais de taille, quand l'une des représentations proposées est plus grande que l'autre.
- Le biais d'interaction, quand les images ne correspondent pas dans leurs caractéristiques d'interactivité.

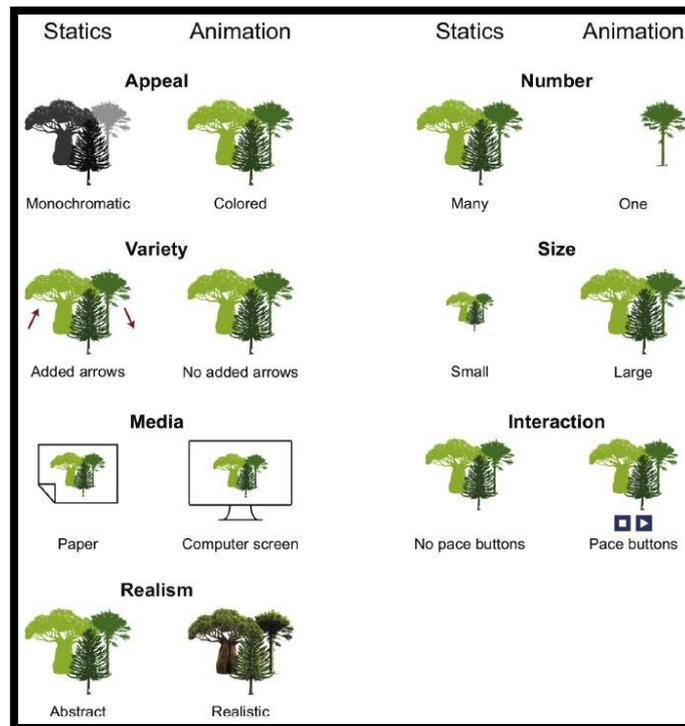


Figure 2.6. Illustration des 7 biais énumérés par Castro-Alonso & al. (2016).

Dans tous les cas, l'impact de ces biais est à double tranchant : leur présence peut pénaliser ou avantager l'un des groupes expérimentaux et peuvent rendre les résultats obsolètes. Par conséquent il est nécessaire d'être attentif à ces facteurs en amont de manière à produire des travaux rigoureux, mais aussi à postériori dans l'interprétation que l'on peut faire des résultats. Pour pouvoir conclure de manière adéquate, il faut pouvoir comparer ce qui est comparable et donc proposer des dispositifs étroitement contrôlés.

En plus du dispositif lui-même, la question de la composition des échantillons s'est posée. En plus de la comparaison de l'acquisition de connaissances à partir de présentations animées versus statique, le nombre d'hommes et de femmes dans les groupes expérimentaux a été analysé (Castro-Alonso & al., 2019). Selon les auteurs, il y a un déséquilibre de représentativité de chaque genre au sein des études réalisées. Ils ont alors regroupé 46 études dont différents paramètres ont été isolés : les caractéristiques des participants, le type de tâche et le domaine d'apprentissage. Les données recueillies suggèrent que dans la plupart des études, les groupes ayant étudié à l'aide présentations dynamiques avaient des performances supérieures à ceux dont les présentations étaient statiques, ce qui va dans le sens des études déjà réalisées. Concernant l'effet du genre, les auteurs ont montré que les présentations dynamiques semblaient être un avantage dans les échantillons où les femmes étaient moins représentées.

Selon eux, l'effet positif des animations sur l'apprentissage pourrait disparaître dans les échantillons où les femmes seraient plus nombreuses, ce qui les a menés à dire que les hommes seraient plus avantagés que les femmes par les visualisations dynamiques. Les chercheurs ont également montré un effet du niveau d'éducation des participants : plus les apprenants étaient jeunes, plus ils bénéficiaient des présentations animées. La présentation de matériel animé serait donc plus efficace pour l'apprentissage chez les élèves de primaires que chez les élèves de collèges, eux-mêmes plus avantagés que les étudiants en université. Cela pourrait s'expliquer par le fait que plus le niveau d'étude est élevé, plus les connaissances sont développées, un certain nombre de schémas mentaux seraient donc déjà intégrés (Kalyuga & al., 1998, 2000 ; Kalyuga & Sweller, 2005 ; Schnotz & Rash, 2005) et les animations seraient donc moins efficaces dans ce cas. Il y aurait ainsi un effet des connaissances préalables sur la qualité d'apprentissage observée. Les auteurs ont également montré des différences entre les domaines d'études abordés : les résultats obtenus étaient supérieurs dans les domaines de la géologie et autres sciences, de la biologie et de la médecine plutôt qu'en physique chimie, technologie, ingénierie et mathématiques. On peut donc penser qu'effectivement, les animations pourraient être plus avantageuses que les présentations statiques tout dépendant de la population à laquelle on s'adresse, des différences interindividuelles ainsi que du domaine traité.

2.4.3 Animations et habiletés visuo-spatiales

Dans les parties précédentes, nous avons mentionné les habiletés visuospatiales ; elles semblaient jouer un rôle important dans la perception et donc dans la compréhension des outils éducatifs que l'on peut proposer. Pour commencer, il est important de définir ce que sont les habiletés visuospatiales et de souligner certains des nombreux aspects qui les composent. Selon l'AQNP (Association Québécoise des NeuroPsychologues), les fonctions visuospatiales « *permettent de percevoir adéquatement les objets dans l'espace en déterminant leur orientation par les angles, la distance à laquelle se trouve un objet ou la direction dans laquelle un objet se déplace. Ces fonctions nous permettent, par exemple, de nous orienter correctement dans une ville ou dans un lieu* ». Selon le NRC (National Research Council, 2006), la pensée spatiale ou le raisonnement spatial implique la position et le déplacement d'objets et de soi, soit mentalement ou physiquement, dans l'espace. Il ne s'agit pas d'une procédure ni d'une habileté unique, mais en fait, d'un nombre considérable de concepts, d'outils et de processus. Pour Hegarty et Kozhevnikov (1999), cela implique aussi des procédures d'encodage et de transformation mentale des formes spatiales. Parmi les différents critères que Carroll (1993) a

proposé à ce sujet, nous retrouvons par exemple la visualisation spatiale, les relations spatiales, la perception de la vitesse, la mémoire visuelle, ou encore l'estimation de la longueur.

Höffler et Leutner (2007) ont réalisé une méta-analyse à partir de 26 études comparant les performances d'apprenants à l'aide de présentations statiques versus animées. Les auteurs ont distingué plusieurs types d'animations : les animations sous forme de vidéos ou les animations réalisées sur ordinateur. Les auteurs ont également catégorisé le type de réalisme des animations proposées selon quatre catégories : schématique, simple, réaliste, photo-réaliste. Aussi, les chercheurs ont veillé à différencier les vidéos représentationnelles dans lesquelles les informations étaient clairement fournies aux apprenants, des vidéos dites décoratives dont le but principal était de motiver les élèves. Les deux dispositifs éducatifs comparés, statiques et animés, devaient être équivalents ; c'est-à-dire que si l'un contenait des explications verbales il devait en être de même pour le second, de même que pour le temps d'apprentissage dédié à la tâche. Le type de connaissances, procédurales motrices, déclaratives et en matière de résolution de problème était également pris en compte. Les résultats obtenus ont montré un avantage des animations non interactives comparées aux images statiques. Les auteurs ont globalement obtenu un effet significatif moyen en faveur des présentations animées. Le type de tâche demandé a également montré un impact sur les apprentissages selon que la présentation était animée ou statique. En effet, les résultats ont montré un bénéfice des animations plus important dans le cas de tâches faisant appel à des connaissances procédurales motrices par rapport à des tâches qui requéraient des connaissances en résolution de problèmes ou de connaissances déclaratives. Les auteurs sont malgré tout restés prudents à propos de ces résultats du fait que les animations étaient représentationnelles pour les tâches motrices, ce qui n'était pas le cas des deux autres, combinant tantôt des animations représentationnelles tantôt des représentations décoratives. Cependant, les chercheurs ont comparé les performances des participants en fonction de leurs habiletés visuospatiales. Cette comparaison a montré que les participants ayant de bonnes habiletés visuospatiales réussissaient mieux les épreuves de visualisation statique que les sujets plus faibles. Aussi, les résultats ont indiqué que la visualisation 3D réduisait les performances des apprenants ayant de bonnes habiletés visuospatiales. Ceci conforte l'hypothèse de Höffler & al. (2007) selon laquelle l'animation fournirait un modèle « tout fait » aux apprenants ayant de moins bonnes habiletés spatiales. Cela expliquerait donc le fait qu'ils aient bénéficié d'une présentation animée, contrairement aux participants ayant de hautes habiletés visuospatiales qui eux avaient créé une représentation mentale du support présenté. Les auteurs ont alors évoqué un effet compensatoire. On peut donc dire que le niveau des habiletés visuospatiales des apprenants est très important dans l'apprentissage, notamment

quand celui-ci fait appel à des supports illustrés. Il faut donc prendre en compte les différences interindividuelles des apprenants afin d'adapter les supports d'apprentissage. On remarque que les mêmes éléments ont été relevés par Lowe (2003) lors d'une comparaison de présentations statiques versus animées dans l'apprentissage de phénomènes météorologiques. En effet, les résultats ont indiqué il n'y avait pas de différence significative en faveur de la présentation animée pour des sujets novices, cela dit, l'idée de l'élaboration d'un modèle mental en fonction du type d'habiletés spatiales était bien présente. Aussi, l'auteur a pu montrer un effet de saillance perceptive, c'est-à-dire que les apprenants orientaient plus volontiers leur attention sur des éléments perceptivement saillants plutôt que sur des relations de cause à effet cohérentes. Höffler (2010) a précisé ces recherches sur l'influence des habiletés visuospatiales dans l'apprentissage à partir d'un matériel éducatif animé versus statique. L'auteur a réalisé son étude à partir de 27 expérimentations issues de 19 études. Les résultats ont montré que les habiletés visuospatiales jouaient bel et bien un rôle dans l'apprentissage. Les apprenants qui avaient de faibles habiletés-visuo-spatiales étaient avantagés par les présentations animées et même par les animations 3D. Les apprenants ayant de bonnes habiletés visuospatiales étaient quant à eux moyennement avantagés par les présentations animées, ils montraient par ailleurs des performances similaires pour les présentations statiques. C'est également ce qu'ont montré les travaux de Berney, Bétrancourt, Molinari & Hoyek (2015) dans une tâche d'anatomie en 3D. Les résultats obtenus ont révélé un effet compensateur de la présentation animée 3D pour les apprenants ayant de faibles habiletés spatiales mais également une potentielle inversion de l'effet d'expertise pour les apprenants ayant de bonnes habiletés spatiales. On retrouve donc sensiblement le même profil de résultats que dans l'étude précédente, avec un effet compensatoire des présentations animées chez les apprenants ayant de moins bonnes habiletés spatiales. L'auteur indique malgré tout que les tests utilisés pour évaluer ces habiletés étaient différents d'une étude à l'autre, et que ceux-ci évaluaient les capacités spatiales statiques. Évaluer les habiletés spatiales dynamiques en évaluant par exemple la discrimination de vitesse ou l'appréciation des distances (ex : Gibson, 1947 ; Hunt & Pellegrino, 1988 ; Pellegrino & Hunt, 1989 ; Sanchez & Wiley, 2014 ; Porte, Boucheix & col., 2016) pourrait alors nuancer quelque peu les résultats obtenus. En effet, les résultats de ces recherches tendent à confirmer que les habiletés visuospatiales seraient différentes entre les modalités statiques et dynamiques.

2.5 Caractéristiques des animations dans l'apprentissage.

Les études que nous avons abordées précédemment ont montré un avantage des apprentissages à partir d'animations comparativement aux apprentissages à partir d'images statiques. Cet effet reste de taille moyenne dans la majorité des cas et la diversité des animations proposées n'ont pas toujours permis de déterminer en quoi les animations étaient plus favorables à l'apprentissage que les présentations statiques. C'est donc sur cette thématique que Ploetzner, Berney & Bétrancourt (2020) se sont centrés: « *Qu'est-ce-qui constitue exactement un argument pédagogique convaincant pour l'animation ?* ». Pour cela, les auteurs ont réalisé une méta-analyse à partir de 194 études portant sur l'apprentissage à partir d'animations, en prenant en compte les caractéristiques suivantes : les animations sélectionnées devaient respecter les caractéristiques déterminées par Ploetzner & Lowe (2012) c'est à dire, être visuelles, solliciter la perception d'un changement continu au moins pendant une certaine durée, être composées d'entités modélisées et permettant de visualiser simplement les changements affichés. L'animation devait avoir une visée pédagogique et non pas simplement motivationnelle ou artistique, elle devait également être fournie aux apprenants pour la tâche d'apprentissage et non pas être réalisée par ces derniers. Les animations sélectionnées devaient être centrales dans l'étude réalisée, c'est-à-dire qu'elles devaient représenter l'outil d'apprentissage principal de la tâche et non pas un prétest, un post-test ou un encart d'informations complémentaires par exemple. L'animation en question devait donc être évaluée de manière à fournir une quantification de l'apprentissage relatif aux informations qu'elle pouvait proposer. Pour finir, conformément aux remarques précédemment émises par Berney & Bétrancourt (2016), l'animation ainsi que les tests utilisés pour évaluer l'apprentissage qui en résultait devaient être décrits de manière suffisamment précise. Les caractéristiques de chaque étude sélectionnée ont été classées, prenant en compte le domaine d'apprentissage, le type de participants, le type d'animations, le type de post-tests (évaluation), le type de réponses données aux évaluations, les catégories perceptives et cognitives ainsi que les processus cognitifs engagés.

Cette recherche a montré que l'animation était typiquement utilisée pour enseigner du contenu scientifique. La plupart des animations proposées étaient schématiques et accompagnées de narration. L'utilisation de ces animations avait pour but, dans la plupart des cas, d'aider l'apprenant à créer un modèle mental conceptuel du processus présenté. Concernant les tests d'évaluation, ils étaient majoritairement verbaux, que les réponses soient à sélectionner ou à construire. Selon les auteurs, les animations schématiques guideraient l'apprenant vers un

raisonnement de type analytique et conceptuel, et la narration permettrait aux apprenants novices de compenser leur manque de connaissances sur le sujet dispensé ; la narration expliquerait alors les éléments abstraits, conceptuels, qui ne peuvent pas être modélisés par l'image. Les animations seraient un avantage pour l'apprentissage parce qu'elles offrent un plus grand nombre d'informations que les images statiques. Elles permettent de percevoir les changements directement alors que dans le cas de présentations statiques, il faut les inférer. Lorsqu'on indique un changement sur une image statique, à l'aide de flèches par exemple, il n'est pas certain que le mouvement soit compris d'emblée ce qui rend la tâche plus difficile et peut donner lieu à des erreurs. Les auteurs précisent cependant que bien évidemment, les changements et le déroulement dans le temps des éléments présentés doivent être pertinents pour que l'animation permette une compréhension supérieure à la présentation statique. Comme nous l'avons mentionné par les modèles de Schnotz & Bannert (2003) et de Lowe & Boucheix (2010), le traitement des informations statiques est différent de celui des informations animées. Cela mène les auteurs à soulever un point crucial qui est que bien souvent, on présente des animations pour mener les apprenants à créer un modèle mental conceptuel. Seulement, les méta-analyses semblent indiquer que celles-ci auraient une valeur éducative plus importante dans l'enseignement de modèles cinématiques. Il est donc nécessaire de poursuivre et d'approfondir les recherches à ce sujet de manière à comprendre plus spécifiquement les processus en jeu dans les animations pour les adapter au mieux aux enseignements.

Dans le cadre de nos travaux, nous nous sommes demandé si le fait d'optimiser les chaînes causales de transformations biologiques représentées par des visualisations dynamiques en réalité virtuelle, serait bénéfique pour l'apprentissage. En d'autres termes, est-ce que créer de la cohérence sémantique par le biais la présentation même des informations animées améliore l'apprentissage à partir d'animations en réalité virtuelle ? Nous aborderons cette problématique dans les expériences 2 et 3 de la partie empirique.

**CHAPITRE 3 - APPRENDRE AVEC LA
REALITE VIRTUELLE IMMERSIVE ET
NON IMMERSIVE.**

3.1 Définition et caractéristiques de la réalité virtuelle.

La réalité virtuelle, tout comme les animations, est composée de nombreux éléments visuels qui peuvent être plus ou moins dynamiques, avec la présence de sons, de légendes sous forme de texte ou encore de narration. Elle est définie comme « *un système multimédia complexe qui englobe une configuration pour l'immersion sensorielle ainsi qu'un moyen de représentation de contenu sophistiqué qui est capable de simuler ou d'imiter des mondes réels ou imaginaires* » (Mikropoulos & Natsis, 2011). Elle se distingue des animations par divers aspects, on peut d'ailleurs mentionner trois dimensions caractéristiques de la réalité virtuelle que sont la présence, l'immersion et le réalisme (Burkhardt, 2003). Lors de la création du matériel RV, les développeurs tentent de permettre un sentiment d'immersion et de présence à l'utilisateur pour que l'environnement soit perçu comme réel (Freina & Canessa, 2015). La présence renvoie à l'impression de l'utilisateur que l'environnement dans lequel il se trouve est réel, et l'immersion, c'est le sentiment de présence induit par la technologie utilisée (Mestre, 2006). En réalité virtuelle, « *l'immersion spatiale fait référence au sentiment d'être réellement là, physiquement présent dans un monde qui n'est pas réel* » (Freina & al., 2016). Ce sentiment viendrait du fait que l'apprenant ne voit pas simplement l'image, il est dedans et peut l'explorer à 360°, à échelle humaine dans le cas de l'utilisation d'un casque RV par exemple. Si la réalité virtuelle peut être physiquement immersive, elle peut aussi être qualifiée de « non immersive » quand « *elle place l'utilisateur dans un environnement 3D qui peut être directement manipulé, mais il le fait avec une station de travail conventionnel en utilisant un moniteur, un clavier et une souris* » (Robertson & al., 1993). Plus généralement, la réalité virtuelle « *peut englober des programmes qui sont simplement visualisés sur un écran plat, comme un moniteur de bureau ou une tablette, ainsi que ceux qui nécessitent l'utilisation de lunettes* » (Fabris & al., 2019). Quand l'environnement présenté sur un écran d'ordinateur est réaliste et qu'on implique le participant dans l'apprentissage, un sentiment d'immersion peut éventuellement émerger. La notion d'interactivité est également soulignée par Fabris & al. (2019), définissant la RV comme « *une expérience interactive dans laquelle on peut s'immerger dans un environnement généré par ordinateur* ». Dans la partie suivante, nous allons revenir plus précisément sur les notions de présence, d'immersion et de réalisme à travers un modèle dédié aux apprentissages en réalité virtuelle.

3.2 Le CAMIL, Makransky & Petersen (2021).

Quand on évoque la réalité virtuelle aujourd'hui, nous pensons d'emblée à l'innovation technologique ou à la nouveauté, mais en réalité, c'est dans les années 1950 que le premier dispositif de réalité virtuelle a vu le jour. La nouveauté réside véritablement dans la démocratisation de ce dispositif tant pour son usage personnel que scientifique. Peu d'études ont été réalisées à l'aide de ce type de matériel, du moins en psychologie cognitive. C'est en fait la démocratisation récente des casques de réalité virtuelle qui a suscité l'intérêt de son utilisation pour les apprentissages et la volonté de comprendre les mécanismes que cela implique. La recherche en psychologie cognitive a permis l'émergence de nombreux modèles relatifs au traitement de l'information dans les apprentissages, mais ce n'est que très récemment qu'a pointé le « Cognitive Affective Model of Immersive Learning » (CAMIL) par les travaux de Makransky & Petersen (2021). Ce modèle (Figure 3.1) répertorie les facteurs pertinents et distinctifs impliqués dans le processus d'apprentissage en réalité virtuelle immersive (RVI), dans une tentative de permettre d'optimiser la conception des dispositifs éducatifs à venir. Dans ce modèle, les chercheurs partent du principe que la méthode utilisée pour l'apprentissage en RVI peut être plus ou moins pertinente. Ils vont évoquer deux facteurs centraux : la présence et l'agence. Ces deux facteurs découlent du sentiment d'immersion, du fait de pouvoir contrôler le dispositif et de la fidélité de la représentation par rapport à la réalité. Ils vont à leur tour entraîner chez l'apprenant des sentiments plus ou moins (in)dépendants que sont : l'intérêt, la motivation, l'auto efficacité, l'incarnation, la charge cognitive et l'auto régulation. Cet ensemble de facteurs va avoir des répercussions plus ou moins importantes sur l'acquisition de connaissances factuelles, conceptuelles, procédurales, ainsi que sur le transfert de connaissances.

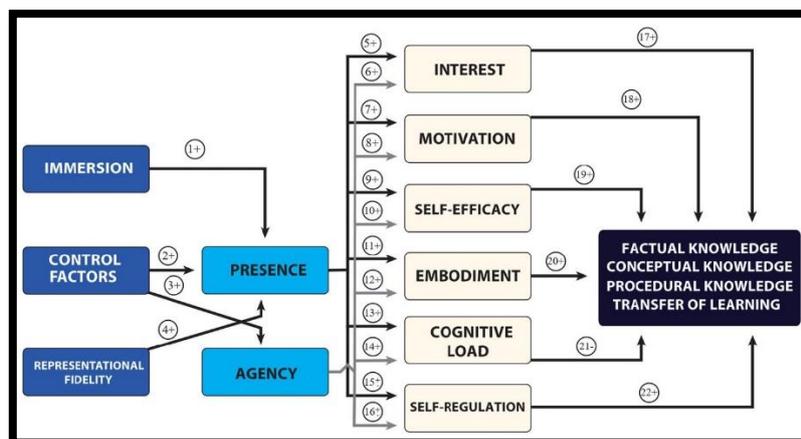


Figure 3.1. Modèle CAMIL de Makransky & Petersen (2021).

De manière plus détaillée, La présence renvoie au « *sentiment d'être là* » (Freina & al., 2016) et l'agence au « *sentiment de générer et de contrôler des actions* ». Les facteurs qui vont favoriser le sentiment de présence sont le degré d'immersion, le degré de contrôle alloué à l'apprenant et la fidélité représentationnelle de l'environnement proposé. Le facteur qui va mener à l'agence est l'interactivité, autrement dit, c'est le fait que l'apprenant puisse contrôler ses actions au sein du dispositif. La présence et l'agence vont également influencer un certain nombre de facteurs.

Pour commencer, ils vont influencer l'intérêt situationnel de l'apprenant, c'est-à-dire que certains stimuli vont attirer son attention à un moment donné. Cela peut amener l'apprenant à développer des comportements de recherche de connaissances, accroissant le désir d'en apprendre plus à propos de l'environnement proposé. Les auteurs indiquent également la disposition du participant à réengager le contenu au fil du temps. Ici, la présence renvoie à l'intérêt pour l'instant et l'agence provoque un effet positif sur l'intérêt situationnel.

Selon ce modèle, l'agence influence directement la motivation de l'apprenant qui va être étroitement liée à l'activité elle-même. Celle-ci peut s'avérer plus plaisante pour elle-même plus que pour ce qui peut en résulter en terme d'apprentissages. Les auteurs parlent alors de motivation intrinsèque.

Vient ensuite la notion d'auto-efficacité qui fait référence aux performances de l'apprenant. Il peut alors considérer l'activité réalisée dans l'environnement comme l'accomplissement de performances du fait que celle-ci soit vécue comme réelle.

L'incarnation quant à elle renvoie au corps, plus précisément à la possession d'un corps virtuel dans le cas de la RVI. Le contrôle des actions corporelles, l'agence, est positivement lié à l'incarnation dans le cadre de la réalité virtuelle immersive.

En RVI, on peut aussi observer une augmentation de la charge cognitive (cf. théorie de la charge cognitive, Sweller & al., 2011 ; théorie cognitive de l'apprentissage multimédia, Mayer, 2014). Quand la quantité d'informations à traiter surcharge la mémoire de travail, l'apprentissage peut être mis à mal. Les auteurs précisent qu'en RVI, le champ visuel est plus étendu que dans les présentations traditionnelles ce qui peut entraver la recherche d'informations pertinentes au profit de possibles distracteurs (détails attrayants). La charge cognitive extrinsèque peut donc être plus importante et perturber l'apprentissage. Cependant, tout dépend de la manière avec laquelle l'information est présentée.

Pour finir, ils décrivent l'autorégulation du participant comme sa capacité à s'adapter à l'environnement de manière à rester concentré sur la tâche d'apprentissage sans se laisser distraire par la présence possible de détails attrayants. Par contre, lorsque le niveau de présence et d'agence est très élevé, il devient très difficile pour l'apprenant de se concentrer sur les informations pertinentes. Il faut donc de nouveau veiller à concevoir un environnement pédagogique qui permette au mieux l'autorégulation en évitant l'intégration d'un surplus d'informations attrayantes et/ou inutiles. Une nouvelle fois, il faut rester prudent quant à l'organisation des informations au sein de l'environnement.

Concernant ce qui en résulte au niveau des apprentissages, les auteurs indiquent que la RVI n'est pas l'outil le plus propice à la création de connaissances déclaratives ou factuelles. En fait, l'efficacité de l'outil dans la création ou l'application de connaissances dépendra de sa conception. Par contre, comme nous pourrons le voir par la suite, la RVI peut s'avérer être un outil très utile dans l'apprentissage et/ou l'application de connaissances procédurales. Elle peut permettre au participant d'apprendre et de s'entraîner sans risque dans la réalisation de tâches procédurales complexes et/ ou potentiellement dangereuses. On peut par exemple penser aux interventions chirurgicales, comme la simulation de soins dentaires (Roy, Bakr & George, 2017) ou au pilotage d'avions. Cela permet un transfert des connaissances ainsi acquises en situation réelle et non pas une simple mémorisation de faits et de procédures (Thomas & Hooper, 1991). Le but est de créer un environnement d'apprentissage dans lequel tous les facteurs qui influencent la façon dont les personnes apprennent, sont équilibrés pour l'apprentissage. (Brandsford & al., 1999).

S'il est indispensable d'adapter les dispositifs de RVI aux apprenants de manière à ce qu'ils aient assez de ressources cognitives pour traiter correctement l'information, cette tâche de conception reste très complexe. Il est donc important de décrire de manière plus précise les travaux réalisés en réalité virtuelle et de comprendre les résultats à la lumière des facteurs énoncés dans le modèle CAMIL.

3.3 Un gain de connaissances en RV ?

Selon Makransky, Terkildsen & Mayer (2019), la simulation en réalité virtuelle « cause plus de présence, mais moins d'apprentissage ». Est-ce réellement ce qui est observé dans la majorité des travaux réalisés ?

Saputro & Setyawan (2020) ont intégré des sessions d'apprentissage en RV au sein d'un cours de physique dans le but d'évaluer si la réalité virtuelle offrait un gain de connaissances comparé aux supports pédagogiques classiques. La leçon portait sur le système solaire et a été dispensée à un seul et unique groupe de 22 élèves. Les auteurs faisaient passer un test et un post-test aux élèves, dans lequel ils devaient décrire le système solaire ainsi que la position et les caractéristiques des différents éléments de ce système. Les chercheurs ont comparé les performances des élèves avant et après la session RV et ont noté un gain de connaissances significatif après la leçon en réalité virtuelle par rapport au cours traditionnel. Ces résultats sont très limités du fait qu'il n'y ait qu'un groupe de petite taille et que les comparaisons n'aient pas été réalisées auprès de deux groupes homogènes distincts. La méthodologie employée était certes écologique du fait que les apprenants étudiaient en classe, en conditions réelles, mais elle n'était pas rigoureusement expérimentale.

Teranishi & Yamagishi (2017) ont également testé l'effet de la réalité virtuelle sur l'acquisition de connaissances procédurales dans une tâche de montage d'ordinateur. Les auteurs ont proposé une application permettant d'apprendre la tâche à l'aide d'un casque Oculus Rift et d'un capteur de mouvement « Leap Motion ». Suite à l'apprentissage, les auteurs ont proposé un post-test composé de questions sur la position des pièces et des questions sur le nom des pièces. Les résultats ont montré de meilleures performances en post-test concernant le bon emplacement des pièces, cependant les apprenants n'avaient pas retenu le nom de ces pièces. Les apprenants avaient donc bien intégré la procédure pour monter correctement l'ordinateur mais n'avaient pas mémorisé le nom des pièces. Cela dit, apprendre, « C'est s'assurer que les concepts clés soient compris plus que mémorisés » (Brandsford, 1999).

La réalité virtuelle immersive a également été utilisée pour évaluer l'apprentissage au cours de 4 expériences sur le thème du changement climatique, plus précisément sur l'acidification des océans (Markowitz & al., 2018) (voir Figure 3.2).



Figure 3.2. Exemple d'exploration en RVI, Markowitz & al., 2018

Les chercheurs ont sollicité 270 participants afin de tester l'efficacité de la RVI comme support d'enseignement. L'exploration océanique simulait une session de plongée pendant laquelle les élèves exploraient les fonds marins comme s'ils étaient immergés sous l'eau. Globalement, les auteurs ont relevé un gain de connaissances suite à l'exploration. Ils indiquent également que plus les élèves exploraient l'environnement, plus leurs connaissances augmentaient, qu'il s'agisse de collégiens, de lycéens ou d'adultes. Cela nous renvoie au facteur « intérêt » décrit par Makransky & Petersen (2021), stipulant que l'intérêt peut amener à rechercher le développement des connaissances.

Pour finir, la méta-analyse effectuée sur 29 études portant sur les apprentissages en RV d'Hamilton, McKechnie, Edgerton & Wilson (2021) a montré un effet significatif de la RVI dans le domaine de l'éducation. Le nombre d'études dont les résultats ne sont pas significatifs est moins important. Les auteurs notent cependant que la plupart des études mettaient en place des interventions courtes dans l'utilisation de la RV. Aussi, les performances de rétention n'ont pas été examinées. Concernant le domaine d'apprentissage, il portait majoritairement sur des enseignements scientifiques. Les études que nous aborderons ici vont d'ailleurs porter très largement sur des domaines scientifiques.

Ici, la réalité virtuelle est utilisée comme seul support qu'il s'agisse d'apprentissage procédural ou d'un complément pédagogique en classe. Il reste important de comprendre en quoi l'apprentissage en RV diffère des apprentissages traditionnels.

3.4 Comparaison de la RV à des supports traditionnels.

Pulijala, Pears, Peebles & Ayoub (2017) ont utilisé la réalité virtuelle afin d'évaluer les apprentissages à partir la simulation d'une pratique de chirurgie maxillo-faciale auprès de 95 internes en médecine.



Figure 3.3. Simulation de chirurgie maxillo-faciale (Pulijala & al., 2017).

Cette simulation a permis aux chercheurs d'évaluer le rôle de la RV sur l'apprentissage ainsi que sur la confiance en soi dans la pratique de ladite intervention. Les auteurs ont constitué deux groupes expérimentaux : le premier apprenait à l'aide d'un dispositif RV et l'autre groupe apprenait la même intervention à l'aide d'un PowerPoint. Les résultats ont montré que les apprenants ayant bénéficié de la simulation de l'intervention en RV avaient une confiance en eux accrue, c'est-à-dire qu'ils étaient plus confiants dans l'idée de pouvoir transférer leurs connaissances en situation réelle. Les auteurs observent également une augmentation des scores de connaissances pour les deux groupes qui eux, ne différaient pas significativement. Dans le cas présent, la RV n'offrait pas un meilleur apprentissage que les supports traditionnels, mais elle apportait une plus grande confiance aux futurs praticiens en vue d'une possible intervention chirurgicale en situation réelle ce qui pourrait dans ce cas, favoriser une meilleure réalisation de l'acte.

Au contraire, la RV a montré un effet plutôt positif sur l'apprentissage dans l'étude de Liu, Wang, Lei, Wang & Ren, (2020). Les auteurs ont proposé une expérimentation écologique en salle de classe où ils ont instauré des sessions d'apprentissage RV. Ils ont travaillé auprès de 90 élèves moyennant 6 cours de 45 min chaque semaine pendant 8 semaines sur le thème de la science.

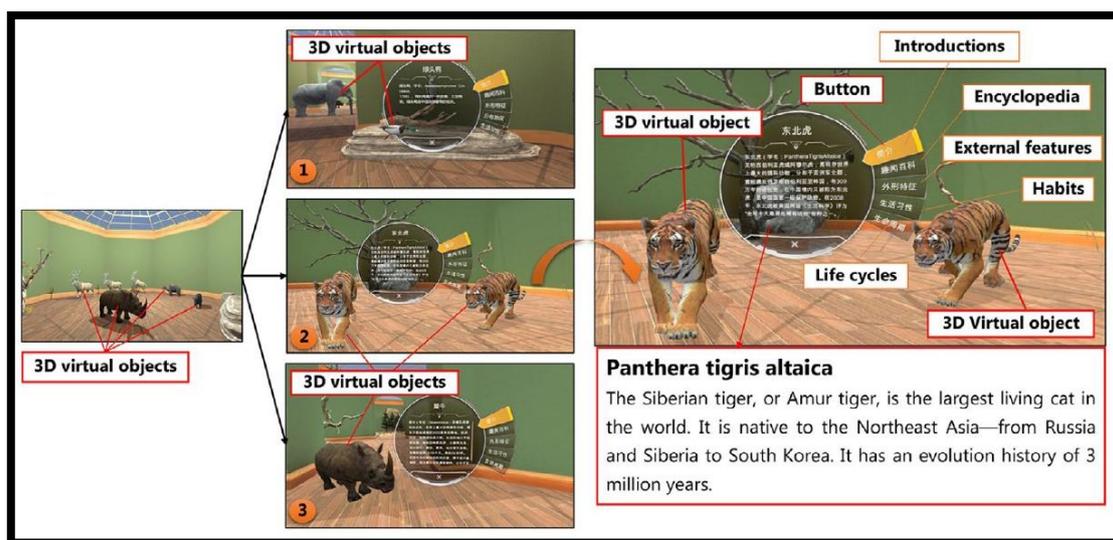


Figure 3.4. Exemple d'interface RV proposé par Liu & al. (2018) portant ici sur le monde des animaux.

Les élèves travaillaient par groupe de 5 et échangeaient le casque après une exploration de 4 à 5min. Les chercheurs ont comparé cet apprentissage en RVI à un apprentissage traditionnel. Ils ont mesuré les connaissances académiques via un prétest et un post-test (questionnaire vrai ou faux), ainsi que l'engagement et l'acceptation du matériel par les élèves. Les résultats ont

montré une supériorité des performances en connaissances académiques ainsi qu'un engagement supérieur en RV par rapport à l'enseignement traditionnel.

Grassini, Laumann & Skogstad, (2020) ont évalué le transfert de connaissances en proposant une tâche de construction d'avions miniatures à un échantillon de 30 apprenants. Les auteurs ont créé deux groupes : le premier visualisait une vidéo en 2D d'une durée d'environ 2min. Le deuxième groupe était immergé dans un environnement virtuel dans lequel, pendant 10min, il recevait des consignes de construction. La tâche de post-test consistait à construire l'avion en conditions réelles. Concernant le nombre d'erreurs commises pendant l'assemblage, aucune différence n'a été constatée entre les deux groupes. La part des participants qui parvenait à construire correctement l'avion ne différait pas non plus d'une condition à l'autre. Ici, la présentation vidéo 2D serait équivalente à la présentation en réalité virtuelle. Il est tout de même important de noter que le temps d'apprentissage était différent dans chaque condition ce qui pourrait expliquer ces résultats.

Maresky, Oikonomou, Ali, Ditzkofsky, Pakkal & Dallyk (2019) ont quant à eux comparé l'apprentissage en RV à un apprentissage traditionnel sur le thème de l'anatomie cardiaque, auprès de 42 étudiants de première année de médecine. Deux groupes expérimentaux participaient à l'expérience : le groupe contrôle composé de 14 participants étudiait l'anatomie à l'aide d'un support conventionnel. L'autre groupe, composé de 28 participants étudiait l'anatomie cardiaque à l'aide d'une session de 30 minutes en réalité virtuelle. Un QCM composé de 5 questions sur l'anatomie cardiaque conventionnelle et 5 questions visuospatiales était complété par les étudiants en prétest et en post-test. Le groupe RV a obtenu des scores supérieurs au groupe contrôle concernant les questions sur l'anatomie cardiaque conventionnelle. Cela dit, il est préférable de rester prudents dans l'interprétation de ces résultats sachant que les deux groupes expérimentaux étaient déséquilibrés.

Un bénéfice de la RV a également été constaté dans les travaux de Brewer-Deluce, Dak, Simms, Sinha, Mitchell, Shin, Saraco & Wainman, (2021). Les chercheurs ont comparé un enseignement traditionnel d'observation anatomique en laboratoire (sur des cadavres) à une application permettant d'explorer l'anatomie du corps humain soit en 2D, soit en 3D à l'aide de lunettes RV adaptées au smartphone. Le groupe de 38 étudiants était évalué dans les trois conditions. Les résultats obtenus ont montré que les étudiants réussissaient mieux en situation réelle et avec l'application en 3D (RV) qu'avec la présentation 2D. Cela a permis de montrer que la RV pouvait permettre un apprentissage équivalent à la pratique directe d'observations sur être humain en médecine. Les auteurs ont également précisé que plus de la moitié des

étudiants ont été malades avec la présentation RV (63%) ce qui a pu jouer un rôle sur la qualité de l'apprentissage. On peut également émettre une réserve sur les résultats obtenus compte tenu du fait que l'échantillon soit restreint.

La RVI a de nouveau été comparée à un support conventionnel dans l'étude de Parong & Mayer (2020). Leur étude portait sur l'apprentissage de la circulation sanguine humaine (Figure 3.5).

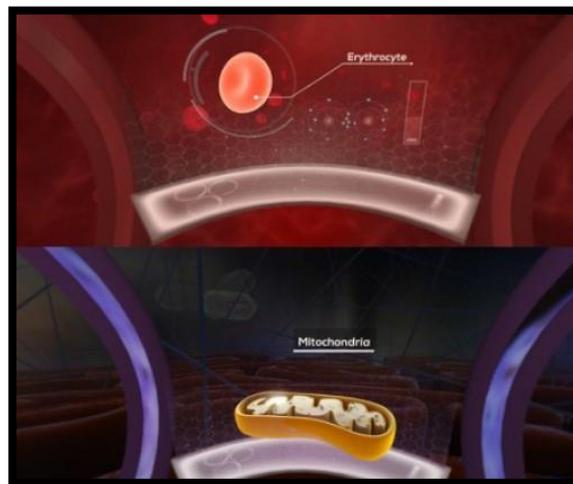


Figure 3.5. Exemple du matériel utilisé par Parong & Mayer (2020) sur la circulation sanguine.

Les auteurs ont créé quatre groupes expérimentaux : deux groupes apprenaient à partir d'une leçon en RVI, l'un voyait la leçon de 12min en continu alors que l'autre voyait 6 segments de 2min entre lesquels l'expérimentateur posait une question simple aux apprenants. Les deux autres groupes respectaient ces mêmes conditions à la différence que le matériel qui leur était proposé était un PowerPoint (PPT) composé de 24 diapositives issues de l'environnement RVI. Les groupes étudiant en RVI bénéficiaient d'explications narrées alors que les groupes PPT disposaient des mêmes informations sous forme écrite. Globalement, les résultats n'ont pas montré de différence significative entre les deux types de groupes pour le test de rétention. Par contre, la version PPT s'est montrée supérieure à la RVI en test de transfert. Pour la condition RVI, les auteurs ont montré que le degré d'enthousiasme et la charge extrinsèque étaient significativement supérieurs à la version PPT. Cela est en accord avec le modèle CAMIL (Makransky & Petersen, 2021) qui stipule que la motivation peut interférer avec l'apprentissage au sens où l'apprenant serait plus centré sur l'exploration elle-même que sur la tâche d'apprentissage. Il serait alors plus difficile pour l'élève de s'autoréguler afin de focaliser son attention sur l'information pertinente. Ceci dit, aucune différence n'a été observée entre les deux groupes pour la charge germane, ce qui indique que dans les deux cas, les apprenants

parvenaient à produire un modèle mental similaire. Ici, la charge extrinsèque aurait pu avoir un impact négatif important sur les apprentissages en RVI du fait du caractère interactif ponctuel de certains éléments de l'environnement virtuel.

Araiza-Alba, Keanel, Matthews, Simpson, Strugnell, Chen & Kauffman (2021) ont également étudié l'apprentissage en RV. L'objectif des chercheurs était d'évaluer les apprentissages de jeunes élèves en réalité virtuelle sur le thème de la sécurité aquatique à travers des ateliers de sensibilisation à l'école. Les auteurs ont constitué trois groupes expérimentaux : un groupe étudiait la leçon à partir d'une vidéo 3D à 360° dans un casque RV adapté à un smartphone, un autre groupe étudiait à partir d'une vidéo et le dernier groupe étudiait à partir d'un poster illustré. Les résultats ont indiqué une différence significative de l'intérêt et du plaisir pris par les élèves durant l'atelier entre le groupe RV et le groupe vidéo, ainsi qu'entre le groupe RV et le groupe poster. Concernant le sentiment de présence, les résultats ont montré que seuls 15,2% des élèves avaient indiqué ressentir un fort sentiment de présence en RV. Pour finir, les résultats ont montré que les scores aux post-tests d'apprentissage ne différaient pas entre les trois groupes qui, immédiatement après l'atelier, une semaine puis 8 semaines plus tard montraient un gain de connaissances sur le thème abordé.

3.5 La prise en compte du genre en RV.

Dans certaines études, le genre a été pris en compte pour déterminer si celui-ci avait une importance dans les apprentissages en RV, qu'il s'agisse du gain de connaissance ou de la conception du matériel lui-même. C'est le cas des deux études que nous allons évoquer ici. Pour commencer, Brian, Forest, Epstein & Brown (2018) qui ont travaillé sur le thème d'opérations militaires en RV.



Figure 3.6. Interaction du joueur avec l'officier dans le scénario RV (Brian & al., 2018).

Les auteurs ont créé quatre groupes expérimentaux et ont comparé le genre au mode de présentation du dispositif : support diffusé sur PC ou RV. Le jeu proposé consistait en un

scénario d'entraînement militaire dans lequel le participant jouait le rôle d'un officier de l'armée américaine en charge d'une mission humanitaire. Dans chacun des deux scénarios proposés il devait collaborer avec un officier de l'armée populaire de libération chinoise pour atteindre le but souhaité. Le participant disposait d'options de dialogue dont le degré d'adéquation variait. Après avoir choisi une option de dialogue, un feedback explicatif lui était donné sur l'adéquation de l'option choisie. Un prétest et un post-test étaient réalisés via un dialogue en temps réel et un jugement de situations sous forme de QCM. Globalement, les auteurs n'ont pas observé de différences d'apprentissage significatives entre les supports. Cela peut s'expliquer par le fait que dans ce genre de tâche, la RV n'est pas forcément indispensable, c'est-à-dire qu'elle n'aurait pas d'influence sur une prise de décision étant donné que le contexte ne nécessitait pas d'exploration particulière. Les auteurs ont malgré tout noté une interaction entre la RV et le sexe des participants, indiquant que la RV aurait un impact plus positif chez les hommes que chez les femmes.

Cet impact pourrait s'expliquer par le type d'interface ou d'interlocuteur proposé à l'apprenant. C'est ce qu'ont étudié Makransky, Wismer & Mayer, (2019). Les auteurs ont travaillé auprès de 66 élèves sur la sécurité en laboratoire en étudiant la relation entre le genre et le type d'agent pédagogique présent dans l'environnement RV (Figure 3.7).

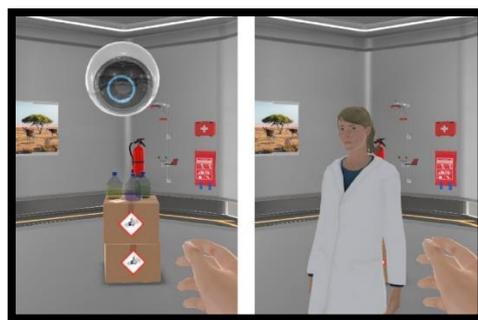


Figure 3.7. Type d'agent, drone ou Marie (Makransky & al. 2019).

L'agent avait pour rôle de guider les apprenants dans le laboratoire, de leur donner des consignes, d'expliquer les concepts théoriques et de poser des questions. Deux groupes mixtes équivalents étaient immergés dans l'environnement RV et étaient accompagnés pour leur apprentissage soit d'un drone, soit de Marie, une technicienne de laboratoire. Les résultats obtenus ont montré que dans les tâches de rétention et de transfert, les femmes obtenaient de meilleurs résultats lorsqu'elles apprenaient avec Marie. De même, les hommes obtenaient de meilleures performances lorsqu'ils étaient accompagnés du drone. Globalement pour les deux

groupes, tous sexes confondus, les chercheurs n'ont pas observé de différence d'apprentissage quel que soit l'agent. Les chercheurs ont donc conclu qu'un design spécifique au genre pourrait jouer un rôle important dans les apprentissages en RVI.

3.6 RVI et interactivité.

C'est avec le même matériel RVI utilisé par Parong & Mayer (2020) que Krassmann & al. (2020) ont étudié l'effet de l'interactivité en réalité virtuelle auprès de 36 élèves. Ils ont créé trois groupes expérimentaux : un groupe voyait une présentation RV interactive, avec la possibilité d'interagir avec certains objets. Le deuxième groupe voyait le même environnement mais ne pouvait pas interagir avec les éléments. Pour finir, le troisième groupe voyait un enregistrement 2D de la vue centrale de l'animation RV visionnée par les deux autres groupes. Chaque groupe étudiait la leçon pendant 12 min. Ils ont évalué les connaissances préalables et le gain de connaissances des élèves par le biais d'un QCM composé de 10 questions. Leurs résultats n'ont montré aucun effet concernant les tendances immersives, cependant un gain de connaissances a été observé pour chacun des groupes en posttest. Au sujet de la condition immersive, le fait de pouvoir observer les alentours naturellement dans l'environnement virtuel semble avoir permis une meilleure acquisition des informations nécessaires à la compréhension des concepts importants. Les performances étaient donc meilleures en condition RV qu'en condition vidéo. Cependant, l'interactivité en RV demandait plus de ressources en mémoire de travail ce qui conduisait à une surcharge de la mémoire de travail entraînant un apprentissage plus difficile. Les auteurs ont également étudié le sentiment de présence et ont observé que celui-ci était significativement plus important en RV interactive qu'en RV non interactive qui elle ne différait pas de la présentation vidéo.

Un profil de résultat sensiblement similaire a été observé dans l'étude d>Allcoat & Mühlennen (2018) sur l'apprentissage de cellules végétales (Figure 3.8).



Figure 3.8. Exemple de présentation RV proposé par Allcoat & Mühlennen (2018).

Les auteurs ont proposé en prétest et en post-test un questionnaire composé de 17 questions dont 12 étaient des questions de rappel et 5 étaient des questions de compréhension. Entre ces deux étapes, les chercheurs ont proposé à trois groupes de travailler soit à partir d'un support traditionnel (manuel scolaire), soit à partir d'une présentation RVI interactive, soit à partir d'une vidéo en 2D issue de la présentation RVI. Dans les trois conditions, les élèves disposaient exactement des mêmes informations verbales écrites et devaient tenter d'apprendre au mieux pendant 7 minutes. Globalement, dans les trois conditions les auteurs ont observé un gain de connaissances avec une supériorité de la RV et du manuel scolaire sur la présentation vidéo. Les élèves rapportaient également un sentiment de confiance plus élevé en ayant appris avec la RV plutôt qu'avec la vidéo, lui-même plus élevé que celui lié à l'apprentissage à partir d'images. Concernant le post-test de rappel de connaissances, les performances observées étaient supérieures pour la RV comparée aux conditions vidéo et livre qui elles ne différaient pas entre elles. Par contre, concernant les questions de compréhension les apprenants ayant appris avec un manuel scolaire avaient sensiblement les mêmes performances que ceux ayant appris en RV, ces deux groupes ayant des scores significativement plus élevés que le groupe vidéo. La réalité virtuelle offrait un meilleur rappel, mais ne permettait pas une meilleure compréhension que les manuels scolaires. Cependant, elle offrait un degré de motivation et d'engagement plus important. Les auteurs sont malgré tout restés prudents quant à cette interprétation, soulignant le caractère nouveau de la RV pour les élèves.

3.7 L'effet de modalité en RV.

L'effet de modalité renvoie au fait que les performances d'apprentissage sont supérieures lorsque, dans une présentation multimédia, l'information verbale est présentée de manière auditive plutôt qu'écrite (Mayer & Moreno, 2002). Baceviciute, Mottelson, Terkildsen & Makransky, (2020) ont testé cet effet sur l'apprentissage en réalité virtuelle immersive. Les chercheurs ont proposé un environnement RV reproduisant un milieu hospitalier dans lequel le thème du sarcome était abordé auprès de 73 participants. Tous les participants étaient immergés dans le même environnement (Figure 3.9), mais trois groupes ont été constitués afin de tester l'effet du type de présentation de texte sur l'apprentissage.

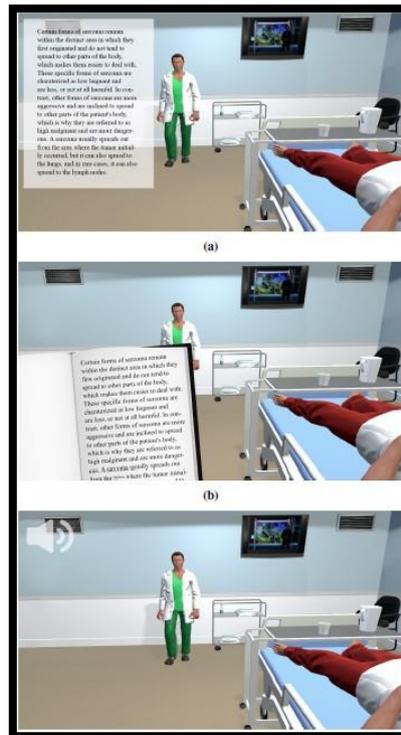


Figure 3.9. Conditions de présentation du texte (Baceviciute & al., 2020)

Un groupe étudiait le texte sous forme d'overlay (Figure 3.9a), le deuxième étudiait le texte à l'aide d'un livre (Figure 3.9b) et le troisième entendait l'information verbale sous forme de narration (Figure 3.9c). Les résultats n'ont rien montré de probant en tâche de transfert, par contre, en tâche de rétention la condition overlay montrait des résultats supérieurs aux conditions livre-audio. A l'inverse, pour les performances d'efficacité personnelle, la condition livre-audio montrait des résultats supérieurs à la condition overlay. Le fait de lire permettrait de mieux retenir les informations verbales que lorsqu'on ne fait que de les entendre. Les participants avaient également plus confiance en leur apprentissage lorsqu'ils consultaient le texte à l'aide du livre. Les auteurs ont expliqué que le contenu des représentations d'apprentissage provoquait des variations au niveau de la charge cognitive. Ici, le livre était le support pour lequel les charges intrinsèques et extrinsèques étaient les plus basses. Un effet du sujet et du format en relation avec le contexte de l'environnement présenté a été montré. Dans le cas présent, le contexte pouvait réduire la perception de la difficulté. Dans la condition narration, les apprenants prêtaient plus attention aux objets visuels de l'environnement que dans la condition overlay par conséquent, comme aucune image présente à l'écran n'était en rapport direct avec l'information verbale entendue cela provoquait une perte d'informations.

3.7.1 Epurer l'environnement RV pour une meilleure allocation des ressources attentionnelles ?

Cho, Lee, Ku, Jang, Kim, Kim, Lee & Kim (2002) ont travaillé sur le thème du contrôle attentionnel auprès de 50 enfants ayant des troubles de l'attention (TDA/H). Les chercheurs ont créé 5 groupes de 10 participants ; il y avait un groupe contrôle, un groupe RV sur PC avec formation biofeedback, un groupe RV sur PC avec formation cognitive, un groupe RV avec EEG biofeedback et un groupe RV avec formation cognitive. Ici nous ne nous centrerons que sur la formation cognitive et pas sur le biofeedback. Pour la formation cognitive en RV, l'élève était immergé dans une salle de classe intime et immersive (Figure 3.10) de manière à favoriser la focalisation de l'attention sur les éléments pertinents. Les chercheurs ont proposé une formation d'attention soutenue aux participants.



Figure 3.10. Exemples du matériel de Cho & al. (2002), avec les tâches VRCT et VRST.

Le participant était alors assis devant un bureau virtuel et pouvait voir trois drapeaux colorés (rouge, jaune et violet). Il n'était autorisé à répondre que lorsque le drapeau était violet. S'il répondait avant, un bip signalait le fait que la consigne n'avait pas été respectée. La formation VRCT est une formation visant à améliorer l'attention focalisée et sélective. Ici, deux objets 3D apparaissaient sur le bureau ; ils étaient soit identiques soit différents. La tâche du participant était de cliquer avec le bouton gauche de la souris s'ils étaient identiques et avec le bouton droit s'ils étaient différents. A chaque étape, le temps de présentation des items était raccourci. La formation VRST renvoyait à une tâche d'attention soutenue. Ici un chiffre arabe apparaissait sur le bureau. Lorsque le chiffre 0 apparaissait, l'apprenant devait cliquer sur le bouton gauche de la souris sauf lorsque celui-ci était précédé du chiffre 8 auquel cas, on demandait au sujet de ne pas répondre. Ici, la durée de présentation augmentait

progressivement. Les résultats obtenus ont montré que la réalité virtuelle aidait les enfants ayant des troubles de l'attention à se focaliser sur une tâche donnée. Ici, il n'y avait pas de distracteurs, l'apprenant ne traitait donc que ce qu'il était censé traiter. En revanche, ce genre de tâche ne nécessitait pas de faire apparaître (et donc de traiter) de nombreuses informations en interaction à l'écran. Cela confirme cependant l'importance de la conception de l'environnement virtuel, qui doit respecter exactement les besoins de l'apprenant pour favoriser un bon apprentissage.

3.8 L'effet du guidage par signalisation en réalité virtuelle.

Comme nous avons pu le constater dans le premier chapitre, le guidage est un principe important dans les apprentissages. Il était donc légitime de se demander s'il était préférable de l'inclure ou non dans les apprentissages en réalité virtuelle.

Chen & Toh (2005) ont étudié la signalisation en réalité virtuelle auprès de 184 apprenants sur le thème des règles de circulation routière et des panneaux de signalisation. Les chercheurs ont proposé trois modes d'apprentissage différents : RV guidé, RV non guidé et non RV (support traditionnel). Les apprenants devaient également compléter un prétest et un post-test composé de 15 questions sur le thème abordé. Le mode guidé intégrait une aide à la navigation à savoir : un traceur indiquant en temps réel la position du véhicule sur une carte ainsi que des flèches directionnelles. Les résultats obtenus ont montré un avantage du guidage sur les performances des apprenants. Les auteurs ont alors conclu que les apprenants s'adaptaient à travers une expérience appréhendée par les sentiments et transformée par l'action. Selon les auteurs, le guidage permettrait donc de pouvoir proposer des environnements adaptés plus favorables à l'apprentissage.

Le guidage par signalisation en réalité virtuelle a également été étudié par Wan Ahmad Jaafar Wan Yahaya & Awaatif Ahmad (2017). Leur objectif était de tester l'efficacité de la signalisation sur l'apprentissage d'un rituel funéraire en réalité virtuelle non immersive. Ils ont proposé un tutoriel en réalité virtuelle de ce rituel funéraire à deux groupes d'élèves : le premier apprenait le rituel sans informations verbales, le deuxième étudiait le même rituel mais bénéficiait de l'ajout d'informations verbales (Figure 3.11). Les auteurs ont montré que les apprenants ayant bénéficié d'un guidage verbal surpassaient le groupe n'ayant pas bénéficié en tâche d'apprentissage et de transfert. Les résultats ont également indiqué un bénéfice de la pratique répétée de l'apprentissage en réalité virtuelle, après trois semaines de pratique.



Figure 3.11. Exemple de présentation du rituel, avec ou sans guidage verbal.

Albus, Vogt & Seufert, (2020) ont quant à eux étudié l'effet de la signalisation sur l'apprentissage en RV auprès de 107 élèves, sur le thème de la désalinisation de l'eau de mer (Figure 3.12).

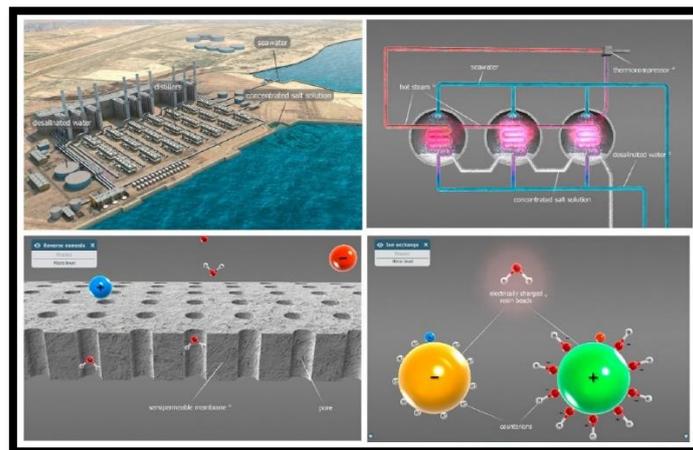


Figure 3.12. Exemples d'annotations ajoutées à l'environnement (Albus & al. 2020).

Les auteurs ont comparé deux groupes étudiant en réalité virtuelle dont l'un des deux groupes bénéficiait d'annotations sous forme de texte. Concernant la tâche de rappel, les apprenants ayant bénéficié des annotations écrites obtenaient des résultats supérieurs à ceux qui n'avaient bénéficié d'aucun signalement verbal. L'effet, bien que significatif reste faible. Aucune différence n'a été observée entre les deux groupes pour la tâche de transfert. La charge cognitive intrinsèque était la même pour les deux groupes, tout comme la charge extrinsèque. Les auteurs ont indiqué que, concernant la charge extrinsèque, leurs observations étaient en désaccord avec la théorie de la charge cognitive qui indique que celle-ci devrait être plus forte lorsqu'on signale les éléments pertinents. Les auteurs ont malgré tout conclu que la signalisation par annotations textuelles favorisait l'apprentissage en tâche de rappel, mais pas en tâche de compréhension ni en tâche de transfert.

Ces deux expériences nous mènent tout de même à penser que le guidage peut être un facteur déterminant dans les problématiques d'apprentissage en réalité virtuelle. Globalement, les travaux que nous venons de voir nous ont montré que, même si la RV semble avoir une valeur de motivation chez les élèves, elle n'est pas toujours plus propice aux apprentissages que les supports conventionnels. Il convient donc de l'utiliser dans des domaines qui nécessitent de l'immersion sans surajouter d'informations.

PARTIE EMPIRIQUE

Introduction.

Les travaux de recherche réalisés dans cette thèse visent à évaluer et optimiser les processus cognitifs liés aux apprentissages en réalité virtuelle dans le cadre du projet E-fran **SILVA NUMERICA**. Le but de ce projet est de créer et tester un simulateur de forêt vivante (et dont le développement évolue en fonction des actions humaines) en réalité virtuelle. Notre objectif est donc de tester les conditions de présentation des informations complexes de cet outil afin d'optimiser son efficacité et son utilisation par un large éventail d'apprenants et d'enseignants.

Le simulateur.

Le simulateur **SILVA NUMERICA** développé pour les besoins pédagogiques des collèges, lycées et établissements de formations professionnelles sylvicoles est composé d'images réalistes représentant un environnement forestier fidèle à ceux que l'on peut observer en France (voir Figure 1).



Figure 1. Exemple d'environnement forestier présenté dans le simulateur.

Ce simulateur offre un large éventail de possibilités d'exploration. L'apprenant peut se déplacer dans cet environnement comme s'il était réellement dans la forêt et l'explorer à 360 degrés. Il peut alors observer les houppiers¹ de l'intérieur de la forêt ; il dispose également d'une carte pour se repérer (Figure 2).

¹ Le houppier représente l'ensemble des branches qui forment la cime d'un arbre.



Figure 2. Exemple de la vue des houppiers et de la carte forestière.

De plus, l'apprenant peut consulter les éléments présents dans l'environnement virtuel, ainsi, il dispose des caractéristiques spécifiques à la parcelle explorée (Figure 3) mais aussi des caractéristiques relatives à chacun des éléments (Figure 4).

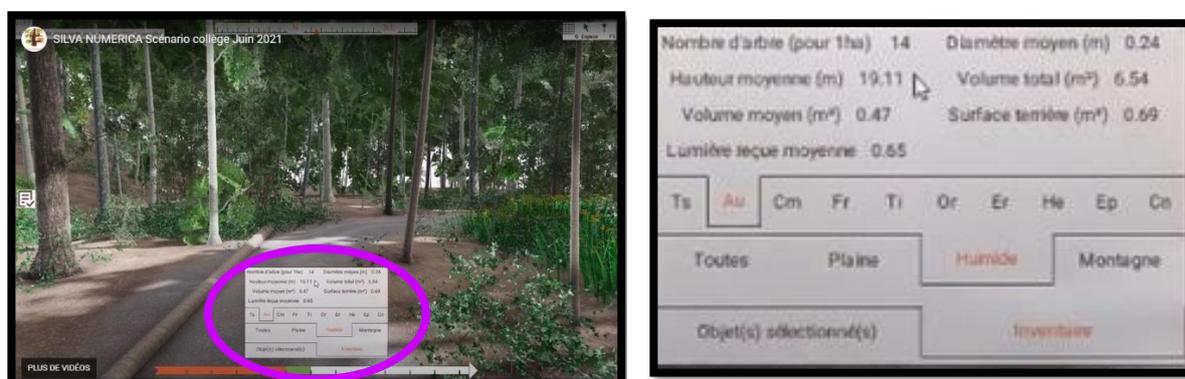


Figure 3. Exemple de tableau informatif détaillant les caractéristiques d'une parcelle.

Comme on peut le voir dans la Figure 3, le tableau informatif disponible permet de savoir le nombre total d'arbres à l'hectare, leur hauteur moyenne, leur diamètre moyen, le volume moyen et total en m³, la surface terrière en m² ainsi que le pourcentage de lumière reçue. Les onglets disponibles permettent d'obtenir ces informations moyennes pour chaque essence² d'arbre. Lorsqu'un onglet spécifique à une essence est sélectionné, le type de terrain approprié est précisé par le biais d'un onglet sélectionné juste au-dessous : plaine, humide, montagne.

Dans le cas d'un arbre isolé, l'apprenant peut prendre connaissance des caractéristiques suivantes : l'essence de l'arbre sélectionné, son âge, sa hauteur, et le pourcentage de lumière

² L'essence désigne une espèce d'arbres. <https://www.onf.fr/onf/raconte-moi-la-foret/la-langue-des-bois/+7f3::lessence-lespece-dun-arbre.html>

reçu. Ce cadre informatif indique également si l'arbre a l'écorce décollée, une coulée de sève, une cavité de pied, des micro-habitats ou des branches mortes dans le houppier. La hauteur de chaque élément est indiquée le cas échéant (Figure 4).

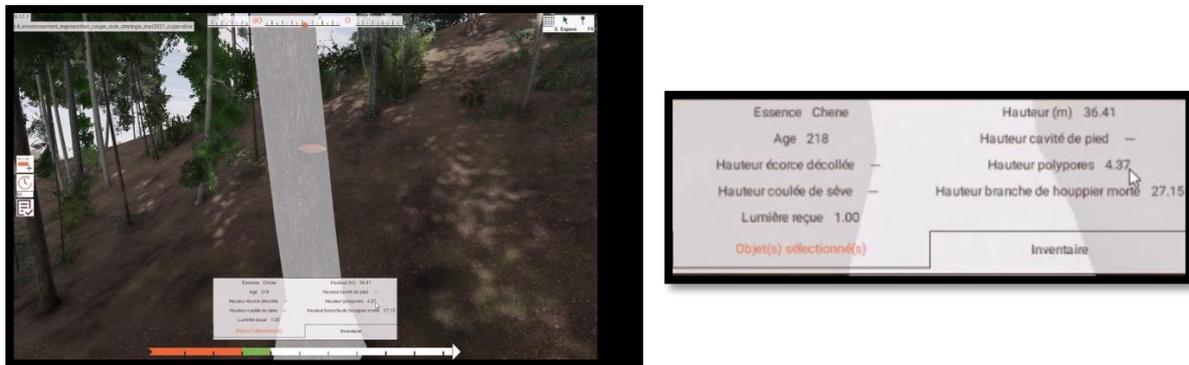


Figure 4. Exemple de tableau informatif pour un élément spécifique de l'environnement forestier.

L'apprenant peut aussi choisir de marquer des arbres. Il peut le faire dans l'objectif de couper ces arbres, ou de les rendre saillants pour indiquer qu'ils sont importants pour diverses raisons (la présence de micro-habitats, de champignons, d'oiseaux ou de branches mortes dans le houppier). Une fois marqués, les arbres prennent la couleur rouge (Figure 5), il est également possible de supprimer le marquage.



Figure 5. Exemple de marquage d'un groupe d'arbres.

Il est également possible de changer de point de vue et de type de représentation graphique. On peut observer la forêt du dessus, et changer la représentation graphique des arbres en activant le mode « sucettes » (Figure 6).



Figure 6. Visualisation de la forêt du dessus en mode normal (à gauche) ou en mode « sucettes » à droite.

Le fait de pouvoir analyser la forêt du dessus sous un mode classique permet de visualiser la densité des arbres. Lorsque certains arbres sont marqués, ce mode d'observation permet d'évaluer l'action réalisée : est-ce que je compte couper trop, ou pas assez d'arbres ? Est-ce que la zone est appropriée ?

Sous le mode « sucette », le fait de visualiser la forêt du dessus permet d'estimer, grâce à la légende, quelles essences d'arbres sont présentes et quelles essences d'arbres ont été sélectionnées. Ainsi, l'apprenant peut mieux comprendre la démarche qu'il a entreprise de plein pied au sein de la forêt et choisir de désélectionner les arbres qu'il juge devoir garder, ou en sélectionner d'autres dans le cas où sa sélection serait trop restreinte pour le type d'aménagement choisi (Figure 7).

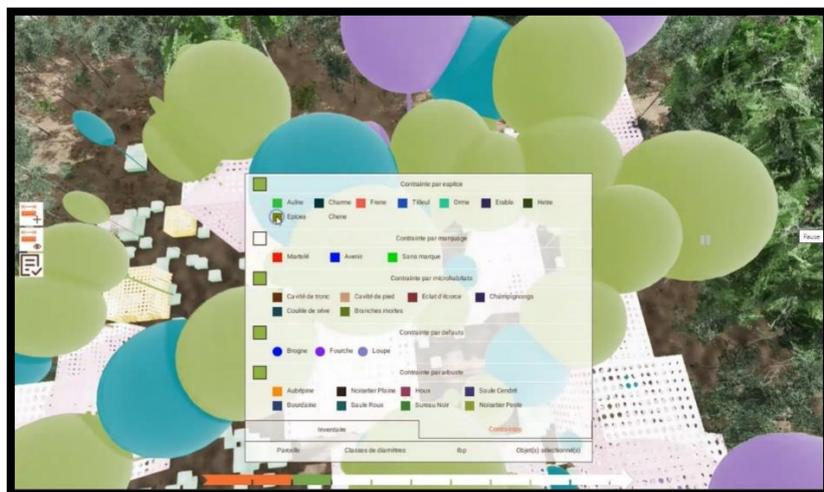


Figure 7. Tableau de sélection des légendes relatives aux arbres dans le mode « sucettes ».

Comme on peut le voir dans la Figure 7, il est possible de sélectionner chaque essence d'arbre en cliquant sur la légende appropriée pour le faire apparaître ou disparaître. Ainsi, l'apprenant peut évaluer la densité de chacune des essences présentes et réaliser le diagnostic correspondant à la problématique qui lui est posée.

Pour finir, le simulateur permet de visualiser une action à plus ou moins long terme. Que se passe-t-il lorsqu'on coupe des arbres ? c'est la question à laquelle les enseignants ont tenté de répondre par l'élaboration d'un scénario spécifique aux élèves de collège. Ce scénario montre un environnement avant, et après une coupe rase (Figure8).

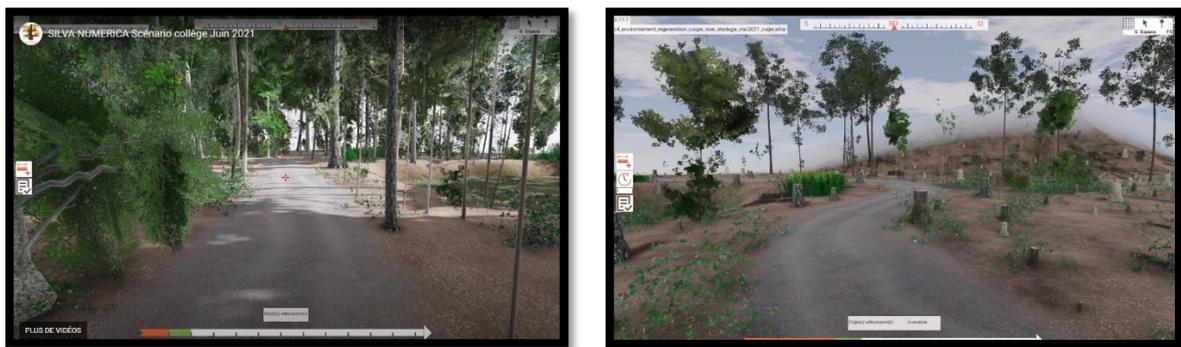


Figure 8. Modélisation de la forêt avant et après une coupe rase.

Le scénario indique qu'après cet aménagement, aucun arbre ne serait replanté de manière à laisser la biodiversité se développer. Pour visualiser l'impact de cette coupe rase sur l'environnement forestier, le simulateur permet d'avancer le temps sur plusieurs années, de cette manière, les élèves ont la possibilité de pouvoir se représenter l'évolution de l'environnement ainsi que le temps que cela représente (Figure 9).

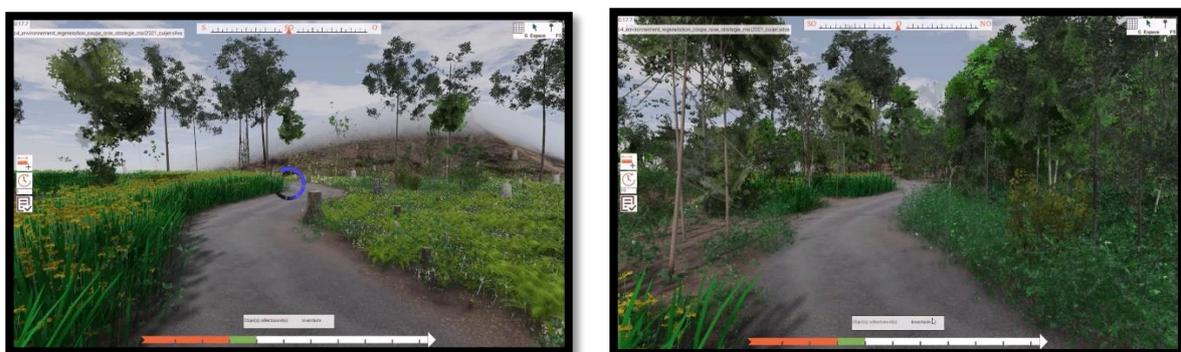


Figure 9. évolution de la forêt un an (à gauche) et 20 ans (à droite) après la coupe rase.

Un fichier Excel est également disponible pour relever les données observées pendant et après l'exploration (Figure 10).

Bilan d'étapes		Etape 0	Etape 1	Etape 2	Etape 3	Etape 4
		La forêt a été coupée	La forêt a 1 an	La forêt a 20 ans	La forêt a 50 ans	La forêt a 100 ans
Arbres et buissons						
	Les arbres sont présents, mais très espacés les uns des autres					
	Les arbres sont présents et très rapprochés					
	Les buissons sont présents, mais en petite quantité					
	Les buissons sont présents, en grande quantité					
Plantes herbacées						
	Le sol n'est pas recouvert de plantes herbacées					
	Le sol est recouvert de quelques petites taches de plantes herbacées					
	Le sol est recouvert de grandes taches de plantes herbacées					
Croissance des arbres *						
	Les arbres ont grandi (hauteur plus importante)					
	Les arbres ont grossi (tronc de plus grand diamètre)					
Lumière au sol						
	Il y a beaucoup de lumière au sol (plus de 4 « taches »)					
	Il y a peu de lumière au sol (3 ou 4 « taches »)					
	Il y a très peu de lumière au sol (1 ou 2 « taches »)					
Souches						
	Les souches (restes des arbres coupés) sont visibles sur le sol					
	Les souches (restes des arbres coupés) ne sont pas visibles sur le sol					
Micro-habitats						

Figure 10. Exemple de tableau Excel mis à la disposition des élèves.

Les apprenants de formation professionnelle disposent de quelques options supplémentaires permettant un aménagement plus fin des zones sylvicoles. Ils ont la possibilité de planter différentes essences d'arbre dans différentes stations³ forestières (Figure 11), de créer des zones touristiques ou des places à bois ou de couper des arbres.

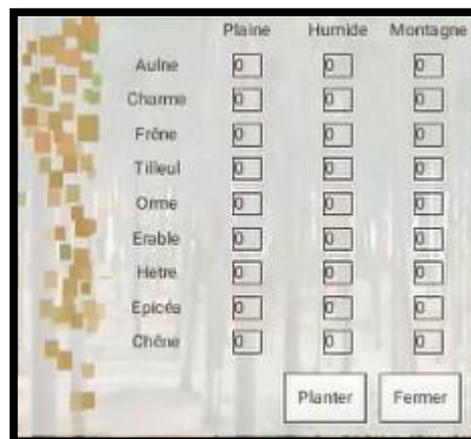


Figure 11. Fenêtre permettant de choisir les essences à planter ainsi que la zone forestière.

Ils peuvent choisir entre deux types de martelage : un martelage pour avenir qui colore les arbres sélectionnés en bleu signifiant qu'ils doivent être conservés, et un martelage pour coupe qui

³ La notion de station forestière désigne « une étendue de terrain homogène sur les plans du climat, du relief, du sol et de la végétation spontanée » (<https://www.cnpf.fr/n/les-stations-forestieres/n:1030>)

colore les arbres en rouge signifiant qu'ils vont être coupés. Le martelage peut être réalisé arbre par arbre ou par zone selon un diamètre déterminé par l'élève. Lorsqu'un martelage est réalisé sur une zone entière, l'apprenant a la possibilité de dé-marteler certains arbres, ou de leur attribuer la qualité de « martelage avenir » en les sélectionnant (Figure 12).

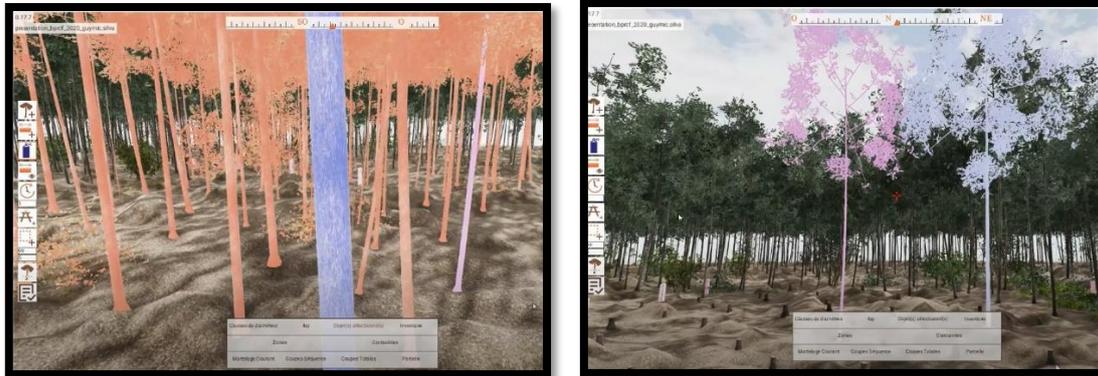


Figure 12. Zone d'arbre martelée pour coupe dans laquelle deux arbres ont été sélectionnés pour avenir, et résultat après la coupe.

Lorsque des arbres sont sélectionnés pour martelage, l'apprenant peut avoir accès aux données qui s'y rapportent dans l'hypothèse d'une coupe. Il peut savoir le nombre de tiges sélectionnées, la surface terrière, le diamètre moyen ou encore le volume total que représente la sélection. Il a également accès à l'IBP c'est-à-dire, à l'Indice de Biodiversité Potentielle qui regroupe des informations sur le nombre de micro-habitats, le nombre d'arbres morts au sol, le type de milieu ainsi que les strates. Les strates représentent les principaux niveaux d'étagement d'un peuplement végétal (Figure 13).

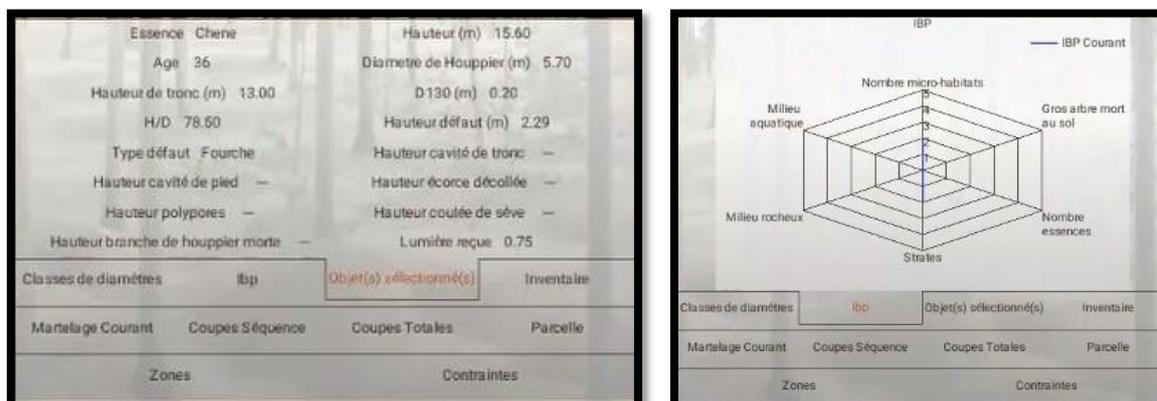


Figure 13. Informations disponibles sur les éléments sélectionnés après martelage et IBP.

Pour un arbre donné, les informations sont plus précises pour les apprenants professionnels que pour les collégiens. Ils disposent par exemple du diamètre du tronc à une certaine distance du sol, du diamètre du houppier ou encore de la présence de fourches.

Les apprenants ont également la possibilité de mesurer la distance qui sépare deux arbres (Figure 14) ou encore de sélectionner un arbre pour avoir la distance qui le sépare des autres arbres alentour (Figure 15).

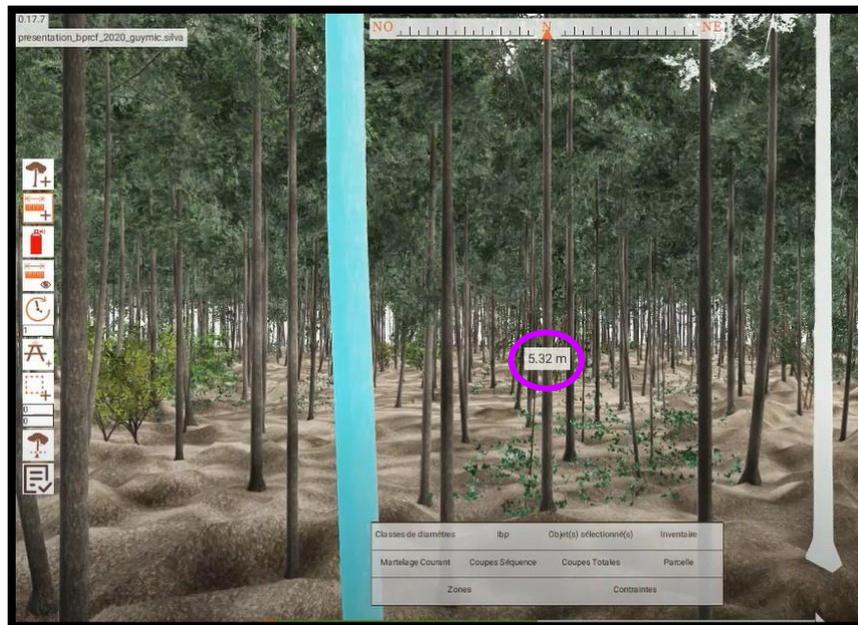


Figure 14. Illustration de la distance séparant deux arbres sélectionnés.

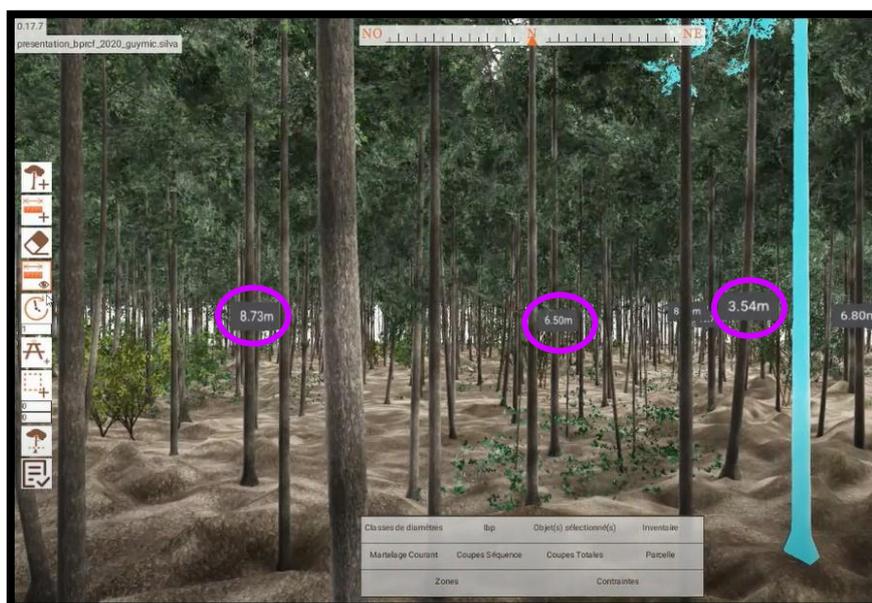


Figure 15. Illustration de la distance des arbres par rapport à un arbre référent sélectionné.

Le simulateur permet également aux élèves de placer des cloisonnements à l'aide de grilles selon une largeur et une distance choisie. Cela leur permet de délimiter une zone qu'ils ont la possibilité de marteler (Figure 16).

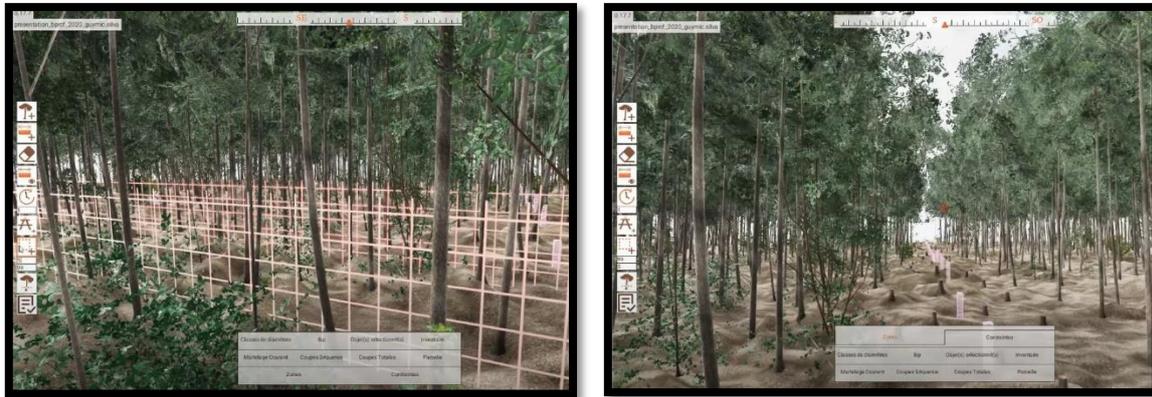


Figure 16. Zone de cloisonnement (à gauche) et zone de cloisonnement après coupe (à droite).

Les apprenants ont également accès à des ressources comme des vidéos, des photos, des tutoriels ou encore des tableaux Excel (Figure 17).

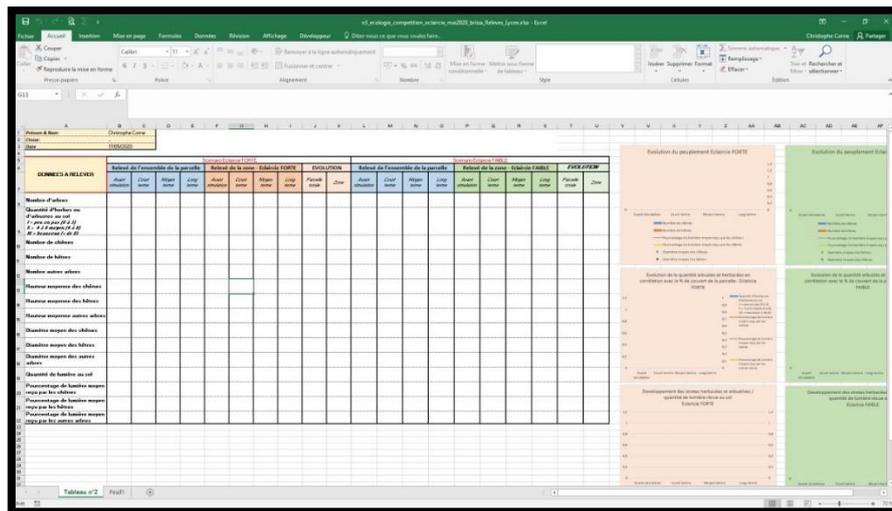


Figure 17. Exemple de ressources dont disposent les apprenants.

Ces tableaux peuvent être complétés par les élèves, et d'autres parties sont complétées directement par le simulateur qui enregistre les données au fil des activités de l'apprenant.

Toutes ces actions sont réalisées dans le cadre de scénarios pédagogiques réalisés par les enseignants. En effet, le simulateur donne la possibilité aux enseignants de créer leurs propres scénarios pédagogiques de manière à enrichir leurs enseignements en contextualisant les différents aménagements forestiers possibles. Cela leur permet également d'évaluer la mise en

application des connaissances acquises par les apprenants au cours d'enseignements réalisés en classe ou sur le terrain.

Notre matériel expérimental.

Dans le but de réaliser nos expérimentations, nous avons conçu et développé un matériel issu de la logique de ce simulateur, que nous avons adapté d'une part aux collégiens et d'autre part aux lycéens. Bien qu'offrant de nombreuses options de créations de scénarios pédagogiques, le simulateur n'était pas encore suffisamment développé pour être utilisé lorsque nous avons débuté nos travaux. En outre, cet outil pédagogique n'était pas adapté à nos besoins en termes d'expérimentations au sens où il ne permettait pas de tester les facteurs cognitifs que nous souhaitions étudier. Le simulateur tel que présenté ci-avant est la dernière version (V2) de l'année 2021. Il s'agit d'une version Alpha dont l'utilisation n'a été possible qu'à partir de 2021, bien après les deux premières expériences que nous avons réalisées. Quant à la première version de ce simulateur, son utilisation n'a été possible qu'au cours de l'année 2020, par conséquent, nous nous sommes inspirés des informations dont nous disposions à propos de celui-ci dans le but de créer le matériel expérimental de nos deux premières études. Concernant l'expérimentation 3, le matériel de réalité virtuelle immersive (RVI) que nous avons utilisé est une adaptation du simulateur réalisée par l'équipe en charge du développement du simulateur initial.

Nous avons commencé par créer une leçon en RV non immersive sur la décomposition de la matière organique du sol forestier (Figure 18). Ce thème faisait partie du programme scolaire des élèves de 5^{ème} et a été sélectionné conformément aux besoins mentionnés par les enseignants. Selon les enseignants de Sciences de la Vie et de la Terre, les élèves montraient des difficultés à se représenter et par conséquent à relier les différents acteurs de la chaîne alimentaire responsable de la décomposition du sol.

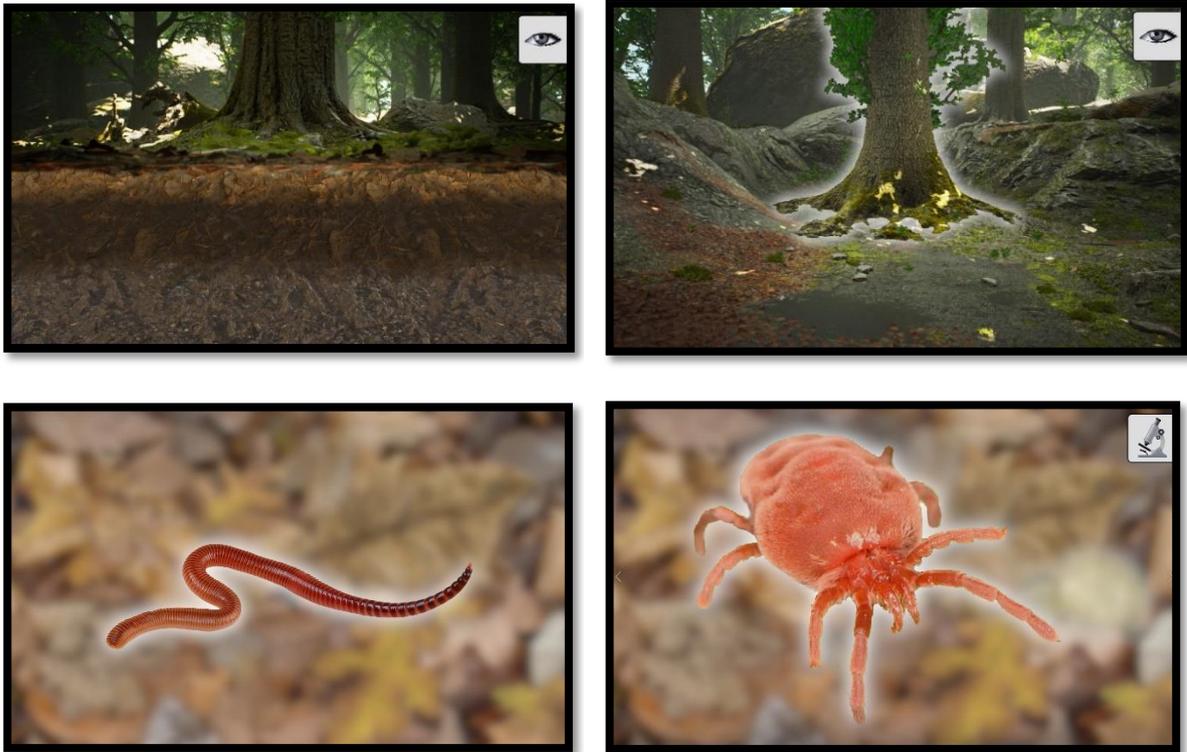


Figure 18. Captures d'écran du matériel expérimental en RV non immersive sur la décomposition de la matière organique destiné aux collégiens.

Nous avons ensuite créé une version en réalité virtuelle immersive (RVI) du simulateur **SILVA NUMERICA** à destination des lycéens, centré sur l'apprentissage du diagnostic des écosystèmes forestiers et de l'action humaine sur ces écosystèmes (Figure 19).



Figure 19. Capture d'écran de l'environnement forestier en réalité virtuelle immersive destiné aux lycéens.

Notre démarche expérimentale.

Comme nous l'avons évoqué dans la partie précédente, apprendre efficacement à partir d'environnements virtuels complexes, visuellement riches et dynamiques nécessite une approche basée sur des principes testés préalablement de manière à ne pas surcharger les canaux sensoriels et la mémoire de travail des apprenants. La réalité virtuelle repose sur l'exploration d'environnements généralement très complexes, et comme nous avons pu le constater, peu d'études ont encore été menées à ce sujet et les résultats obtenus sont plutôt mitigés. C'est pourquoi nous avons sélectionné les principes essentiels issus des recherches réalisées antérieurement sur les apprentissages multimédias. Notre but étant de valider ou non leur efficacité en réalité virtuelle qu'elle soit immersive ou non immersive. Nous avons donc sélectionné trois principes essentiels dont l'objectif général est le guidage des traitements perceptifs, attentionnels et cognitifs : (i) le principe de cohérence sémantique de la structure conceptuelle des informations présentées, (ii) le principe de contiguïté temporelle dans l'intégration des informations verbales et picturales, et enfin (iii) le guidage par signalisation visuelle accompagné d'un feedback informatif sur les éléments cruciaux de l'environnement.

Pour commencer, notre objectif était de mettre en évidence l'importance de l'effet de la cohérence sémantique du guidage de la leçon sur la compréhension et la mémorisation lors d'un apprentissage complexe chez de jeunes élèves. Pour cela nous avons réalisé une première étude exploratoire, sur le thème de la décomposition de la matière organique et nous avons testé les principes de cohérence sémantique de la structure conceptuelle de la leçon en réalité virtuelle non immersive auprès d'élèves de 5^{ème}. Cela nous a également permis de mieux cerner comment les élèves s'appropriaient le matériel de manière à organiser au mieux les expérimentations suivantes.

Nous avons ensuite mené une étude, également auprès de collégiens de 5^{ème}, sur le principe de contiguïté temporelle en réalité virtuelle non immersive. Nous avons de nouveau abordé le thème de la décomposition de la matière organique. Cela nous a permis de nous positionner par rapports aux résultats mitigés obtenus dans les travaux antérieurs et de déterminer l'amplitude des perturbations provoquées par un décalage entre la narration et l'image dans l'apprentissage. Les résultats obtenus nous permettent également de faire le lien avec les performances des élèves dans l'expérience réalisée en réalité virtuelle immersive.

Pour finir, nous avons expérimenté avec des lycéens de première et terminale en Sciences et technologies de l’Agronomie et du Vivant (STAV) en réalité virtuelle immersive, pour tester le principe de signalisation et de feedback. Ici, le thème abordé était celui de l’aménagement forestier dans le respect de l’écosystème. Nous avons alors pu comprendre dans quelle mesure le guidage peut aider les apprenants à raisonner sur l’environnement qu’ils explorent, en relation avec le feedback. Le feedback permettant aux élèves de réajuster leur jugement après une prise de décision lors de l’exploration de la forêt virtuelle.

En résumé, cette partie sera composée de 4 chapitres (voir Tableau 1).

Tableau 1

Résumé des expérimentations réalisées.

Expérience n°	Chapitre n°	niveau	Principe(s) étudié(s)	Thème	Type de RV
1	4	5 ^{ème}	Cohérence sémantique	Décomposition de la matière organique.	Non immersive
2	5	5 ^{ème}	Contiguïté temporelle		
3	6	1 ^{ère} & T ^{le} STAV	signalisation & feedback	Gestion de l’écosystème forestier.	Immersive

Nous tenterons alors de répondre à la question globale que nous nous posons : Est-ce que les principes d’apprentissage multimédia s’appliquent en réalité virtuelle ?

**Chapitre 4 - expérience 1, Le principe de
cohérence sémantique en réalité virtuelle non
immersive.**

4.1 Introduction.

Le but de cette première étude était de tester un principe de guidage de la cohérence sémantique en l'appliquant à une présentation en réalité virtuelle non immersive. Nous étudierons également les mécanismes cognitifs sous-jacents à l'apprentissage dans ces conditions de guidage.

Bien que non immersive, la réalité virtuelle (RV) expose un grand nombre d'informations ce qui peut accroître le coût de l'apprentissage en termes de ressources attentionnelles, d'autant plus pour des collégiens (Figure 4.1).



Figure 4.1. exemple de quelques-uns des nombreux éléments présentés dans la leçon en réalité virtuelle.

En effet, les jeunes élèves possèdent moins de schémas internes automatisés que des apprenants plus âgés ayant généralement une quantité de connaissances préalables plus importantes sur un plus vaste éventail de domaines (Kalyuga & al., 2000). Dans la théorie cognitive de l'apprentissage multimédia (CTML, Mayer, 2009, 2014), le principe de cohérence fait référence au fait que l'on apprend et que l'on comprend mieux à partir de sources d'informations multiples à condition que les informations superflues et/ou non pertinentes dans une leçon multimédia ne perturbent pas le traitement des informations pertinentes, et donc la compréhension et l'apprentissage (Mayer, Heiser & Lonn, 2001 ; Mayer, 2001 ; Mayer, 2002). Il est donc préférable de sélectionner et de relier entre elles de façon pertinente les informations essentielles par rapport aux notions à apprendre (Sweller, Ayres, Kalyuga & Chandler, 2003) de manière à rendre le dispositif présenté plus facilement traitable et intelligible cognitivement. La cohérence consiste à présenter les informations pertinentes par rapport aux concepts ou aux connaissances à acquérir, « au bon endroit et au bon moment » pour que les apprenants puissent les assimiler plus correctement.

Dans le cas de la cohérence sémantique, plus rarement étudiée dans le domaine de l'apprentissage multimédia (Mayer, 2014, chapitre 4), l'objectif est de concevoir une « chaîne »

conceptuelle « logique » structurée et cohérente des notions à acquérir à partir des informations imagées et verbales. Il s'agit de cohérence intégrée verbale et picturale. Cette exigence implique également une représentation de l'information imagée à laquelle elle correspond (Chandler & Sweller, 1991).

Le guidage que nous dénommons sémantique vise à accompagner l'apprenant dans le processus de sélection des informations et d'organisation de l'exploration d'un dispositif multimédia, à travers l'ordre de présentation des notions, la création de liaisons explicites entre des concepts, et éventuellement par le biais d'ajout d'indices qui peuvent par exemple se présenter sous forme de mots clés. Le guidage sémantique peut ainsi prendre les formes suivantes : l'insertion de chaînes causales (ou chronologiques) ordonnées, structurées et en surbrillance, l'intégration de mots clés, l'ajout de flèches ou de couleurs (ex : Boucheix & Rouet, 2007 ; Boucheix & Lowe, 2010 ; Jamet & al., 2010). Ce type de guidage vise également à faciliter l'intégration des informations verbales et picturales.

Aujourd'hui encore, une « tendance », présente chez les « concepteurs-designers » d'outils d'apprentissage (et les formateurs) qui consiste à « laisser les élèves apprendre par la découverte » dans les environnements multimédias reste forte (Schmidt & al., 2007), d'autant plus en RV. Il apparaît cependant important de rappeler que de très nombreux travaux montrent que sans accompagnement, les apprentissages peuvent être très perturbés : insuffisamment circonscrits, incomplets, inappropriés et incohérents (Kirschner, Sweller, & Clark, 2006 ; Kirschner & Clark, 2007). Il semble donc nécessaire de guider l'apprenant, en particulier lorsqu'il est novice. En situation réelle, l'enseignant accompagne l'élève, qu'il s'agisse de l'exploration d'un lieu, d'une situation ou dans le cas de résolutions de problèmes. Ce qui paraît « naturel » dans les conditions d'enseignement classique l'est parfois beaucoup moins lorsqu'on propose un dispositif pédagogique numérique sur écran, simulant un aspect de la réalité. Dans le cas présent, le caractère motivationnel de la présentation RV (Makransky & Petersen, 2021) peut prendre le pas sur l'efficacité de l'apprentissage. Malgré un engagement important de l'élève, il se peut que cela ne favorise pas la compréhension des concepts développés dans l'environnement proposé (Allcoat & Mühlénen, 2018). De plus les environnements multimédias ne sont pas destinés à remplacer les enseignants, ce sont des outils pédagogiques dont le but est d'enrichir et d'offrir de nouvelles perspectives à l'enseignement conventionnel.

Si la cohérence a été étudiée dans le cadre des apprentissages multimédias (Mayer, 2009, 2014) à partir de documents numériques incluant des images et des textes (Harp & Mayer, 1997 ;

Lehman & al., 2007 ; Sheiter & Eitel, 2010 ; Jamet & al., 2008) ou encore des animations (ex : Schneider & Boucheix, 2007), tel n'est pas le cas de la réalité virtuelle pour laquelle très peu d'études ont été menées (Chen & Toh, 2005 ; Albus & al, 2020). Les travaux réalisés sur le principe de cohérence (Mayer, 2014) en réalité virtuelle sont donc très peu nombreux et aucun n'a porté sur l'effet de la cohérence sémantique dans l'apprentissage en réalité virtuelle.

La cohérence sémantique dans les apprentissages multimédias, que sait-on ?

Cohérence et ajout d'informations au dispositif d'apprentissage.

Une surcharge d'informations verbales, délivrée au cours d'un temps contraint, peut conduire à l'altération de l'apprentissage (Mayer & al. 1996). C'est également le cas lorsque d'autres types d'informations inutiles ou incohérentes sont ajoutés. C'est ce qu'a montré l'étude de Harp & Mayer (1997) qui consistait à ajouter des images attrayantes à un texte explicatif avec également des éléments de texte attrayants. Cette étude a montré que lorsqu'on pense qu'ajouter des éléments attrayants à un dispositif éducatif pourrait améliorer l'apprentissage, on obtient en fait l'effet inverse. Dans cette étude, les auteurs ont aussi montré l'existence d'une corrélation entre le type d'image présenté et les performances : les performances étaient d'autant plus élevées que les images proposées étaient cohérentes avec le contenu verbal approprié. L'attrait pour les images « décoratives », dénommées « *seductive details* », peut favoriser la motivation mais pas nécessairement la qualité de l'apprentissage. De plus, ce type d'images pourrait perturber la compréhension.

Lehman, Schraw, McCrudden & Hartley (2007) Ont également étudié l'effet de l'ajout d'informations attrayantes sur l'apprentissage. Les auteurs ont repris et adapté le matériel utilisé par Harp & Mayer (1997) sur le thème de la formation de la foudre et ont comparé les performances d'apprentissage en fonction du type de présentation : une version de la leçon multimédia avec une présentation originale standard, sans ajout de détails, était comparée à une version standard enrichie de phrases contenant des détails attractifs. Par exemple, la phrase initiale « *Les courants d'air ascendants et descendants au sein du nuage peuvent provoquer la formation de grêlons* » était comparée à la phrase « *Ces courants d'air ascendants et descendants au sein du nuage peuvent provoquer la formation de grêlons, car les gouttelettes d'eau sont ramenées dans la haute atmosphère froide* ». Les résultats ont montré que les apprenants ayant étudié un texte dans lequel des détails « attrayants » avaient été ajoutés, lisaient plus rapidement les phrases du texte initial que les apprenants ayant étudié uniquement le texte initial. La présence de ces phrases attrayantes semblait également interférer avec le

rappel des éléments importants du texte principal. Le nombre d'informations rappelé par les apprenants ayant lu le texte pourvu de phrases attrayantes était moins élevé que celui des apprenants n'ayant lu que le texte original. Les auteurs ont conclu que les détails attrayants interrompaient la construction de la chaîne causale interne du processus présenté, réduisant ainsi la compréhension globale du texte principal. Les ajouts de détails réduisent donc la capacité des lecteurs à mémoriser puis intégrer les aspects importants relatifs au processus présenté. En somme, on peut dire que les informations attrayantes augmentent la charge extrinsèque (Sweller, 2011) du fait de leur caractère non pertinent, ce qui perturbe l'apprentissage et la compréhension.

Les travaux réalisés sur la cohérence nous montrent que le fait d'ajouter des informations, et/ou des détails attrayants aux dispositifs multimédias perturbe l'apprentissage et la compréhension. Il est donc préférable de ne présenter que les informations pertinentes nécessaires à la compréhension.

Cohérence et guidage par l'ajout de repères représentationnels externes.

L'étude de Stull & Mayer (2007) a montré les bénéfices d'informations complémentaires au matériel multimédia dans les apprentissages, à travers une leçon sur le thème de la formation de la foudre. Leurs résultats ont révélé que proposer un schéma récapitulatif des notions abordées conjointement au texte permettait un apprentissage plus rapide et plus efficace que lorsque les élèves devaient détecter et synthétiser eux-mêmes les informations pertinentes. Dans le cas de cette étude, le schéma synthétisant l'information était ajouté conjointement au texte, il ne faisait pas partie du matériel initial.

Dans le cas des visualisations dynamiques, comme les animations, Boucheix & Lowe (2010) ont émis l'hypothèse selon laquelle il serait préférable de centrer l'attention des apprenants sur une chaîne conceptuelle plutôt que sur des aspects isolés des animations. Les auteurs ont proposé différentes versions d'un mécanisme de piano, faisant varier le type de signalisation causale du système dynamique. Les chercheurs ont montré que la propagation progressive d'une flèche mobile de couleur parcourant la chaîne causale offrait une meilleure compréhension que lorsque les mouvements des pièces du mécanisme étaient indiqués par des flèches conventionnelles, externes à la chaîne causale. Cette dernière modalité ne favorisait d'ailleurs pas plus l'apprentissage qu'une présentation sans signalisation physique. Les auteurs ont donc conclu que lorsque l'on guide l'apprenant dans la compréhension et l'intégration d'un processus animé, il est préférable que le guidage respecte la coordination spatio-temporelle

entre le début de l'évènement animé pertinent du mécanisme et l'attention de l'apprenant. Cette étude suggère l'idée que présenter une chaîne causale cohérente du contenu à comprendre en RV permettrait un meilleur apprentissage qu'une exploration plus libre du dispositif.

Albus, Vogt & Seufert (2020) ont comparé deux versions d'une présentation en réalité virtuelle sur le thème de la désalinisation de l'eau de mer. Une version dans laquelle étaient insérées des annotations écrites sur les éléments clés de l'environnement visuel proposé était comparée à une version de la même leçon mais sans annotation. Leurs résultats ont montré un effet significatif, bien que de petite taille, des performances des apprenants ayant bénéficié de la présentation avec annotations.

La réalité virtuelle comporte de nombreuses visualisations dynamiques, qui de surcroît sont souvent « hyper-réalistes », ce qui peut augmenter à la fois le nombre, la diversité et la complexité des informations présentées. De plus, en RV immersive, il est possible d'explorer l'environnement visuel à 360 degrés. Ces propriétés sont différentes de celles d'un document multimédia « simple » de type texte-illustration. Il pourrait donc être nécessaire de limiter la liberté de contrôle du participant et/ou de le guider dans l'environnement en RV à explorer.

Dans le cadre de notre première étude en réalité virtuelle non immersive, nous comparerons un mode de présentation libre des notions à apprendre et à comprendre, dans lequel les apprenants peuvent choisir d'explorer des séquences d'informations dans l'ordre qu'ils souhaitent, à un mode de présentation fixe, dans lequel les séquences sont présentées selon une organisation structurale de type chaîne causale, sémantiquement cohérente. Le mode libre fait écho aux méthodes d'enseignement rapportées par les enseignants de SVT du collège dans lequel nous sommes intervenus. Le fait est que les notions du processus biologique de la décomposition du sol étaient présentées de manière isolées lors de plusieurs séances de cours. Cette méthode conduirait à une intégration morcelée des informations de ce processus complexe, rendant difficile la mise en relation des différents éléments qui le composent.

Les habiletés visuospatiales dans le traitement d'informations multimédia.

Höffler (2010) a réalisé une méta-analyse dont les résultats ont montré que les apprenants ayant de faibles habiletés-visuo-spatiales étaient avantagés par les présentations animées de même que par les animations 3D. Les apprenants ayant des habiletés visuospatiales élevées étaient quant à eux moyennement avantagés par les présentations animées ; par ailleurs ils obtenaient des performances similaires pour les présentations statiques.

En résumé.

Les études sur le principe de cohérence (Mayer, 2014) ont principalement montré que les informations attrayantes mais non pertinentes perturbent l'apprentissage. En revanche, dans ces travaux, il est plus rare que l'on manipule l'organisation structurale et sémantique de la « chaîne causale » des informations présentées. Des ajouts d'informations externes au matériel peuvent être fournis, mais ils ne se sont pas intégrés au matériel initial : ces informations sont complémentaires. La cohérence sémantique, au sens de chaîne sémantique organisée et cohérente, a peu été étudiée dans le domaine des apprentissages multimédia faisant usage de visualisations dynamiques accompagnées de textes, et encore moins dans le domaine des apprentissages et de la compréhension en réalité virtuelle.

Questions de recherche et hypothèses.

Pour cette première étude, nous nous sommes interrogés sur la manière avec laquelle nous pourrions organiser la cohérence sémantique directement dans la leçon donnée, sans ajouter d'informations supplémentaires. Ainsi, nous nous sommes demandé si un guidage sémantique cohérent et fixe des notions à apprendre, selon une chaîne causale délivrée automatiquement par le système, était plus efficace pour la compréhension d'un phénomène complexe en réalité virtuelle qu'une présentation libre de ces notions, et dont la chaîne causale était déterminée par l'apprenant.

Par ailleurs, bien que les études antérieures portant sur la cohérence dans le domaine des apprentissages multimédias représentent de bons indicateurs pour nos travaux, il n'en reste pas moins que le matériel proposé dans ces recherches est relativement simple comparé à ce que la réalité virtuelle, immersive ou non, est en mesure de mettre en œuvre. La plupart du temps, les recherches ont porté sur des représentations de mécanismes ou de mouvements biologiques très schématisés avec des animations présentant un dynamisme relativement lent voir peu représentatif (Lowe & Ploetzner, 2016).

Cette étude a été menée conformément aux besoins communiqués par les enseignants de collège à propos de la thématique complexe que représente la décomposition de la matière organique. Ces enseignants ont rapporté des difficultés de la part des élèves à visualiser et à mettre en relation les différents acteurs de la décomposition du sol, du fait qu'ils ne soient pas tous perceptibles et donc identifiables à l'œil nu. Il était donc difficile pour les enseignants de

présenter les différents acteurs de cette chaîne alimentaire complexe tout en rendant compte de leur échelle et de leur rôle au sein du processus.

De ce fait, nous avons conçu une leçon sur la décomposition de la matière organique en forêt, composée de plusieurs séquences distinctes mais interdépendantes. A partir du dispositif RV, nous avons comparé une présentation fixe des séquences qui suivaient la chaîne causale du processus de décomposition du sol, à une présentation libre de ces séquences. Lorsque la présentation de l'information est cohérente, l'ordre dans lequel les différents éléments sont explorés peut-il avoir un impact sur la chaîne causale finale construite par les apprenants ?

Lors de la conception de notre matériel, nous avons synchronisé temporellement la narration et les images virtuelles. Comme l'ont montré les travaux d'Hegarty & Just (1993) et de Moreno & Mayer (1999), les performances de l'apprenant sont plus élevées lorsque le texte est présenté conjointement à l'information visuelle à laquelle il correspond. Aussi, on obtiendrait des résultats plus élevés quand l'animation est accompagnée d'informations verbales orales plutôt qu'écrites (Mayer, Heiser & Lonn, 2001). Afin de ne pas proposer d'informations superflues à l'apprenant, nous avons également pris soin de n'insérer que les mots clés pertinents, conjointement aux éléments visuels auxquels ils correspondent, simultanément à la narration. Ainsi, nous pensons provoquer un effet de redondance sans pour autant surcharger le canal visuel des élèves (Mayer & al., 2001 ; Lehmann & al., 2007).

Nous avons utilisé un paradigme expérimental de type (i) prétest, (ii) apprentissage, compréhension (ii) posttest. Le prétest nous a permis de recueillir et de contrôler les connaissances préalables de chaque élève ainsi que les habiletés visuospatiales. Cela nous a permis de créer des groupes expérimentaux homogènes et d'analyser l'influence des habiletés visuospatiales sur la compréhension de l'environnement, selon que l'exploration est fixe ou libre.

Nous avons émis deux hypothèses principales concernant les performances d'apprentissage et de compréhension des élèves, et les différences interindividuelles.

Hypothèses.

Notre première hypothèse, **H1**, prédit un effet du type de présentation sur les performances d'apprentissage et de compréhension entre le pré-test et le post-test. Nous pensons que les performances d'apprentissage des élèves ayant visualisé la leçon dans un ordre cohérent, prédéfini fixe, et suivant la chaîne causale du processus biologique, auront des performances

supérieures aux élèves qui choisissent l'ordre d'exploration des différentes séquences de la leçon.

Notre deuxième hypothèse, **H2**, prédit un effet des différences interindividuelles sur les performances d'apprentissage. Nous nous attendons à des performances d'apprentissage plus élevées pour les élèves ayant des habiletés visuospatiales élevées par rapport aux élèves ayant de plus faibles habiletés visuospatiales.

4.2 Méthode.

4.2.1 Participants.

111 élèves de classe de 5^{ème} issus de quatre classes d'un collège français ont participé à cette étude. L'expérience s'est déroulée selon deux sessions toutes deux espacées d'une semaine : (i) une session de pré-test (ii) une phase d'apprentissage suivie immédiatement d'une session de post-test. En raison de l'absence de quelques élèves à l'une des deux sessions, nous avons retenu un total de 103 élèves ayant terminé les tâches des deux sessions (âge moyen = 11.92 ans, 57 filles et 46 garçons, respectivement : 55.34% et 44.66%). Les élèves ont été classés selon leurs performances aux test d'habiletés spatiales et de connaissances préalables afin de créer deux groupes homogènes. Concernant les performances au test spatial, les élèves ayant obtenu un score supérieur ou égal à la médiane (67.50%) étaient considérés comme ayant un niveau élevé d'habiletés spatiales ; les élèves ayant obtenu un score inférieur étaient considérés comme ayant de plus faibles habiletés spatiales. Concernant les scores obtenus au questionnaire de connaissances préalables, les élèves ayant obtenu un score supérieur ou égal à la médiane (37.50%) étaient considérés comme ayant un niveau élevé de connaissances préalables. Les élèves ayant obtenu un score inférieur au score médian de 37.50% étaient considérés comme ayant de plus faibles connaissances préalables. Ainsi, nous avons déterminé un groupe d'élèves ayant un niveau élevé de connaissances préalables et d'habiletés spatiales (n= 50) et un groupe ayant un niveau plus faible de connaissances préalables et d'habiletés spatiales (n = 53). Dans chacun des deux groupes expérimentaux, nous avons attribué aléatoirement le même nombre d'élèves de chacun des groupes préalablement établis : performances élevées, performances faibles.

L'expérience s'est déroulée en salle de classe lors de la première session, et dans la salle informatique de l'établissement lors de la deuxième session. Au cours de la première session, les élèves ont répondu à une série de prétests comprenant un test d'habiletés spatiales et un questionnaire à choix multiples de connaissances préalables portant sur le sujet de la leçon en

réalité virtuelle à savoir : la décomposition de la matière organique. Les performances des élèves à ces tests ont été prises en compte et utilisées pour créer deux groupes homogènes (avec N = 53 et 50) correspondant aux deux conditions d'apprentissage, respectivement : présentation cohérente fixe et présentation libre. Une semaine plus tard, lors de la deuxième session, les élèves ont étudié individuellement la leçon multimédia RV sur écran d'ordinateur, sur les mécanismes de décomposition de la matière organique dans l'une des deux conditions expérimentales (une condition par groupe), après quoi ils ont complété les post-test. Chaque session a duré 60 minutes pour une durée totale de 120 minutes pour l'expérience dans sa totalité. La session expérimentale s'est déroulée par demi-groupes de 14 élèves en salle informatique. A la fin de la première session, c'est-à-dire après les prétests, les élèves n'ont pas été informés des activités prévues pour la deuxième session.

4.2.2 Matériel expérimental.

4.2.2.1 Pré-tests utilisés lors de la session 1.

Lors de la session de prétests, les participants ont complété deux tests : (i) un test d'habiletés spatiales de type rotation mentale, basé sur le matériel mis en œuvre par Shah & Miyake (1996) et adapté pour les enfants et les adolescents par Guignard & Boucheix (2005). Ce test évalue la mémoire de travail spatiale. (ii) un test de connaissances préalables sur la décomposition de la matière organique : il s'agit d'un questionnaire à choix multiples composé de 36 questions ciblant les connaissances factuelles et la compréhension, c'est-à-dire la qualité du modèle mental.

Le test d'habiletés spatiales comprenait 18 items présentés dans des livrets individuels et était organisé en deux subtests (Annexe A). Ces subtests ont été conçus à partir d'une tâche de mémoire de travail spatiale de Shah et Miyaké (1996). Dans le premier subtest, une lettre était présentée à l'élève (qui devait alors garder cet item en mémoire) et il devait déterminer, à partir de 5 lettres identiques qui occupaient toutes des positions différentes dans l'espace, celles qui représentaient la première lettre présentée après une rotation ; les autres lettres étaient identiques mais présentées en miroir. (Figure 4.1).

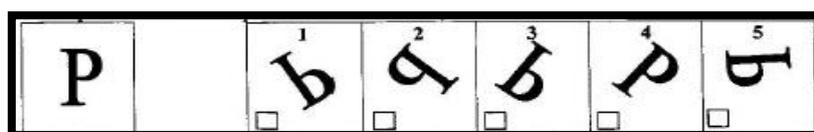


Figure 4.1. Exemple d'item du test d'habiletés spatiales, Shah & Miyaké (1996).

Dans le deuxième sous-ensemble, après la présentation de lettres, des objets communs puis des figures simples étaient présentés. Les élèves disposaient de 15 secondes pour répondre à chaque ligne d'items. L'expérimentateur indiquait lorsqu'il fallait passer à l'item suivant. Ce test adapté avait déjà été utilisé avec succès par Guignard et Boucheix, 2005. Le nombre de bonnes réponses (sur 40) a été calculé pour chaque participant.

Les pré-tests (et post-tests) de Connaissances factuelles et Compréhension étaient composés de 36 items (Annexe B) permettant de mesurer la compréhension des élèves. Ce test était présenté sous forme d'une série d'affirmations liées au thème de la leçon abordée en phase expérimentale. Pour chaque affirmation, les participants devaient choisir entre les réponses « vrai », « faux » ou « ne sait pas » (Figure 4.2).

14	Le humus est une couche de terre humide qui permet aux végétaux de se nourrir.	VRAI	FAUX	NSP
15	Le ver de terre est le seul animal à manger les feuilles de la litière.	VRAI	FAUX	NSP
16	Les bactéries du sol rendent les animaux de la litière malades.	VRAI	FAUX	NSP

Figure 4.2. Exemple d'items issus du pré/post-test de connaissances factuelles et de compréhension.

Certaines des questions posées concernaient le rappel de connaissances factuelles sur les mécanismes de décomposition du sol, tandis que d'autres faisaient appel à une compréhension plus approfondie, ce qui nécessitait la production d'inférences. Le type d'affirmation faisant appel à des connaissances factuelles était par exemple : « *Le pseudoscorpion mange d'autres animaux comme les collemboles et les oribates* » ou « *Sur les feuilles mortes on peut voir de petits fils blancs : ce sont des filaments de champignon. On appelle cela le mycélium* ». Concernant les affirmations qui nécessitaient une compréhension plus approfondie du processus écosystémique il y avait par exemple : « *Les êtres vivants et ce qu'ils rejettent sont de la matière organique* » ; ou concernant les animaux « *Ils mangent, digèrent puis rejettent la matière organique de la litière et créent ainsi une belle couche d'humus.* ». Le sujet des questions était lié au contenu (mots et images) délivré par la leçon de réalité virtuelle et concernait les écosystèmes du sol forestier en décomposition produisant de la matière organique et nourrissant les arbres de la forêt. Cela comprenait : la constitution et l'évolution des couches

du sol de la litière aux couches plus profondes ; le « travail » des micro-organismes, bactéries, champignons (mycélium) et animaux ; la chaîne alimentaire allant des micro-organismes aux animaux (tels que les vers de terre) et l'effet de la décomposition du sol sur la croissance des arbres. Ce type d'affirmations pourrait être déduit de la série de phrases et d'images de connaissances factuelles. Le questionnaire complet est présenté en Annexe B. Le même test a été utilisé pour le pré-test et l'un des deux post-tests. L'ordre des questions différait du pré-test au post-test et selon les participants : trois questionnaires contenaient les mêmes affirmations mais dans un ordre différent. Pour une moitié, les affirmations étaient vraies et pour l'autre moitié elles étaient fausses. Le nombre de bonnes réponses, d'erreurs et de réponses « Je ne sais pas » (sur 36 items) ont été codés par trois évaluateurs sur la base des bonnes réponses fournies par les enseignants de biologie. La concordance entre les évaluateurs était très élevée (Cronbach alpha = .98). Il a été demandé aux participants d'éviter les réponses aléatoires et d'utiliser les trois catégories de réponses proposées (Vrai/Faux/Je ne sais pas –NSP-).

4.2.2.2 Matériel dédié à l'apprentissage.

Compte tenu du nombre d'élèves inclus dans l'échantillon, la leçon RV était présentée sur un écran et était de ce fait non immersive. Cependant, la présentation multimédia vidéo était composée d'images très réalistes comprenant⁴ (i) une navigation dynamique autour des arbres de la forêt, (ii) une présentation animée des éléments des couches du sol (les éléments comme les feuilles et les plumes ont été zoomés pour l'apprenant comme s'il se penchait pour les ramasser), (iii) une immersion progressive dans les différentes couches du sous-sol ; (iv) une présentation dynamique des composantes de base de l'écosystème responsable de la décomposition des sols, en partant de la surface (feuilles, branches et brindilles, poils et plumes d'animaux) jusqu'aux couches plus profondes ; (v) des animations du processus de décomposition du sol et de la formation des couches du sol (litière, humus, matière minérale) ; (vi) une présentation dynamique du rôle de chaque animal, des bactéries aux collemboles, pseudo-scorpions et vers de terre (Figure 4.3) d'autres exemples sont disponibles en Annexe C.

⁴ Matériel développé par Stéphane Argon, LEAD, Dijon.

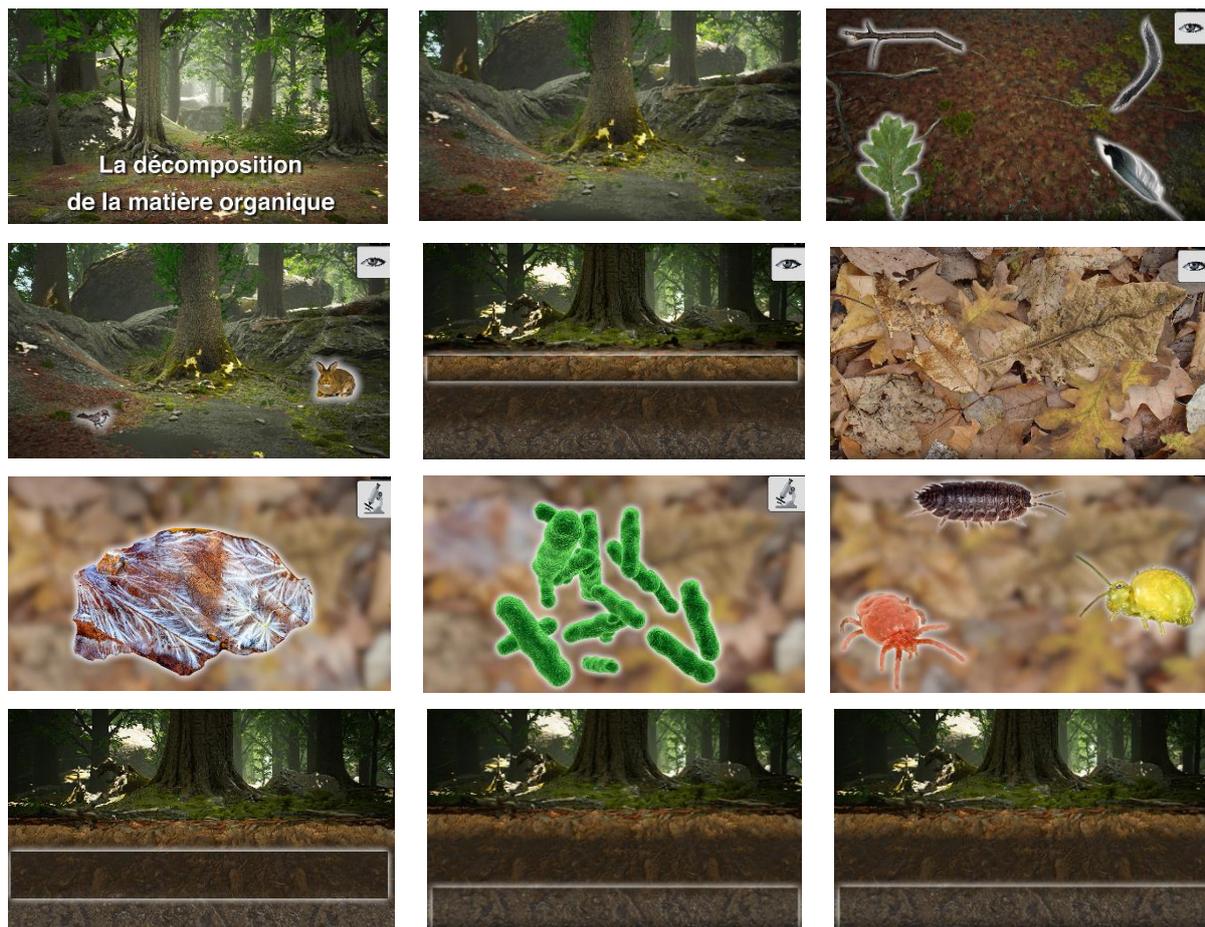


Figure 4.3. Captures d'écran de différentes étapes de la leçon.

La leçon se termine par un schéma de la chaîne alimentaire responsable la décomposition du sol et de la production de matière organique et minérale, superposée à l'image RV de la forêt (Annexe C). En somme, la vidéo proposait une alternance de navigations le long du chemin forestier, d'animations, d'images fixes montrant de très petits détails, de zoom avant, de zoom arrière et d'autres animations. De plus, dans le coin supérieur droit de l'écran, des pictogrammes évoluant selon le contexte représentaient un œil humain (Annexe D), une loupe ou un microscope pour indiquer de quelle manière il était possible d'observer les éléments présents à l'écran ce qui donnait donc une idée de la taille de l'élément en question (par exemple pour les bactéries, il s'agissait d'un microscope).

La narration (voix féminine) accompagnant la leçon de réalité virtuelle tout au long de la vidéo était entendue à l'aide d'un casque audio. Les mots clés de la narration étaient écrits à l'écran conjointement aux éléments auxquels ils correspondaient. Les titres des séquences présentées étaient également inclus dans la narration. Chaque élève disposait d'un poste informatique et d'un casque audio individuel. La narration consistait en une série de blocs composés de

quelques phrases courtes (deux à cinq en moyenne) expliquant les composants et les processus impliqués dans la décomposition du sol. Le texte parlé était composé de 30 petits blocs de phrases (1245 mots). Le texte correspondant à la narration complète est disponible en Annexe **E**. Bien entendu, sémantiquement, chaque bloc correspondait exactement à une série d'images de la vidéo (principe de cohérence-congruence, voir Mayer, 2009, chapitre 4). La leçon sous forme de vidéo multimédia durait 12 minutes et était composée de 10 séquences. Chaque séquence abordait une étape de la décomposition de la matière organique : la surface du sol, la matière organique, les acteurs de la décomposition du sol, les micro-organismes, les animaux décomposeurs, les prédateurs, le ver de terre, le résultat du travail des êtres vivants, la matière minérale et la chaîne alimentaire. Avant de commencer la leçon, l'élève visionnait les pictogrammes qui allaient être présentés dans le coin supérieur droit de l'écran accompagnés de leur fonction (Annexe **D**). Nous avons conçu deux versions de la même leçon correspondant aux deux conditions expérimentales : un mode fixe et un mode libre. Dans la version fixe, la vidéo était interactive mais la diffusion des séquences thématiques était automatique et imposée (Figure 4.4). La vidéo suivait son cours automatiquement et s'arrêtait pour indiquer à l'élève de cliquer sur un élément particulier pour déclencher sa description. Aucune autre action n'était possible et la leçon ne pouvait être visionnée qu'une seule fois.

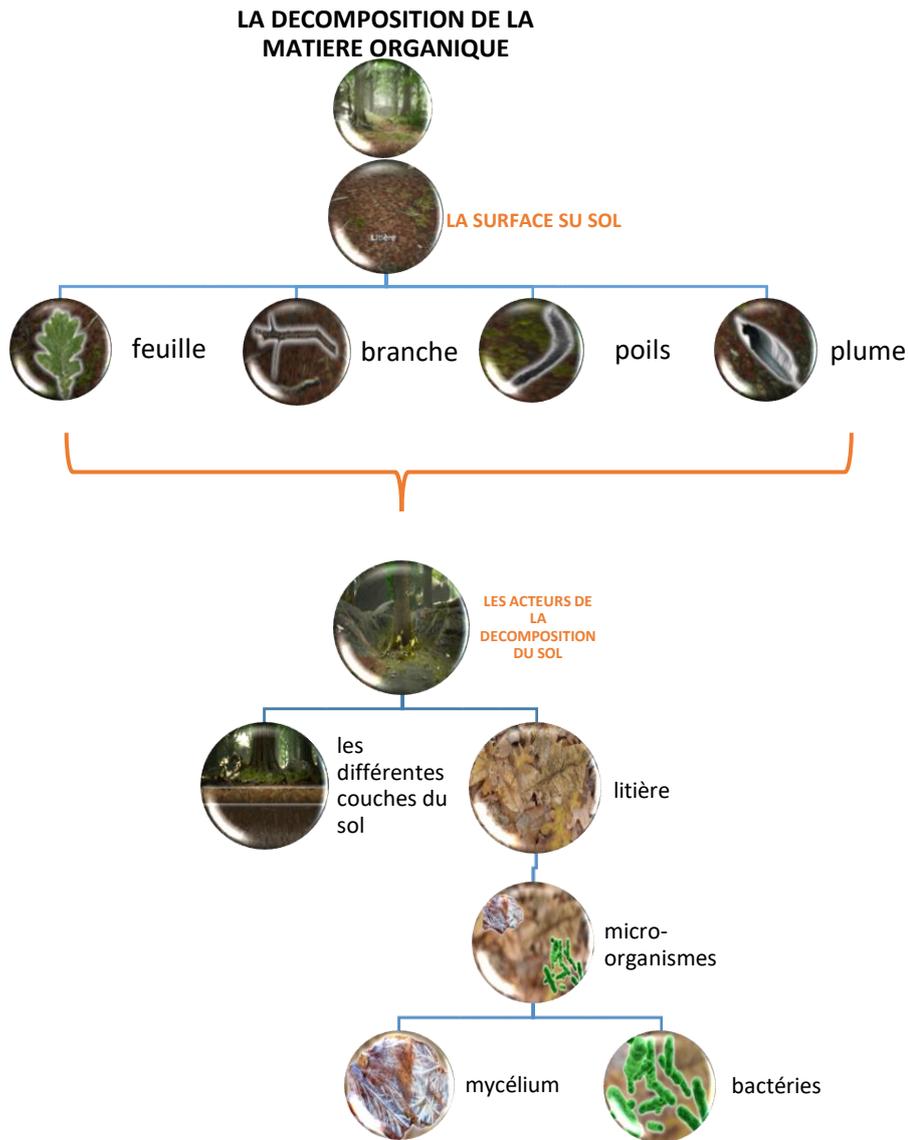


Figure 4.4. Schématisation du mode Fixe, suivant un ordre chronologique cohérent de manière automatique.

La Figure 4.4 représente un exemple de l'enchaînement de deux séquences de la leçon. Chaque séquence était composée d'éléments organisés hiérarchiquement. Elles débutaient par un écran sur lequel l'intitulé de la séquence était annoncé ; par exemple, la séquence 1 commençait par un écran où était indiqué en son centre « *La surface sur sol* ». Le thème abordé était ensuite introduit : « *Qu'y a-t-il sous tes pieds ? [Animation mimant l'abaissement du regard] ce qu'on voit ici, c'est ce que l'on appelle la litière. Elle est composée d'éléments que tu connais bien ! clique dessus pour en savoir plus* ». Le groupe d'éléments feuille, plume, poils, branche, apparaissait à l'écran et l'élève avait pour consigne de cliquer sur ceux-ci dans le but d'obtenir plus d'informations. Lorsqu'un élément était sélectionné, il était zoomé au centre de l'écran puis décrit. Lorsque tous les éléments avaient été consultés, la leçon passait automatiquement

à la séquence 2. De même, cette séquence débutait par un écran indiquant le titre de la thématique abordée, ici « *les acteurs de la décomposition du sol* » suivi d'une courte introduction sur le thème abordé, permettant de présenter et de situer les éléments dans la couche du sol à laquelle ils correspondaient. Ensuite, les bactéries et les champignons apparaissaient à l'écran, et l'élève devait les sélectionner un à un afin d'obtenir leur description respective. Les dix séquences s'enchaînaient ainsi dans l'ordre donné jusqu'à la fin de la leçon. Dans le mode libre, après la visualisation des pictogrammes, l'élève voyait apparaître à l'écran le titre des différentes séquences disposés aléatoirement : les titres étaient présentés selon une disposition aléatoire différente pour chaque élève (Figure 4.5)



Figure 4.5. Écran présenté dans la condition libre.

L'élève devait alors cliquer sur l'intitulé de son choix pour lancer la diffusion de la séquence correspondante. Une fois la séquence terminée, l'écran proposant les intitulés était de nouveau affiché. Chaque séquence ne pouvait être visionnée qu'une seule fois et la leçon était terminée une fois que l'élève avait visionné toutes les séquences. Dans les deux conditions, les élèves étudiaient donc exactement les mêmes éléments.

4.2.3 Procédure.

Lors de la première session, les participants ont réalisé les trois pré-tests : le test d'aptitudes spatiales, le test de mémoire de travail verbale et le pré-test de connaissances préalables. Ils ont également complété un questionnaire démographique comprenant des questions sur leur âge, leur sexe, l'utilisation des ordinateurs et des outils numériques tels que les réseaux sociaux et les jeux vidéo. Nous leur avons également demandé d'indiquer s'ils portaient des lunettes et s'ils avaient connaissances de difficultés d'apprentissage. Les tests

étaient effectués en salle de classe ce qui représente des groupes d'environ 25 élèves à qui on distribuait des livrets individuels spécifiques aux tâches à réaliser. Pour chaque groupe, la séance était encadrée par trois expérimentateurs qui indiquaient explicitement aux élèves qu'ils devaient faire abstraction de cette première séance et qu'ils ne devaient pas communiquer avec les autres à propos des tâches réalisées après avoir quitté la salle.

La deuxième session a eu lieu une semaine après la première séance et a également duré 60 minutes. Cette fois, les élèves étaient accueillis par groupes de 14 dans la salle informatique de l'établissement. Ils ont été placés individuellement devant un ordinateur (15 pouces) équipé d'écouteurs. Le bon fonctionnement de tous les équipements a été vérifié, puis les élèves ont reçu les instructions générales sur la tâche d'apprentissage. Ils ont ensuite été invités à apprendre du mieux qu'ils pouvaient la leçon qui leur serait présentée sur l'ordinateur. On leur a également indiqué qu'ils devraient répondre à des questions sur la leçon après la phase d'apprentissage. La leçon RV durait 12 minutes dans la modalité fixe, et pouvait durer quelques minutes de plus dans la modalité libre en fonction du temps mis par les élèves pour choisir les séquences à étudier. Immédiatement après la phase d'apprentissage, les élèves recevaient un livret individuel pour répondre aux questions du QCM portant sur des connaissances factuelles et de compréhension. Ce QCM était le même que celui proposé en prétest. Une fois le QCM complété, les élèves ayant étudié à partir de la modalité non guidée visualisaient en groupe la leçon originale. La séance se terminait par un court débriefing.

4.2.4 L'analyse des données.

Des ANOVA à un facteur et ANCOVA ont été réalisées pour analyser les données. Les deux conditions d'apprentissage (fixe et libre) constituaient un facteur inter-sujet. Dans les différentes ANOVAs, les variables dépendantes étaient respectivement (i) les performances au pré-test de connaissances préalables, (ii) les performances au post-test de rappel et de compréhension. Des analyses supplémentaires ont été effectuées pour les habiletés spatiales. Pour toutes les analyses, nous avons utilisé $p < 0.05$ comme critère de significativité. Les valeurs de l'êta partiel au carré (η_p^2) et du d de Cohen sont fournies en tant que mesures de taille d'effet pour tous les effets principaux, les interactions et les comparaisons post-hoc. Les valeurs η_p^2 de 0,01, 0,06 et 0,14 représentent respectivement des tailles d'effet petite, modérée et grande (Clark-Carter, 1997), et les valeurs d de 0,20, 0,50, 0,80 et 1,3 représentent respectivement les tailles d'effet petite, moyenne, grande et très grande.

4.3 Résultats.

Nous présenterons dans un premier temps les résultats du test de connaissances factuelles et de compréhension, puis les performances au test d'habiletés spatiales.

Homogénéité des groupes expérimentaux.

Les performances des élèves au questionnaire de connaissances préalables et au test d'habiletés spatiales sont présentées Tableau 4.1.

Tableau 4.1

Scores obtenus, moyennes et écarts types, au test d'habiletés spatiales en fonction de chaque condition.

	Score au test d'habiletés spatiales (%)
Présentation cohérente fixe	65.05 (22.47)
Présentation libre	65.20 (21.90)

Afin de de vérifier l'homogénéité des groupes fixe et libre, des ANOVAs ont été effectuées, avec la condition d'apprentissage comme facteur inter-sujet, et les scores au QCM de prétest et au test d'habiletés spatiales, respectivement, en tant que variables dépendantes.

Le test d'homogénéité ($p = 0.47$) et l'ANOVA ne montrent pas de différence significative des scores au QCM entre les deux groupes expérimentaux établis $F(1, 102) = .44$, $p = .51$, $\eta_p^2 = 0.004$. Il en est de même concernant le test spatial dont le test d'homogénéité ($p = 0.52$) et l'ANOVA ne montrent pas de différence significative entre les deux groupes : $F(1, 102) = .09$, $p = .76$, $\eta_p^2 = 0.0076$.

Performances au QCM.

Le QCM était composé de 36 questions auxquelles nous avons attribué 1 point par réponse juste et 0.5 points par réponse fautive ou « ne sait pas ». Afin de prendre en compte les erreurs et les réponses aléatoires potentielles, nous avons conçu un indice de ratio de

performances qui inclut toutes les catégories de réponses avec la formule suivante : *réponses correctes* / ((*erreurs**0.5 + *NSP**0.5)). Un participant qui avait 23 bonnes réponses, 10 erreurs et 7 réponses « Je ne sais pas » a été noté : 23/ (10*0.5+7*0.5) = 2.71. Les performances des élèves au QCM lors du prétest et du posttest sont présentés dans le tableau 4.2.

Tableau 4.2

Scores obtenus, moyennes et écarts types au QCM pour le prétest et le posttest en fonction de chaque condition.

	Score au prétest ratio	Score au prétest %	Score au posttest ratio	Score au posttest %
Présentation cohérente fixe	1.42 (1.06)	36.20 (16.00)	3.53 (2.94)	55.65 (18.22)
Présentation libre	1.52 (0.96)	40.00 (13.00)	2.82 (2.41)	52.55 (14.44)

L'ANOVA à mesures répétées, incluant un facteur intragroupe prétest-posttest et un facteur intergroupe, type de présentation, respectivement présentation cohérente fixe vs. présentation libre, a montré une augmentation significative des performances d'apprentissage et de compréhension du pré-test au post-test $F(1,102) = 42.51, p < 0.001, \eta_p^2 = 0.29$. S'il n'y a pas d'effet principal (incluant donc les performances des pré et post-test) du type de présentation, l'interaction entre pré et post-test et le type de présentation est significative: $F(1,102) = 3.98, p = 0.048, \eta_p^2 = 0.04$. (Figure 4.6). Une ANOVA à un facteur indépendant, type de présentation, réalisée sur le score de différence entre le pré-test et le post-test révèle un effet significatif du type de présentation, $F(1,102) = 3.98, p = 0.048, \eta_p^2 = 0.04$.

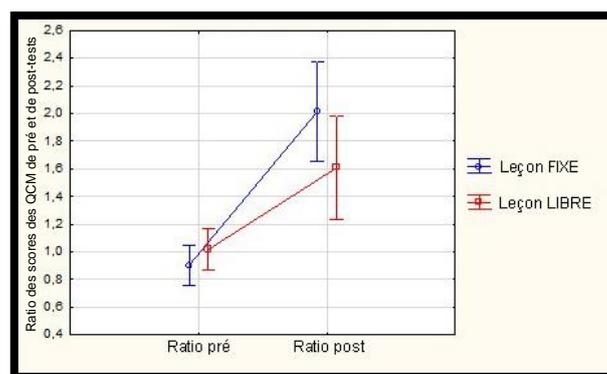


Figure 4.6. Ratio des performances des élèves au prétest et au post-test de connaissances factuelles selon le type d'exploration de la leçon : fixe ou libre.

Apprentissage et habiletés spatiales.

Nous avons réalisé une ANCOVA incluant le facteur intergroupe type de présentation et le score au test d'habiletés spatiales comme facteur covariant, sur la différence des scores entre le prétest et le post-test.

L'ANCOVA réalisée sur la différence de score entre le prétest et le posttest a révélé un effet significatif des habiletés spatiales : $F(1,101) = 7.91, p = <0.006, \eta_p^2 = 0.7$, et aussi une conservation de l'effet du type de présentation en faveur de la modalité cohérente fixe, $F(1, 101) = 4.06, p = .046, \eta_p^2 = 0.04$.

En résumé, les performances des élèves ayant bénéficié d'une leçon cohérente fixe sont supérieures aux performances des élèves ayant bénéficié d'une leçon avec une présentation libre. Les résultats indiquent également que les habiletés spatiales influencent la compréhension des apprenants. Notre première prédiction, **H1**, évoquant un effet principal du type de leçon sur les performances d'apprentissage est confirmée. Les élèves ayant étudié la leçon cohérente fixe ont obtenu des performances plus élevées que les élèves ayant exploré les différentes séquences de la leçon dans l'ordre qu'il souhaitaient. Aussi, notre seconde prédiction, **H2**, évoquant un effet des habiletés spatiales sur les performances des élèves est validée. Les apprenants ayant un niveau d'habiletés spatiales plus important ont également obtenu des résultats supérieurs au questionnaire de connaissances factuelles lors du posttest.

Discussion et conclusion.

L'objectif de cette étude était de tester le principe de guidage par la cohérence sémantique, appliqué à une leçon en réalité virtuelle non immersive. Que le contenu proposé soit statique, dynamique ou en réalité virtuelle, les informations essentielles sont sélectionnées et reliées de façon pertinentes par rapport aux notions à apprendre (Sweller, Ayres, Kaluyga & Chandler, 2003). La cohérence est testée par l'ajout d'informations au dispositif d'apprentissage (Mayer & al. 1996 ; Harp & Mayer, 1997 ; Lehman & al. 2007) et/ ou par l'ajout de repères représentationnels externes (Stull & Mayer, 2007 ; Boucheix & Lowe, 2010 ; Albus & al., 2020). Cependant, le rôle de l'organisation de la chaîne causale des informations en elle-même n'a pas été étudiée dans le domaine des apprentissages multimédia qu'il soit question d'animations ou de la réalité virtuelle. Cette étude a tenté de répondre à plusieurs questions sur le rôle de l'organisation de la chaîne causale des informations dans le cadre de l'apprentissage d'une leçon en réalité virtuelle non immersive sur le thème de la décomposition

de la matière organique. Ce thème met en jeu de nombreuses informations interdépendantes dont les relations doivent être intégrées hiérarchiquement pour être comprises et apprises. Nous avons conçu une leçon composée de séquences distinctes que nous avons pu isoler pour créer nos deux modalités de présentation : fixe ou libre. Le contenu de chaque séquence respectait la théorie de l'apprentissage multimédia (Mayer,2009) : la cohérence, la signalisation et la contiguïté spatiale et temporelle. Les résultats que nous avons obtenus sont en faveur d'un bénéfice de la réalité virtuelle pour les apprentissages. En effet, nous avons d'abord observé un gain de connaissances pour les deux types de présentation de la leçon (fixe et libre). Même lorsque l'élève étudiait la leçon dans la modalité libre, le fait que les informations aient été présentées de manière à ne pas surcharger les ressources cognitives au sein même des séquences permettait de créer des « chunks » d'informations. La difficulté résidait dans le fait de relier ces différents « chunks » de manière à créer une chaîne causale cohérente (Lowe & Boucheix, 2008). Cette interprétation explique en partie le fait que nous n'ayons pas observé un écart de performances très important entre les élèves ayant bénéficié d'une leçon fixe dont la chaîne causale était organisée et cohérente ou libre. De plus, dans la présentation libre, le fait que les apprenants puissent choisir l'ordre d'exploration des différentes séquences a pu engendrer un engagement plus important de leur part dans la tâche d'apprentissage. Cependant, pour mieux comprendre le type d'exploration réalisé par les apprenants de ce groupe (libre), il aurait été utile d'enregistrer l'ordre d'exploration des différentes séquences sur lesquelles ils pouvaient cliquer. Nous aurions également pu proposer une diffusion aléatoire des différentes séquences dans la condition « libre ». Ainsi, nous aurions pu mesurer plus finement le rôle de la cohérence sémantique entre les deux types d'exploration. En outre, les élèves auprès desquels nous avons travaillé possédaient des connaissances préalables sur le processus étudié. En effet, ils avaient étudié le thème de la décomposition de la matière organique l'année précédente ce qui explique le taux de connaissances préalables assez élevé au prétest (respectivement 36.20% pour la condition cohérente fixe et 40 % pour la condition libre) ce qui explique le faible écart de gain de connaissances entre les deux conditions expérimentales. Si, le gain de connaissances est loin d'atteindre les 100%, (il n'y a pas d'effet plafond) on peut se demander dans quelle mesure les résultats que nous avons obtenus sont dus à la complexité des concepts à apprendre. La RV nous a permis de rendre visible un ensemble d'éléments invisibles dans les couches du sol forestier, et que les enseignants ne pouvaient pas forcément montrer de manière contextualisée en classe. Notre leçon permettait de resituer les différents être vivants et micro-organismes dans le sol, tout gardant la notion de leur taille à l'échelle de la forêt. Sachant que le phénomène de décomposition de la matière organique était déjà complexe à intégrer pour les enfants dans une

version simplifiée, il est possible que ces allers retours entre le sol et les différents éléments qui le composent aient provoqué une forte demande en termes de ressources cognitives.

Par ailleurs, les résultats que nous avons obtenus ont montré que les habiletés spatiales des élèves jouaient un rôle important dans les performances de compréhension de la leçon en RV (H2). Il semblerait que les apprenants ayant de faibles habiletés visuospatiales comprendraient moins facilement le processus exposé du fait de la complexité spatio-temporelle induite par l'environnement en RV. La leçon que nous avons proposée expose de nombreux éléments entretenant des relations complexes les uns avec les autres, et ce, pendant une durée de 12 minutes ce qui peut être long pour les élèves. Jusqu'alors, les études menées sur la cohérence sémantique proposaient des explorations assez courtes, comme par exemple l'étude de Liu, Wang, Lei, Wang & Ren (2020) qui était d'une durée de 4 à 5 minutes et dont les résultats ont montré une supériorité du gain de connaissances en RV. A l'inverse, Grassini, Laumann, & Skogstad (2020) qui proposaient une exploration en RV de 10 minutes n'obtenaient pas de différences de performances entre la version RV et une version vidéo 2D. Ces deux études n'ont cependant pas respecté une méthodologie expérimentale rigoureuse, il serait donc nécessaire de tester la durée d'exploration d'une leçon en RV sur les performances de compréhension et d'apprentissage des apprenants. L'étude de Parong & Mayer (2020) ou encore celle de Krassmann, Melo, Peixoto, Pinto, Bessa & Bercht (2020) n'ont pas montré de différences de performances au test de rétention d'informations entre un apprentissage en RV de 12min et un apprentissage de 12 minutes à partir d'un Powerpoint ou d'une vidéo 2D. On peut donc supposer que la durée d'exposition à des dispositifs audiovisuels complexes pourrait donner lieu à une surcharge des ressources cognitives, notamment de la mémoire de travail face à un trop grand nombre d'informations visuelles. Dans notre étude, la vidéo RV durait 12 minutes et était composée de 10 séquences évoquant chacune 3 à 4 éléments différents en moyenne. L'élève devait alors relier les différents éléments entre eux et de surcroît, les mettre en relation avec les couches du sol auxquelles ils appartenaient afin comprendre l'intégralité processus. On peut donc penser à une surcharge du calepin visuo-spatial en mémoire de travail face à un environnement en mouvement. Dans le but de tester l'impact de la durée d'exposition aux environnements multimédias en réalité virtuelle sur les apprentissages, il serait intéressant de présenter les séquences de la leçon une à une, de manière ordonnée et espacée dans le temps. Il est possible que le fait de présenter les séquences une par une laisse le temps à l'apprenant d'intégrer l'information de manière à créer une base de connaissances solide permettant, au fil du temps, de relier les différentes séquences entre elles pour créer un modèle mental cohérent.

Il est tout à fait possible que la RV puisse être bénéfique dans l'apprentissage à condition de n'être utilisée que sur de courts laps de temps.

Conclusion.

Dans cette étude, l'utilisation d'un matériel riche et non immersif a montré que même lorsque les élèves avaient des connaissances préalables, une présentation fixe, hiérarchique de l'information était préférable à une exploration libre pour favoriser l'acquisition de connaissances. Des recherches supplémentaires seraient nécessaires pour obtenir une meilleure compréhension des mécanismes cognitifs sous-jacents à l'acquisition, notamment en RV immersive. Les résultats que nous avons obtenus nous ont mené à penser que l'utilisation de dispositifs d'apprentissage en réalité virtuelle pourrait permettre un apprentissage efficace dans le cadre d'une présentation hiérarchisée et cohérente. Il serait également intéressant, dans une perspective développementale, de comparer l'impact de la RV sur l'apprentissage des enfants comparativement aux adultes afin de déterminer si ce format fait appel à la même quantité de ressources cognitives en fonction de l'âge des apprenants.

**Chapitre 5 - La contiguïté temporelle en
réalité virtuelle : Effets des latences
temporelles contrastées narration-animation
et différences individuelles dans
l'apprentissage d'un écosystème biologique au
collège.⁵**

⁵ Ce chapitre a fait l'objet d'une soumission en vue d'une publication dans une revue internationale.

5.1 Introduction

Le but de cette recherche était de tester et d'étendre nos connaissances du principe de contiguïté temporelle dans l'apprentissage multimédia (Mayer & Anderson, 1991, 1992 ; Mayer & Sims, 1994 ; Mayer, Moreno, Boire & Vagge, 1999 ; Michas & Berry, 2000 ; Mayer, 2009, 2014, Fiorella & Mayer, 2017) dans un environnement en réalité virtuelle riche et dans le contexte de l'apprentissage en collège. Dans le cadre de la théorie cognitive de l'apprentissage multimédia (CTML) (Mayer, 2009, 2014), le principe de contiguïté temporelle stipule que les élèves apprennent mieux lorsque les mots et les images d'une leçon multimédia sont présentés de manière synchronisée plutôt que lorsque les mots précèdent les images, ou lorsque les images précèdent les mots. La présente étude, menée auprès de collégiens de 12 ans, s'inscrit dans un projet plus large, qui vise à créer un simulateur forestier interactif en RV immersive, pour l'étude des écosystèmes forestiers dans les collèges. La correspondance et la synchronisation de la narration et de l'animation peuvent être d'une importance cruciale en RV immersive interactive, en particulier parce que pendant la navigation et l'exploration, les mouvements de la tête dirigent la dynamique visuelle par le regard (Makransky & Petersen, 2021). Lorsque des mots sont prononcés en même temps que la visualisation dynamique en RV (Baceviciute, Mottelson, Terkildsen et Makransky, 2020), des effets de transience peuvent se produire dans les canaux sensoriels verbaux et visuels. Par conséquent, l'intégration des deux sources d'information dans l'élaboration d'un modèle mental cohérent pourrait avoir un impact du fait de la limitation des ressources en mémoire de travail (voir par exemple le Time-Based Resource-Sharing Model de Barrouillet, Bernardin, & Camos, 2004).

Des études antérieures sur la contiguïté spatiale (le split attention effect, voir Moreno & Mayer, 1998, et la récente méta-analyse de Schroeder & Cenki, 2018), et sur la contiguïté temporelle (Mayer, 2009, 2014 ; Ginns, 2006) font référence à la nécessité d'intégrer des sources d'information disparates. Cependant, la méta-analyse menée par Ginns (2006) comprenait 37 études comparatives sur la contiguïté spatiale mais « seulement » 13 études comparatives sur la contiguïté temporelle. Il est clair que les recherches portant sur la contiguïté temporelle sont moins nombreuses que celles portant sur l'attention spatiale divisée. Plusieurs études plus récentes, telles que Owens & Sweller, 2008 ; Schüler, Scheiter, Rummer & Gerjets, 2012 ; Meyerhoff et Huff, 2016 ; Arndt, Schüler et Scheiter, 2019 ; Xie, Mayer, Wang et Zhou, 2018 ; Khacharem, Trabelsi, Engel, Sperlich & Kalyuga, 2020, ont examiné la contiguïté temporelle en utilisant différents contenus d'apprentissage, dans différents domaines. Cependant, leurs résultats semblent mitigés (voir ci-après). Que savons-nous de

l'effet de contiguïté temporelle dans l'apprentissage multimédia ? Un certain nombre de travaux de recherche ont montré un effet significatif de la contiguïté temporelle sur les tâches d'apprentissage et de transfert, tandis que d'autres n'ont trouvé aucun effet ou ont de sérieuses limitations. Au vu de ces résultats, nous avons choisi d'évaluer le principe de contiguïté temporelle en classe, dans un contexte plus écologique que dans les recherches précédentes, afin de déterminer les limites de ce principe dans des conditions plus « authentiques ».

5.1.1 Quand et pourquoi l'effet de contiguïté temporelle fonctionne-t-il, et quels sont les processus cognitifs sous-jacents ?

Dans une série d'études de Mayer & Anderson (1991), les participants devaient apprendre comment fonctionnait une pompe, à partir d'une animation accompagnée d'une narration. La narration a été présentée soit simultanément à l'animation, soit de manière séquentielle (avant ou après l'animation). Dans la condition simultanée, par exemple, lorsque le narrateur a dit « *La soupape d'admission s'ouvre* », l'animation montrait l'ouverture de la soupape d'admission, voir Mayer & Fiorella, 2014, p. 304. Les résultats indiquent que les apprenants de la condition simultanée ont surpassé les apprenants des conditions séquentielles successives dans les tests d'apprentissage et de transfert. L'effet de contiguïté temporelle a été reproduit avec succès dans plusieurs autres études (Mayer et Sims, 1994; Mayer, Moreno, Boire et Vagge, 1999; Michas et Berry, 2000). L'explication sous-jacente était que la liaison et l'intégration cognitive d'un modèle mental verbal et d'un modèle mental pictural (Mayer, 2014, Schnotz, 1999) étaient facilitées dans la condition simultanée. La méta-analyse de Ginn's (2006), qui comprenait 13 comparaisons expérimentales, a confirmé l'avantage de la contiguïté temporelle avec une taille d'effet de $d = 0,87$.

Dans une étude pionnière, Baggett, 1984 (voir aussi, Baggett & Ehrenfeucht, 1983) a demandé à des élèves d'apprendre à partir d'un film qui présentait un kit de montage, ses différentes pièces, leurs noms et certaines de leurs utilisations. L'étude a comparé différentes latences temporelles entre les informations verbales et les illustrations (par exemple 7, 14 ou 21 secondes avant ou après le film, et 0 seconde (synchronie)) et a montré qu'un délai de plus de sept secondes entre une explication narrative et une image était préjudiciable en tâche de rappel immédiat et différé.

Récemment, Xie, Mayer, Wang & Zhou (2018) ont étudié les effets de la coordination des signaux visuels et auditifs dans une présentation multimédia composée d'un schéma contenant une courte animation ainsi qu'une narration décrivant le processus de transmission

neuronale. Dans l'expérience 3, les auteurs ont comparé un groupe « signal visuel-avant-auditif » à un groupe « signal visuel-après-auditif » ; un groupe « synchronisé » et un groupe « sans indice » ont également été inclus. La latence temporelle de présentation entre les indices visuels (code couleur) et les indices auditifs (mots sélectionnés prononcés plus fort) était de 3 secondes. Par exemple, dans la leçon où le signal visuel était présenté avant le signal auditif, lorsque le narrateur prononçait les mots clés « *potentiels d'action* » plus fort, l'élément correspondant dans le schéma avait été signalé visuellement 3 secondes auparavant (Xie & al., p. 13-14). La leçon dans laquelle l'information était présentée dans la modalité visuel-après-auditif, inversement, les éléments picturaux ont été signalés visuellement 3 secondes après que les mots clés correspondants aient été prononcés. Les résultats ont montré que les apprenants qui ont reçu des signaux doubles coordonnés ont obtenu de meilleurs résultats aux tests de rétention et de transfert que les apprenants qui n'ont reçu aucun signal et que les apprenants ayant reçu les signaux doubles successivement. Fait intéressant, du point de vue du traitement cognitif, les mesures de suivi oculaire ont révélé que les élèves ayant reçu des signaux doubles coordonnés ont regardé plus rapidement la zone pertinente et ont passé plus de temps à regarder cette zone pertinente, que les étudiants n'ayant reçu aucun signal ou que les étudiants ayant reçu les signaux visuels après les signaux auditifs. Ainsi, Xie, Mayer & al. (2018) ont réduit le laps de temps déterminé par Baggett (1984) de 7 secondes à 3 secondes pour les apprenants traitant du matériel complexe.

5.1.2 Limites et résultats mitigés dans l'effet de contiguïté temporelle : quand et pourquoi ça ne fonctionne pas ?

Malgré cet ensemble de preuves en faveur de l'effet de contiguïté temporelle, une autre série d'études n'a montré aucun effet de la présentation successive ou simultanée d'informations verbales et picturales, et n'a trouvé aucune condition limite importante aux effets de contiguïté temporelle.

Premièrement, lorsque la présentation est auto-rythmée (vs. Rythmée par le système), l'effet de contiguïté temporelle semble être moins prononcé (Michas & Berry, 2000; Mayer & Fiorella, 2014). Deuxièmement, lorsque la présentation séquentielle successive d'informations verbales et picturales est composée d'une série de petits morceaux d'informations qui ne dépassent pas les limites de la mémoire de travail, plutôt qu'un long texte (parlé ou écrit) présenté avant ou après une longue animation, un certain nombre d'études n'ont montré aucun effet de contiguïté temporelle (Mayer & al. 1999; Moreno & Mayer, 1999; Schüler, Scheiter, Rummer & Gerjets,

2012). Dans cette dernière étude, les participants ont étudié le processus de développement des tornades, et les résultats ont montré que les présentations simultanées n'étaient pas supérieures aux présentations séquentielles. Ce résultat contredit l'effet de contiguïté temporelle. Selon ces auteurs, lorsque les conditions de présentation du texte écrit et de l'image étaient simultanées, les apprenants semblaient s'être davantage concentrés sur le texte et ont négligé les images, ce qui se traduisait par une mémorisation picturale plus faible. Dans une expérience très récente, Khacharem, Trabelsi, Engel Sperlich & Kalyuga (2020), ont étudié si la longueur des segments d'apprentissage et le niveau d'expertise des apprenants (novices *vs* experts) influençait l'occurrence de l'effet de contiguïté temporelle. L'étude a été menée sur le thème du sport avec du matériel vidéo montrant des stratégies d'attaque de basket-ball. Les résultats ont montré que les novices ont plus appris de la présentation simultanée (audio et vidéo synchronisés) que des présentations successives. Dans les présentations en macro-étapes, les novices ont mieux appris lorsque l'audio était précédé par le segment de clip vidéo. Les experts ont obtenu de meilleurs résultats dans la condition de présentation en micro-étapes. (Effet d'inversion de l'expertise).

Troisièmement, et comme on peut le voir dans la méta-analyse de Ginns (2006), plusieurs études n'ont trouvé aucun effet de contiguïté temporelle lorsque le matériel à mémoriser et à comprendre était trop simple. Arndt, Schüler et Scheiter (2019) ont utilisé des séries d'articles composées de quelques phrases simples et d'images uniques dans leur étude. Celles-ci étaient présentées simultanément ou séquentiellement, l'image précédant ou suivant la phrase correspondante. Les résultats ont montré que l'intégration texte-image n'était pas influencée par la contiguïté temporelle. Meyerhoff et Huff (2016) ont utilisé de brefs extraits de films audio-vidéo (films hollywoodiens), à partir desquels ils ont présenté uniquement les pistes sonores (audio), uniquement la piste vidéo (visuel) ou les pistes son et vidéo ensemble (audio-visuel). Ils n'ont trouvé aucun effet de la synchronie temporelle, mais ont observé un effet de congruence sémantique entre les mots et les images sur la construction d'un modèle mental approprié.

En somme, les résultats de la recherche sur l'effet de contiguïté temporelle sont mitigés, et une étude plus approfondie est nécessaire pour mieux préciser et définir les limites à l'intérieur desquelles le principe de contiguïté temporelle peut être appliqué. Plus important encore, d'autres recherches pourraient se concentrer davantage sur l'analyse du traitement cognitif dans des conditions impliquant la désynchronisation du contenu verbal et pictural, en particulier dans des contextes tels que la réalité virtuelle, qui contiennent des informations riches, dynamiques et variées qui pourraient donner aux images une plus grande influence sur le texte que des documents multimédia texte-image plus simples et conventionnels.

5.1.3 Questions de recherche et hypothèses

Les images animées dans les environnements de réalité virtuelle, de même que les vidéos non immersives présentées sur écran d'ordinateur, ont beaucoup plus d'impact en termes de réalisme, de mouvements et de détails que les images utilisées dans les études précédentes (voir Figure 5.1). Nous nous sommes d'abord demandé si l'effet de contiguïté temporelle s'appliquerait toujours dans des conditions aussi réalistes.

Un examen plus approfondi des études susmentionnées montre également que le principe de contiguïté temporelle a été étudié principalement chez les adultes, les novices ou les experts (voir Khacharem & al., 2020), et beaucoup moins fréquemment chez les apprenants plus jeunes, tels que des élèves de collège (11-12 ans) ou des enfants d'école primaire. La mémoire de travail des jeunes apprenants pourrait encore se développer, et une désynchronisation temporelle entre les informations verbales et visuelles pourrait nuire à l'intégration des mots et des images correspondantes.

Les différences temporelles critiques séparant la narration de l'animation dans les études de Baggett (1984) et de Xie, Mayer & al. (2018) sont substantielles, 7secs et 3secs respectivement. Cela nous a donc conduit à émettre l'hypothèse selon laquelle avec une latence plus courte, un effet néfaste de la synchronisation serait trouvé chez des apprenants plus jeunes. Des recherches antérieures (comme la méta-analyse de Ginns, 2006) ont suggéré que l'impact de la latence entre les informations auditives et visuelles sur l'apprentissage peut dépendre de la complexité et de la longueur du type de matériel utilisé. Nous avons donc choisi de tester expérimentalement le principe de contiguïté temporelle (Mayer, 2014) dans cinq nouvelles conditions temporelles, avec une leçon animée plus longue que dans la plupart des autres études. Pour ce faire, nous avons utilisé une vidéo en RV non immersive que les élèves ont pu visualiser au sein de leur salle de classe pendant un cours de SVT.

L'analyse des procédures expérimentales utilisées dans les études précédentes suggère qu'au moins deux types de présentations temporelles désynchronisées (conditions non simultanées) ont été utilisés : (i) une *présentation séquentielle successive* des mots (texte ou narration) suivie d'images (ou animation) ou vice-versa (y compris les segments longs ou courts), *vs* (ii) une *présentation continue* de l'animation (film, animation ou présentation dynamique) accompagnée d'une narration continue qui se chevauchent (paroles) ; des conditions impliquant un laps de temps différent entre le texte et les images ont également été incluses. Nous avons supposé que ces deux paradigmes expérimentaux donneraient des résultats contrastés.

Dans la présentation séquentielle successive, le texte (texte raconté ou écrit) est généralement présenté seul, avant ou après les images (ou animation), qui sont également présentées seules. Les segments successifs peuvent être constitués de la narration complète et de l'animation complète, ou peuvent être constitués d'une série de petits extraits (pour une présentation détaillée de cette procédure, voir Mayer, 2009, p. 165, Figure 8.2).

Dans ce que nous avons qualifié de *présentation continue*, l'animation (série d'images dynamiques ou vidéo) se déroule en continu tandis que la narration se déroule en parallèle avec différentes latences possibles entre le texte et l'animation (selon les conditions expérimentales), de telle sorte que la narration peut être diffusée avant les images correspondantes, en même temps que les images correspondantes ou après les images correspondantes en continu.

Dans des recherches antérieures, les études ayant utilisé *des présentations séquentielles successives* (Mayer & Anderson, 1991, 1992 ; Mayer & Sims, 1994 ; Mayer, Moreno, Boire & Vagge, 1999 ; Michas & Berry, 2000 ; Owens & Sweller, 2008 ; Schüler, Scheiter, Rummer & Gerjets, 2012 ; Meyerhoff, & Huff, 2016 ; Arndt, Schüler, & Scheiter, 2019 ; Xie, Mayer, Wang & Zhou, 2018 ; Khacharem, Trabelsi, Engel, Sperlich & Kalyuga, 2020) sont nettement plus nombreuses que les études utilisant la présentation continue superposée de mots et d'images (Baggett, 1984). Si ces deux types de présentations temporelles semblent presque similaires pour la condition simultanée (présentation simultanée de la narration et de l'animation), ils sont très différents pour les conditions désynchronisées. Par exemple, dans les recherches de Kahcharem et al. (2019), qui utilisaient une présentation séquentielle successive, les novices apprenaient mieux lorsque l'audio précédait le segment du clip vidéo (voir aussi Xie, Mayer, Wang & Zhou, 2018), que lorsqu'une explication verbale était donnée avant la visualisation. Inversement, dans la recherche de Baggett (1984) utilisant une présentation continue (chevauchement), les participants apprenaient mieux lorsque le film précédait l'audio, que lorsque la présentation d'images pertinentes était diffusée avant l'explication.

Il apparaît que dans la présentation successive, deux facteurs se confondent : la segmentation des deux informations successives verbales et visuelles (notamment dans le cas de séries de petits extraits), et le décalage temporel entre les deux entrées sensorielles. Chaque médium, narration et animation, est traité individuellement et successivement. Par conséquent, pendant et immédiatement après la présentation d'un support, aucune autre entrée d'information n'a lieu avant la présentation du second support. De cette façon, l'apprenant peut répéter le premier ensemble d'informations, par exemple verbales, sans qu'aucun autre type

d'informations ne rivalise simultanément en mémoire de travail. Il peut alors intégrer plus efficacement ce premier ensemble d'informations dans le deuxième ensemble d'informations, par exemple l'animation (voir Ginns, 2006, p. 518). Le traitement cognitif sous-jacent de l'effet de contiguïté temporelle pourrait être différent dans le cas de la *présentation continue* de textes et d'images désynchronisés. Par exemple, lorsque la narration est diffusée avant les images correspondantes, l'apprenant entend les mots prononcés alors qu'il peut visionner des images qui ne correspondent pas réellement - ou ne correspondent pas assez précisément - aux mots. Ces mots prononcés peuvent également être non pertinents ou incompatibles par rapport aux mots prononcés qu'il traite à ce moment-là.

Une conséquence possible peut être que l'apprenant s'engage dans une activité de recherche visuelle inutile afin de trouver des informations visuelles liées aux mots prononcés qu'il vient d'entendre, ce qui augmente la charge extrinsèque. Cette activité pourrait entrer en compétition avec le processus d'intégration, et pourrait entraîner la perte de l'information narrée, des interprétations erronées potentielles (voir Baggett, 1984), ou dans le processus de déclin temporel de la mémoire de travail décrit dans le modèle de Barrouillet, Bernardin & Camos, 2004). Cette hypothèse pourrait également expliquer l'effet *d'asymétrie temporelle* décrit pour la première fois par Baggett (1984) relative aux performances des apprenants. L'effet négatif de la désynchronisation temporelle sur la compréhension s'est avéré plus important lorsque les mots prononcés étaient présentés avant l'animation correspondante que lorsque l'animation était présentée avant la narration, et aucun effet de la condition désynchronisée n'a été trouvé même lorsque le film était présenté avant la narration (Baggett, 1984). Dans ce dernier cas, la narration délivre des informations relatives à l'interprétation des images vues juste avant, facilitant ainsi l'intégration des images et de la narration. La narration peut rendre les images plus significatives. Cette situation peut être comparée à l'utilisation de casques audio par les visiteurs dans les musées d'art (voir Schwan, multimedia in the wild, 2018). Ce type d'asymétrie entre la narration avant et la narration après semble également être en lien avec le modèle de traitement multimédia texte-image récemment développé par Schnotz (Schnotz & Wagner, 2018). Des études empiriques menées dans le cadre de ce développement récent ont montré des preuves soutenant l'idée que, dans le traitement multimédia, le texte prime sur les images. De ce fait, lorsque les informations verbales sont présentées avant les images correspondantes, les apprenants peuvent être induits en erreur si les images visionnées lorsque la présentation de ces informations verbales ne sont pas utiles par rapport aux mots prononcés, ou ne leur correspondent pas suffisamment. Dans un avenir proche, une correspondance précise et verrouillée dans le temps entre le texte parlé (au lieu du texte écrit à l'écran, voir Makransky,

2018) et les images dynamiques pourrait s'avérer de la plus grande importance pour la réalité virtuelle immersive, en raison de leur fort degré d'interactivité (par exemple le contrôle par l'utilisateur de la dynamique de l'environnement).

Les conditions temporelles et la désynchronisation mot-image peuvent augmenter la charge cognitive extrinsèque dans la sélection, l'organisation et l'intégration pertinente des mots et des images correspondantes, et l'effet de contiguïté temporelle peut donc être médiatisé par les différences interindividuelles des élèves, comme les capacités de mémoire de travail (en particulier verbale) et les habiletés spatiales. Des recherches antérieures ont montré un effet significatif des habiletés spatiales sur l'apprentissage via la visualisation (voir par exemple la méta-analyse de Höffler, 2010, Castro-Alonzo, 2019). L'effet des habiletés spatiales pourrait également être prononcé chez les élèves apprenant via des présentations dynamiques en réalité virtuelle.

Enfin, nous voulions déterminer dans quelle mesure les apprenants étaient conscients de la (dé)synchronisation temporelle entre les mots et les images, et étudier leur capacité à se rappeler ultérieurement quelles étaient exactement les images présentées au moment où ils écoutaient les informations verbales diffusées simultanément. Le traitement et la compréhension impliquent la connexion et l'intégration des informations verbales et picturales pertinentes et conceptuellement correspondantes (voir les modèles théoriques de Mayer, 2001, 2014 ; et Schnotz & Bannert, 2003). Ce processus d'intégration pourrait être plus difficile dans les conditions désynchronisées que dans les conditions synchronisées, dans lesquelles les mots et images correspondantes sont affichés simultanément, sans aucun décalage temporel. De plus, si ce mécanisme de traitement d'intégration texte-image continue de fonctionner activement (plus ou moins) avec succès dans les conditions de désynchronisation mot-image, cela pourrait affecter négativement la capacité des apprenants à se rappeler de l'image qu'ils regardaient en écoutant les mots présentés simultanément, dans une sorte d'effet de bascule. Au lieu de cela, les apprenants pourraient être plus susceptibles de se souvenir correctement de la paire texte-image correspondante sémantiquement et conceptuellement que les apprenants ayant vu la condition de présentation mot-image synchronisée. Cependant, nous nous attendions à ce que dans les conditions désynchronisées, les apprenants se souviennent d'une plus grande proportion d'images lorsque les images qu'ils voient correspondent exactement à la narration entendue simultanément que dans les autres conditions désynchronisées. Cependant, en raison de la mise en œuvre du processus d'intégration texte-image, nous avons pensé que tous les apprenants de toutes les conditions, synchronisées et désynchronisées, auraient tendance à se souvenir plus volontiers des correspondances image-texte pertinentes sur le plan sémantique et

conceptuel, tout comme pour la présentation synchronisée. La question de savoir si cette tendance est liée ou non à la capacité de mémoire de travail des apprenants (ici la capacité de la mémoire de travail verbale) et, potentiellement, aux habiletés spatiales des apprenants est également intéressante.

Sur la base de l'analyse ci-dessus, la présente étude a été conçue dans le but d'étudier les effets de cinq nouvelles latences temporelles, différentes de celles utilisées dans les études précédentes, entre la narration et les images RV dynamiques sur l'occurrence de la contiguïté temporelle, de -6, -2, 0, +2, +6 secondes respectivement. Nous avons utilisé un *cadre* expérimental de superposition que nous avons appelé ci-dessus la présentation continue, car mis à part l'expérience pionnière de Baggett, très peu d'études l'ont utilisé et les résultats de Baggett (1984) n'ont pas encore été reproduits. Dans la présentation continue superposée, l'animation (série d'images) se déroule en continu tandis que la narration se déroule en parallèle, avec différentes latences possibles entre le texte et l'animation (selon les conditions expérimentales), de manière à ce que la narration puisse être diffusée avant les images correspondantes, simultanément aux images correspondantes ou après les images correspondantes et ce, de manière continue (la diffusion étant rythmée par le système).

L'étude a été menée sur le thème de la décomposition de la matière organique, à l'aide d'une leçon de RV non immersive présentée sur un écran d'ordinateur et portant sur le mécanisme de décomposition des sols forestiers européens (et son effet sur les arbres). Ce sujet fait partie du programme scolaire français de sciences de la vie et de la terre destiné aux élèves 6^{ème} et 5^{ème} (*âgés de 12 ans*). Il a été choisi à la lumière d'une enquête pilote (entretiens avec des professeurs de sciences et de biologie) qui a révélé que ce domaine était particulièrement difficile pour les élèves de 12 ans en raison de l'introduction d'un certain nombre de nouvelles catégories biologiques et d'éléments en interaction. Des difficultés conceptuelles, telles que fournir des réponses à des questions telles que « *pourquoi la hauteur de la couche de feuilles de la forêt n'augmente-t-elle pas au fil des ans ?* » ont également été relevées par les enseignants. Le matériel RV créé a permis de montrer l'interaction « naturelle » d'organismes vivants microscopiques. Des bactéries invisibles sont devenues visibles, et les différentes couches du sous-sol forestier profond et superficiel, ainsi que la croissance des arbres sont devenues apparentes (voir Figure 5.1, ci-dessous dans la section méthode). Bien que les écarts temporels dans les conditions désynchronisées de l'expérience pilote mentionnée ci-dessus ne constituent pas une bonne pratique pédagogique, ils sont utiles d'un point de vue théorique, et

fournissent des résultats factuels qui nous ont permis (i) de donner la priorité au principe de coordination texte image dans le cadre de la conception d'outils d'apprentissage RV, et (ii) pour mieux comprendre les processus cognitifs sous-jacents au principe de contiguïté temporelle chez les jeunes apprenants, et sa relation potentielle avec la mémoire de travail verbale et les habiletés spatiales. Tous les élèves ont bien entendu reçu le matériel de la condition la plus efficace à la fin de l'expérience (étant donné que nous nous attendions à trouver qu'une condition était plus efficace que les autres).

Nous avons utilisé un paradigme expérimental (i) pré-test (ii) apprentissage (ii) post-test. Le pré-test nous a permis de contrôler les connaissances préalables de chaque élève, ainsi que leur empan de mémoire de travail verbale et leurs habiletés spatiales, respectivement (voir Castro-Alonzo & al. 2019). En plus de l'objectif principal de l'étude concernant l'effet de la contiguïté temporelle sur les résultats d'apprentissage, les mesures de la mémoire de travail et des capacités spatiales nous ont permis non seulement de vérifier l'homogénéité des groupes expérimentaux, mais aussi d'analyser l'influence potentielle de la mémoire de travail et des habiletés spatiales sur la contiguïté temporelle.

Si les conditions de décalage temporel ont une influence sur l'occurrence de l'effet de contiguïté temporelle, nous avons souhaité déterminer dans quelle mesure les apprenants avaient conscience de cet écart, ou cooccurrence, entre la narration et l'animation, et plus précisément s'ils étaient capables de reconnaître les mots qui étaient prononcés lorsqu'une image donnée (ou une série d'images) apparaissait simultanément avec l'information verbale, et vice versa (voir ci-dessus). Si, malgré le décalage entre la narration et l'animation, un fort traitement d'intégration narration-animation se développe au cours de l'apprentissage, alors les apprenants dans les conditions d'écart temporel entre la narration et l'animation pourraient ne pas être en mesure de reconnaître quels mots ont été prononcés lorsque les images ont été montrées. Au lieu de cela, dans une tâche post-test de correspondance texte-image, ils auraient tendance à choisir des paires image-texte qui correspondent exactement sémantiquement, c'est-à-dire des paires correspondant à la version intégrée texte-image, comme dans la version synchronisée de la leçon en réalité virtuelle.

Sur la base de la justification ci-dessus, quatre hypothèses principales ont été formulées concernant les performances d'apprentissage et les effets sur la mémoire de travail (verbale et spatiale).

Performances d'apprentissage

Notre première hypothèse, **H1a**, prédit un effet principal de la contiguïté temporelle sur les performances d'apprentissage et de compréhension aux pré-tests et aux post-tests, avec $0 > 2 > 6$. Nous nous attendions également à un effet d'asymétrie, **H1b** : les performances d'apprentissage seraient meilleures lorsque les images RV dynamiques sont présentées avant les mots prononcés, plutôt que lorsque les images sont présentées après les mots prononcés. Nous nous attendons également à un effet de l'empan de mémoire de travail verbale et des habiletés spatiales sur les performances d'apprentissage, **H2a**. Cet effet pourrait différer selon la longueur des écarts temporels entre les mots prononcés et les images RV, **H2b**. Enfin, une hypothèse plus ouverte et exploratoire, **H3**, a été émise sur un effet potentiel des écarts temporels sur la capacité à reconnaître, dans le post-test, les paires texte-image qui ont été présentées pendant la phase d'apprentissage. Une telle reconnaissance peut être difficile étant donné le caractère transitoire de la présentation multimédia audiovisuelle dynamique de la leçon vidéo RV.

5.2 Méthode.

5.2.1 Participants

244 élèves de classes de 5^{ème} issus de deux collèges français ont participé à l'étude. Dix classes ont été sollicitées pour cette expérimentation à raison de 5 classes par collège. Ces deux collèges étaient comparables quant à l'origine sociale des élèves. L'expérience comprenait deux sessions espacées de 10 à 14 jours (voir ci-dessous) : (i) une session de pré-test (ii) une phase d'apprentissage suivie immédiatement d'une session de post-test. En raison de l'absence de quelques élèves à l'une des deux sessions, nous avons retenu un total de 227 élèves ayant terminé les tâches des deux sessions (Âge moyen = 11,90 ans, 122 filles et 105 garçons, respectivement : 53,74 % et 46,25 %).

L'expérience s'est déroulée dans les salles de classe dédiées aux sciences au sein des collèges, lors de deux sessions distinctes. Au cours de la première session, les élèves ont répondu à une série de prétests de mémoire de travail, d'habiletés spatiales ainsi qu'à un questionnaire à choix multiple de connaissances préalables portant sur le sujet de la leçon en réalité virtuelle à savoir : la décomposition de la matière organique. Les performances des élèves à ces tests et leurs moyennes académiques, y compris leurs notes moyennes globales en biologie et dans toutes les autres matières ont été prises en compte, et utilisées pour créer cinq groupes homogènes

(avec $N = 40$ à 46 participants par groupe) correspondant aux cinq conditions de contiguïté temporelle. Dix jours à deux semaines plus tard, lors de la deuxième session, les élèves ont visualisé et étudié individuellement la leçon multimédia RV sur écran d'ordinateur (vidéo PC) sur les mécanismes biologiques de décomposition de la matière organique dans l'une de nos cinq conditions expérimentales (c'est-à-dire une condition par groupe), après laquelle ils ont effectué les post-tests (voir Figure 5.1). Chaque session a duré 60 minutes, et la durée totale globale de l'expérience était de 120 minutes. Dans chaque session expérimentale, de petits groupes de 10 à 12 élèves ont été affectés à des zones distinctes de la salle de classe. A la fin de la première session, les élèves n'ont pas été informés du contenu de la seconde session, ils ont simplement été informés qu'en seconde session ils auraient des activités différentes.

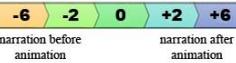
		tests																									
Session 1	Mental rotation test  <i>18 items</i>	MTVE Working memory pretest. Série de 3 phrases Essai 1 : 1) Quand je vais à l'école, je mets ma trousse dans mon (cartable) 2) Pour planter un clou, on se sert d'un (marteau) 3) Le chat miaule et le chien (aboie) <i>15 items</i>	Multiple Choice Questionnaire Pretest. <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 5%;">13</td> <td style="width: 75%;">Le ver de terre digère la matière minérale du sol.</td> <td style="width: 5%;">VRAM</td> <td style="width: 5%;">FAUX</td> <td style="width: 5%;">NSP</td> </tr> <tr> <td>14</td> <td>Le collembole, le cloporte et l'oribato fragmentent la matière organique.</td> <td>VRAM</td> <td>FAUX</td> <td>NSP</td> </tr> <tr> <td>15</td> <td>Le collembole, le cloporte et l'oribato sont les premiers à se nourrir de la matière organique.</td> <td>VRAM</td> <td>FAUX</td> <td>NSP</td> </tr> <tr> <td>16</td> <td>Les prédateurs sont des animaux qui mangent des champignons.</td> <td>VRAM</td> <td>FAUX</td> <td>NSP</td> </tr> </table> <i>36 items</i>		13	Le ver de terre digère la matière minérale du sol.	VRAM	FAUX	NSP	14	Le collembole, le cloporte et l'oribato fragmentent la matière organique.	VRAM	FAUX	NSP	15	Le collembole, le cloporte et l'oribato sont les premiers à se nourrir de la matière organique.	VRAM	FAUX	NSP	16	Les prédateurs sont des animaux qui mangent des champignons.	VRAM	FAUX	NSP			
13	Le ver de terre digère la matière minérale du sol.	VRAM	FAUX	NSP																							
14	Le collembole, le cloporte et l'oribato fragmentent la matière organique.	VRAM	FAUX	NSP																							
15	Le collembole, le cloporte et l'oribato sont les premiers à se nourrir de la matière organique.	VRAM	FAUX	NSP																							
16	Les prédateurs sont des animaux qui mangent des champignons.	VRAM	FAUX	NSP																							
Session 2	Video  time gap in second:  <i>12min</i>	Correspondence text/image test <table style="width: 100%;"> <tr> <td style="width: 50%;">« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »  <input type="checkbox"/></td> <td style="width: 50%;">« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »  <input type="checkbox"/></td> </tr> </table> <i>24 items</i>		« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »  <input type="checkbox"/>	« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »  <input type="checkbox"/>	Multiple Choice Questionnaire Post-test. <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 5%;">13</td> <td style="width: 75%;">Le ver de terre digère la matière minérale du sol.</td> <td style="width: 5%;">VRAM</td> <td style="width: 5%;">FAUX</td> <td style="width: 5%;">NSP</td> </tr> <tr> <td>14</td> <td>Le collembole, le cloporte et l'oribato fragmentent la matière organique.</td> <td>VRAM</td> <td>FAUX</td> <td>NSP</td> </tr> <tr> <td>15</td> <td>Le collembole, le cloporte et l'oribato sont les premiers à se nourrir de la matière organique.</td> <td>VRAM</td> <td>FAUX</td> <td>NSP</td> </tr> <tr> <td>16</td> <td>Les prédateurs sont des animaux qui mangent des champignons.</td> <td>VRAM</td> <td>FAUX</td> <td>NSP</td> </tr> </table> <i>36 items</i>		13	Le ver de terre digère la matière minérale du sol.	VRAM	FAUX	NSP	14	Le collembole, le cloporte et l'oribato fragmentent la matière organique.	VRAM	FAUX	NSP	15	Le collembole, le cloporte et l'oribato sont les premiers à se nourrir de la matière organique.	VRAM	FAUX	NSP	16	Les prédateurs sont des animaux qui mangent des champignons.	VRAM	FAUX	NSP
« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »  <input type="checkbox"/>	« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »  <input type="checkbox"/>																										
13	Le ver de terre digère la matière minérale du sol.	VRAM	FAUX	NSP																							
14	Le collembole, le cloporte et l'oribato fragmentent la matière organique.	VRAM	FAUX	NSP																							
15	Le collembole, le cloporte et l'oribato sont les premiers à se nourrir de la matière organique.	VRAM	FAUX	NSP																							
16	Les prédateurs sont des animaux qui mangent des champignons.	VRAM	FAUX	NSP																							

Figure 5.1. Récapitulatif des différentes tâches réalisées au cours de l'expérimentation.

5.2.2 Matériel expérimental.

5.2.2.1 Pré-test utilisés lors de la session 1.

Lors de la session de pré-test, les participants ont subi trois tests : (i) un test d'aptitude spatiale ; un test de rotation mentale, basé sur le matériel mis en œuvre par Shah & Miyake (1996) et adapté pour les enfants par Guignard & Boucheix, (2005) qui implique en partie la mémoire de travail spatiale. (ii) Un test standardisé de mémoire de travail verbale (MTVE : test de mémoire de travail verbal, adapté aux élèves de douze ans). (iii) un test de connaissances préalables sur la décomposition de la matière organique : un questionnaire à choix multiples composé de 36 questions ciblant les connaissances factuelles et la compréhension.

Le test d'aptitude spatiale comprenait 18 items présentés dans des livrets individuels (voir Figure 5.2) et était organisé en 2 subtests.

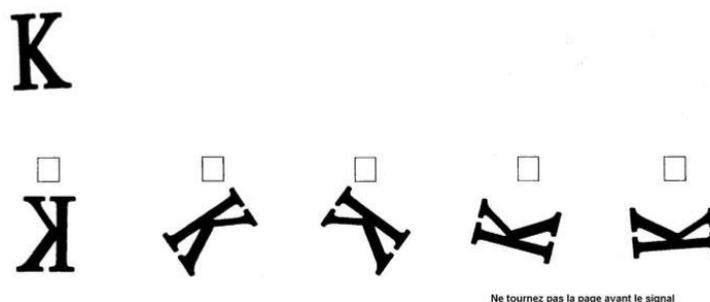


Figure 5.2. Exemple d'item du test d'habiletés spatiales.

Nous avons conçu ces subtests à partir d'une tâche de mémoire de travail spatiale de Shah et Miyaké (1996). Dans le premier subtest, une lettre était présentée à l'élève (qui devait alors garder cet item en mémoire) et il devait déterminer, à partir de 5 lettres identiques qui occupaient toutes des positions différentes dans l'espace, celles qui représentaient la première lettre présentée après une rotation ; les autres lettres étaient identiques mais présentées comme une image miroir. Dans le deuxième sous-ensemble, après la présentation de lettres, des objets communs puis des Figures simples étaient présentés. Ce test adapté avait déjà été utilisé avec succès par Guignard et Boucheix, 2005. Le nombre de bonnes réponses (sur 40) a été calculé pour chaque participant.

Le test de mémoire de travail verbale comprenait une série de 15 items présentés dans un livret de réponses pour chaque élève. L'un des tests le plus utilisé est le test de Daneman et Carpenter (1980). Dans ce test, les participants doivent écouter une série de phrases présentées oralement. A la fin de chaque série de phrases, ils doivent stocker et rappeler le dernier mot de chacune des phrases de la série. Chaque série comprend 2, 3, 4, 5 ou 6 phrases. Le participant doit alors écouter les phrases et en même temps garder le dernier mot de chaque phrase, actif en mémoire de travail. Ce test a été conçu et validé pour les adultes mais contient des phrases longues avec de nombreux mots qui pourraient être inconnus des jeunes élèves. Nous avons donc utilisé un autre test similaire au Daneman et Carpenter (1980), conçu et validé pour des élèves plus jeunes appelé : Le MTVE (Picard, Seigneuric, de Guibert, Megherbi, & Potier, 2008, une tâche de mémoire de travail verbale pour jeunes élèves (Adaptation française du Listening Span Test de Siegel et Ryan (1989)). Les mots utilisés sont connus des jeunes élèves. Les participants devaient écouter la série de phrases, et ce n'est qu'après avoir reçu le signal « go » qu'ils devaient écrire le dernier mot de chaque phrase dans le cahier de réponses. La

série la plus longue de ce test comportait 5 phrases (ce test est disponible en Annexe F). Deux scores ont été calculés à partir des réponses de l'élève : le critère le plus exigeant pour noter les réponses était le nombre de séries de phrases complètement réussies (sans erreurs ni mots manquants, sur 15), et le critère le moins exigeant était le nombre de mots correctement et complètement rappelé, dans le bon ordre, dans toutes les séries de phrases (max. 42). La corrélation entre ces deux scores était élevée, $r_{\text{Bravais-Pearson}}(227) = 0,91$. Les deux scores ont été transformés en pourcentages et nous avons utilisé la moyenne des deux scores pour les analyses de données.

Les pré-tests (et post-tests) de Connaissances factuelles et Compréhension étaient composés de 36 items (Annexe B). Les participants ont répondu à un questionnaire papier-crayon, qui était présenté comme une série d'affirmations. Pour chaque affirmation, les participants devaient choisir entre les réponses « vrai », « faux » ou « ne sait pas ». Certaines des questions posées concernaient le rappel de connaissances factuelles sur les mécanismes de décomposition du sol, tandis que d'autres faisaient appel à une compréhension plus approfondie, ce qui nécessitait la production d'inférences. Le sujet des questions était lié au contenu (mots et images) délivré par la leçon de réalité virtuelle (voir ci-après pour les détails du matériel d'apprentissage), et concernait les écosystèmes du sol forestier en décomposition produisant de la matière organique et nourrissant les arbres de la forêt. Cela comprenait : la constitution et l'évolution des couches du sol de la litière aux couches plus profondes ; le « travail » des micro-organismes, bactéries, champignons (mycélium) et animaux ; la chaîne alimentaire allant des micro-organismes aux animaux (tels que les vers de terre) et l'effet de la décomposition du sol sur la croissance des arbres.

Le type d'affirmation de connaissance factuelle était par exemple : « *Les bactéries mangent ce que les autres animaux ne peuvent pas digérer. Elles prémâchent la litière pour permettre aux petits animaux peuplant le sol de se nourrir.* » ; ou « *Sur les feuilles mortes on peut voir de petits fils blancs : ce sont des filaments de champignon. On appelle cela le mycélium* ». Un exemple d'affirmation qui nécessitait une compréhension plus approfondie du processus écosystémique était : « *Les êtres vivants et ce qu'ils rejettent sont de la matière organique* » ; ou concernant les animaux « *Ils mangent, digèrent puis rejettent la matière organique de la litière et créent ainsi une belle couche d'humus.* » ; ou encore concernant l'humus « *Cette couche de terre est extrêmement importante car elle permet de garder un sol stable et fertile. Cela veut dire qu'elle permet aux végétaux de pousser, de se nourrir.* ». Ce type d'affirmations pourrait être déduit de la série de phrases et d'images de connaissances factuelles. Le même test a été utilisé pour le

pré-test et l'un des deux post-tests. L'ordre des questions différait du pré-test au post-test et selon les participants : trois questionnaires contenaient les mêmes affirmations mais dans un ordre différent. La moitié des affirmations étaient vraies et la moitié étaient fausses. Le nombre de bonnes réponses, d'erreurs et de réponses « Je ne sais pas » (sur 36 items) ont été codés par trois évaluateurs sur la base des bonnes réponses fournies par les enseignants de biologie. La concordance entre les évaluateurs était très élevée (Cronbach alpha = .98). Il a été demandé aux participants d'éviter les réponses aléatoires et d'utiliser les trois catégories de réponses proposées (Vrai/Faux/Je ne sais pas –NSP-). Cependant, afin de prendre en compte les erreurs et les réponses aléatoires potentielles, nous avons conçu un indice de ratio de performance incluant toutes les catégories de réponses, avec la formule suivante : $Réponses\ correctes / (Erreurs * 0,50 + NSP * 0,50)$. Le nombre de bonnes réponses (sur 36) a été divisé par le nombre d'erreurs*0,50 (un demi-point était accordé aux erreurs) plus le nombre de réponses « Je ne sais pas » *0,50 (pour la réponse « Je ne sais pas » rapportait 0,50 point). Cet indice de ratio a été appelé score total pondéré (sTP).

5.2.2.2 Matériel d'apprentissage

Compte tenu du nombre d'élèves inclus dans l'échantillon la leçon RV était présentée sur un écran et était de ce fait non immersive. Cependant, la présentation multimédia des images vidéo comprenait des images très réalistes⁶ (voir une série de captures d'écran de la leçon dans la Figure 5.3) comprenant (i) une navigation dynamique autour des arbres de la forêt, (ii) une présentation animée des éléments des couches du sol (les éléments tels que les feuilles et les plumes ont été zoomés pour l'apprenant comme s'il se penchait pour les ramasser); (iii) une immersion progressive dans les différentes couches du sous-sol ; (iv) une présentation dynamique des composantes de base de l'écosystème responsable de la décomposition des sols, en partant de la surface (feuilles, branches et brindilles, champignons, poils et plumes d'animaux) jusqu'aux couches plus profondes ; (v) des animations du processus de décomposition du sol et de la formation des couches du sol (litière, humus, matière minérale) ; (vi) une présentation dynamique du rôle de chaque animal, des bactéries aux collemboles, pseudo-scorpions et vers de terre. La leçon se termine par un schéma de la chaîne alimentaire responsable la décomposition du sol et de la production de matière organique et minérale, superposée à l'image RV de la forêt. En somme, la vidéo proposait une alternance de navigations le long du chemin forestier, d'animations, d'images fixes montrant de très petits

⁶ Matériel développé et adapté par Stéphane Argon, LEAD, Dijon.

détails, de zoom avant, de zoom arrière et d'autres animations. De plus, dans le coin supérieur droit de l'écran, des symboles évoluant selon le contexte représentaient un œil humain, une loupe ou un microscope pour indiquer de quelle manière il était possible d'observer les éléments présents à l'écran, ce qui donnait donc une idée de la taille de l'élément en question (par exemple pour les bactéries, il s'agissait d'un microscope). (Figure 5.3)

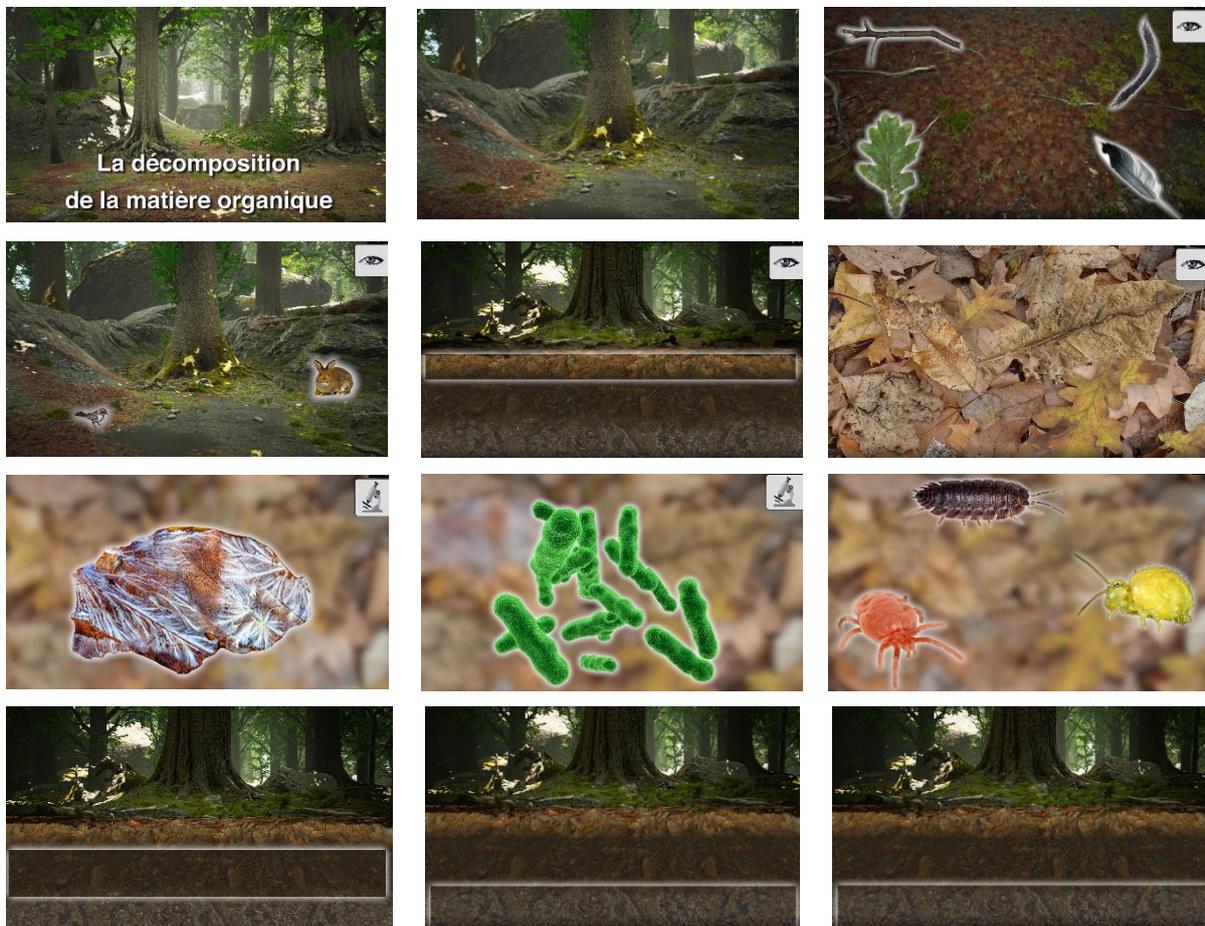


Figure 5.3. Captures d'écran issues de la leçon RV sur la décomposition de la matière organique.

La narration (voix féminine) accompagnant la leçon de réalité virtuelle tout au long de la vidéo consistait en une série de blocs composés de quelques phrases courtes (deux à cinq en moyenne) expliquant les composants et les processus impliqués dans la décomposition du sol. Le texte parlé était composé de 30 petits blocs de phrases (1245 mots). Le script de la narration est disponible en Annexe E. Bien entendu, sémantiquement, chaque bloc correspondait exactement à une série d'images de la vidéo (principe de cohérence-congruence, voir Mayer, 2009, chapitre 4, et chapitre 8, pour les conditions de contiguïté temporelle, fig. 8.2). Le vidéo multimédia durait 12 minutes et son déroulement était automatique. Nous avons conçu 5 versions de la

même vidéo, correspondant aux 5 conditions de latence temporelle établie entre le texte parlé et les images de la vidéo. Dans la condition synchronisation, il n'y avait pas d'écart temporel entre le texte parlé et les images, *condition 0 seconde*. Les quatre autres conditions étaient des présentations successives de la narration et de la vidéo. Dans deux conditions, la narration était diffusée avant l'élément visuel de la vidéo auquel il correspondait. Les écarts temporels entre le texte et l'image étaient respectivement de - 6 et - 2 secondes, (ex. : condition -6 s. et condition -2 s.). Dans deux conditions, la narration était diffusée après l'élément visuel de la vidéo auquel il correspondait ; les écarts temporels entre le texte et l'image étaient respectivement de +2 et +6 secondes, (ex. : condition +2 s et condition +6 s). Afin d'obtenir une homogénéité optimale des groupes dans chaque condition expérimentale, nous avons utilisé quatre indicateurs de performance : (i) les connaissances préalables via les scores du pré-test, (ii) le test de mémoire de travail verbale (iii) le test de capacité spatiale et enfin (iv) la note moyenne scolaire en sciences de chaque élève (biologie) depuis le début de l'année scolaire (fournie par les enseignants). Bien entendu, l'anonymat complet des élèves concernant l'utilisation des données collectées par les expérimentateurs était garanti. Sur cette base, les 227 participants ont été assignés au hasard à chacune des cinq conditions : -6, -2, 0, +2, +6.

5.2.2.3 Matériel post-tests

Lors de la session post-test réalisée après avoir visionné la vidéo, les participants devaient effectuer deux tests. Le premier test était une « nouvelle » tâche post-test dans laquelle les participants remplissaient un questionnaire de correspondance texte/image (Fig. 1, session 2, cellule du milieu). La même phrase (ou bloc de phrases) a été présentée pour chacune des 24 paires d'items de ce test. Les phrases étaient présentées sous forme écrite et étaient extraites de la narration orale. Elles étaient accompagnées de deux images statiques (captures d'écran tirées de l'animation RV, voir Figure 5.3) présentées par paires, soit côte à côte (24 paires et donc 48 images étaient présentées au total). Les participants devaient choisir, pour chaque paire, quelle était l'image de la paire qu'ils pensaient avoir vue ou se souvenaient avoir vue lorsqu'ils avaient entendu la phrase présentée pendant la leçon de RV (Test disponible en Annexe G). Un élément de chaque paire correspondait à la version texte-image synchronisée, et l'autre correspondait à une version texte-image non synchronisée qui était tirée des quatre versions texte-image désynchronisées (quatre images issues de -6, -2, + 2 et +6). Ce test fait appel à la mémoire épisodique, ce qui n'est pas habituel dans les acquis d'apprentissage multimédia. Notre hypothèse exploratoire, H3 explore dans quelle mesure les apprenants sont conscients du décalage temporel entre la narration et la vidéo. Si les apprenants traitent activement

l'intégration sémantique texte-image lors de la phase d'apprentissage, et/ou s'ils ne sont pas conscients du décalage temporel entre la narration et la vidéo dans les conditions -6, -2, +2, +6, ils pourraient être moins susceptibles de se souvenir de l'image présentée lorsqu'ils écoutaient un bloc de phrases, en particulier lorsqu'ils étaient affectés à une condition avec un écart entre la narration et la section vidéo correspondante (par exemple, -6, -2, +2, +6). Par conséquent, ils pourraient choisir beaucoup plus souvent la version synchronisée texte-image de la paire, et cela s'appliquerait également dans toutes les conditions de contiguïté temporelle : -6, -2, 0, +2, +6. À l'inverse, si les apprenants étaient capables de se rappeler qu'il y avait un décalage entre la narration et les animations, ils pourraient être plus susceptibles de choisir le couple texte-image correspondant à leur condition initiale d'apprentissage et/ou au texte-image d'une autre condition avec un écart temporel. Quatre indicateurs de performance ont été utilisés pour ce test : (i) la proportion (%) de choix des items texte-image synchronisés des paires présentées, (ii) la proportion (%) de choix des items texte-image de la condition à laquelle l'apprenant avait été affecté (iii) la proportion (%) de choix des items texte-image d'une autre condition avec un décalage temporel différent de celui auquel l'apprenant avait été affecté et (iv) la proportion de choix des éléments texte-image autres que les éléments texte-image synchronisés, ce qui correspond à la somme des indicateurs (ii) et (iii).

Pour le deuxième post-test, plus conventionnel, un questionnaire similaire au questionnaire de pré-test a été utilisé (voir ci-dessus dans la section « matériel du pré-test »). L'ordre de présentation des questions différait du pré-test au post-test, et selon les participants. La notation du post-test était la même que pour le pré-test (voir ci-dessus).

5.2.3 Procédure

La première séance (Figure 5.1) a duré 60 minutes. Les participants ont effectué les trois pré-tests : un test de mémoire de travail verbale, un test d'aptitude spatiale et le pré-test de connaissances préalables. Ils devaient également remplir un questionnaire démographique qui comprenait des questions sur leur âge, leur sexe, l'utilisation des ordinateurs et des outils numériques tels que les réseaux sociaux et les jeux vidéo. On leur a également demandé s'ils portaient des lunettes et s'ils avaient connaissance de difficultés d'apprentissage. Les tests étaient effectués en petits groupes successifs de 10 à 12 élèves répartis dans des zones distinctes d'une grande classe de sciences, à l'aide de livrets individuels et spécifiques. Pour chaque petit groupe, la séance était encadrée par trois expérimentateurs, qui indiquaient explicitement aux élèves qu'ils devaient faire abstraction de la première séance, et qu'ils ne devaient pas parler

aux autres de cette séance après avoir quitté la salle. Les élèves ont également été informés qu'ils seraient invités à une deuxième session au cours de laquelle différentes activités leur seraient proposées.

La deuxième séance a eu lieu quelques jours après la première séance et a également duré 60 minutes. Les conditions étaient les mêmes que lors de la première session (petits groupes successifs de 10-12 élèves). Les élèves ont été placés individuellement devant un ordinateur portable (15 pouces) équipé d'écouteurs. Tous les équipements ont été vérifiés pour s'assurer de leur bon fonctionnement et les élèves ont reçu des instructions générales sur la tâche d'apprentissage. Les élèves ont ensuite été invités à apprendre du mieux qu'ils pouvaient la leçon qui serait présentée sur l'ordinateur. On les a également informés qu'ils devraient répondre à des questions sur la leçon lors de deux tests après la phase d'apprentissage. Les élèves devaient regarder la vidéo RV de 12 minutes, et immédiatement après la fin de la phase d'apprentissage, ils recevaient des livrets pour répondre aux questions des deux post-tests, dans un ordre fixe : (i) la tâche de correspondance texte/image, (ii) le post-test de connaissances factuelles et de compréhension. La séance se terminait par un court débriefing.

5.2.4 L'analyse des données

Des ANOVA, des ANCOVA et des régressions linéaires ont été réalisées pour analyser les données. Les 5 conditions de contiguïté temporelle (- 6s., -2 s., 0s., +2s., +6s) constituaient le facteur inter-sujets. Dans les différentes ANOVA, les variables dépendantes étaient respectivement (i) les performances au pré-test de connaissances préalables ; (ii) les performances dans chacun des deux post-tests de rappel et de compréhension. Dans des analyses supplémentaires, les performances de la mémoire de travail verbale et des habiletés spatiales ont été utilisées comme covariants dans les ANCOVA, et des analyses de régression ont également été effectuées. Pour toutes les analyses tout au long de cette étude, nous avons utilisé $p < 0,05$ comme critère de significativité. Les valeurs de l'êta partiel au carré (η_p^2) et du d de Cohen sont fournies en tant que mesures de taille d'effet pour tous les effets principaux, les interactions et les comparaisons post-hoc. Les valeurs η_p^2 de 0.01, 0.06 et 0.14 représentent respectivement des tailles d'effet petite, modérée et grande (Clark-Carter, 1997), et les valeurs d de 0.20, 0.50, 0.80 et 1.3 représentent respectivement les tailles d'effet petite, moyenne, grande et très grande.

5.3 Résultats

Nous présenterons dans un premier temps les résultats du test de connaissances factuelles et de compréhension, ainsi que les performances aux tests de mémoire de travail verbale et d'habiletés spatiales. Nous allons ensuite décrire les résultats du test de correspondance texte-image.

5.3.1 Contrôle d'homogénéité des groupes

Les performances au test de mémoire de travail verbale et au test d'habiletés spatiales sont présentées dans le Tableau 5.1. Afin de vérifier l'homogénéité des groupes des conditions de contiguïté temporelle, des ANOVA ont été effectuées, avec la condition temporelle comme facteur inter-sujets, et les scores au test de mémoire de travail verbale et au test d'habileté spatiale, respectivement, en tant que variables dépendantes.

Comme prévu, l'ANOVA n'a montré aucun effet de la contiguïté temporelle sur les scores de mémoire au test de mémoire de travail verbale, $F(4, 222) = 0.31, p = 0.86, \eta_p^2 = 0.006$. Aucune des comparaisons post-hoc entre les conditions n'était significative. De même, l'ANOVA n'a montré aucun effet de la contiguïté temporelle sur les scores au test d'habiletés spatiales, $F(4, 222) = 0.90, p = 0.46, p^2 = 0.02$. Aucune des comparaisons post-hoc entre les conditions n'a été significative. En somme, les groupes expérimentaux ont été considérés comme homogènes en ce qui concerne les tests de mémoire de travail verbale et d'habiletés spatiales. (Tableau 5.1)

Tableau 5.1

Scores (Moyennes et SD), au test de mémoire de travail verbale et au test d'aptitude spatiale pour chaque groupe de conditions de contiguïté temporelle.

	Narration avant		Narration- Animation Synchronisée	Narration après	
	I' animation			I' animation	
	-6 s.	-2 s.	0 s.	+2 s.	+6 s.
	(n = 41)	(n=43)	(n = 40)	(n = 44)	(n = 59)
% MdT verbale (Moyenne : N mots rappelés et N essais réussis).	69.02 (10.60)	67,65 (12.08)	66.19 (10.35)	66,79 (12.10)	67,34 (13.85)
Capacité spatiale /40	19,68 (10.52)	21.23 (10.89)	23.32 (8,98)	22,77 (8.68)	21.44 (9.1)

5.3.2 Connaissances factuelles et compréhension des scores pré et post-test

Les scores de connaissances factuelles et de compréhension au pré-test et au post-test sont présentés dans le Tableau 5.2 et la Figure 5.4. Comme mentionné précédemment, afin de prendre en compte l'effet potentiel des réponses aléatoires, en particulier dans le pré-test, des ANOVA ont été effectuées sur les scores totaux pondérés (sTP, voir ci-dessus) qui incluent tous les types de réponses. Nous avons utilisé le ratio suivant : $\text{Réponses correctes} / (\text{Erreurs} * 0,50 + \text{NSP} * 0,50)$. Par exemple, un participant qui avait 10 bonnes réponses, 12 erreurs et 14 réponses « Je ne sais pas » a été noté : $10 / ((10 * 0,50) + (14 * 0,50)) = 0,83$; un participant qui avait 25 bonnes réponses, 6 erreurs et 5 réponses « Je ne sais pas » a été noté : $25 / ((6 * 0,50) + (5 * 0,050)) = 4,54$. Ce ratio évite d'obtenir des scores négatifs élevés, notamment au pré-test, avec une méthode de notation plus conventionnelle, qui consiste à soustraire le score d'erreur du score de bonne réponse est utilisée. (Voir Tableau 5.2 et Figure 5.4)

Tableau 5.2

Scores (moyennes et SD), pour les scores pré-test et post-test, y compris les pourcentages de réponses correctes (BR), de réponses fausses (MR) et de réponses « Je ne sais pas » (NSP), respectivement, et le score total pondéré (TP) qui prend en compte les trois types de réponses (BR, MR et NSP).

		Narration avant		Narration-	Narration après	
		I' animation		Animation Synchronisée	I' animation	
		-6 s.	-2 s.	0 s.	+2 s.	+6 s.
		(n = 41)	(n=43)	(n = 40)	(N = 44)	(n = 59)
Pré- test	TP	1,00 (0,79)	0,97 (0,56)	1,14 (0,75)	1,13 (0,88)	1,06 (0,71)
	% BR (/36)	29,74 (13,89)	29,91 (13,32)	33,88 (14,15)	32,20 (15,61)	31,35
	% MR	23,51 (13,65)	20,28 (13,24)	21,74 (13,55)	21,27 (13,26)	19,91 (11,71)
	% NSP	46,07 (23,68)	49,74 (21,41)	44,40 (24,08)	46,50 (24,21)	48,63 (23,57)
	Total W	2,98 (2,24)	2,82 (1,54)	4,78 (2,68)	3,68 (2,75)	3,40 (2,71)
Post- test	% BR (/36)	53,34 (14,41)	54,23 (12,40)	65,21 (14,36)	59,34 (13,23)	56,78 (14,73)
	% MR	27,84 (11,71)	24,81 (10,74)	21,80 (9,65)	23,99 (10,08)	23,91 (9,57)
	% NSP	17,81 (14,38)	20,96 (14,59)	12,99 (13,77)	16,67 (12,28)	19,30 (15,69)
	Total W	2,98 (2,24)	2,82 (1,54)	4,78 (2,68)	3,68 (2,75)	3,40 (2,71)
	% BR (/36)	53,34 (14,41)	54,23 (12,40)	65,21 (14,36)	59,34 (13,23)	56,78 (14,73)

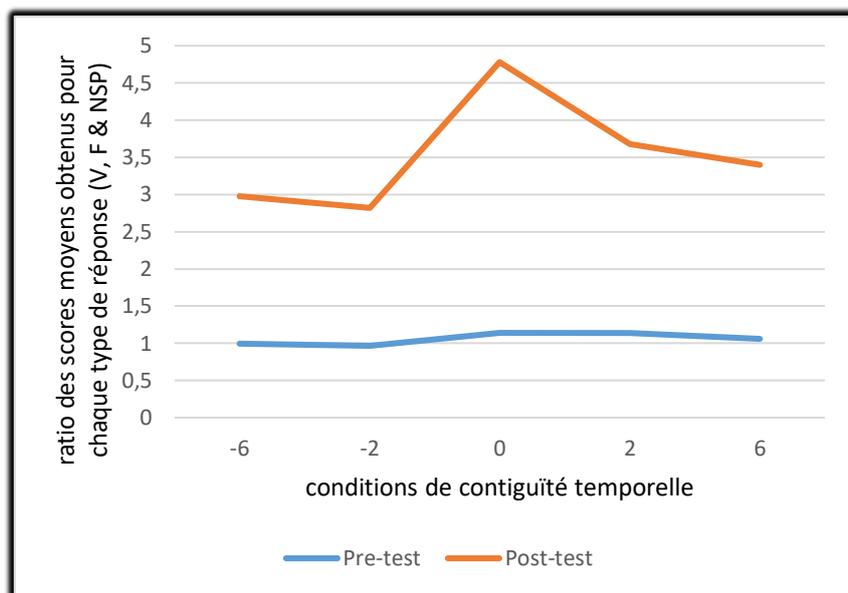


Figure 5.4. Scores (moyennes) pour le score total pondéré dans les cinq conditions de contiguïté temporelle.

Une première ANOVA réalisée sur les scores du pré-test (sTP) uniquement, avec les conditions de contiguïté temporelle comme facteur inter-sujets et le score total au pré-test comme variable dépendante, n'a montré aucun effet des conditions de contiguïté temporelle, $F(4,222) = 0.59$, $p = 0.67$, $\eta^2 = 0.01$, et la comparaison post-hoc n'a indiqué aucune différence entre les 5 groupes.

Une ANOVA à mesures répétées a ensuite été réalisée, avec les conditions de contiguïté temporelle comme facteur inter-sujets, le pré-test et le post-test comme facteurs intra-sujets et les scores totaux (sTP) comme variables dépendantes. Un effet principal global des conditions de contiguïté temporelle a été trouvé, $F(4,222) = 3.95$, $p = 0.004$, $\eta^2 = 0.07$, principalement en raison de la grande amélioration des scores après la phase d'apprentissage dans le groupe synchronisé. Un fort effet d'apprentissage du pré-test au post-test a également été trouvé, $F(1, 222) = 239.38$, $p < 0.00001$, $\eta^2 = 0.51$. Une interaction significative entre les conditions de contiguïté temporelle et les scores pré-post-test (Figure 5.4), $F(4, 222) = 3.59$, $p = 0.007$, $\eta^2 = 0.06$, a révélé que l'effet d'apprentissage du pré-test au post-test différait entre les conditions de contiguïté temporelle : le groupe 0s (synchronisé) a surpassé les autres groupes. Pour les scores du post-test, la figure 5.5 semble montrer une asymétrie entre les délais négatifs et positifs, par rapport à la contiguïté synchronisée. Plus précisément, les comparaisons univariées des scores au post-test ont montré une supériorité de la condition synchronisée sur les conditions désynchronisées, à la fois lorsque la narration a eu lieu avant l'animation (-6 s. et -2 s.), $F(1,222) = 15.95$, $p = 0.0001$, $d = 0.81$) et également lorsque la narration a eu lieu après

l'animation (+6s. et +2s.), $F(1, 222) = 7.33, p = 0.007, d = 0.46$). Cependant, nous n'avons pas trouvé de différences significatives entre les conditions -6s. et -2s., $F(1, 222) = 0.09, p = 0.76$) ou entre +6s. et +2s., $F(1, 222) = 0.33, p = 0.56$. De plus, les participants des groupes des conditions de narration avant animation (-6s. et -2 s.) avaient tendance à avoir des scores post-test inférieurs à ceux des participants des groupes dans les conditions où la narration est diffusée après l'animation (+2s. et +6s.), $F(1, 222) = 3.46, p = 0.07, d = 0.27$.

En résumé, notre première prédiction, **H1a**, évoquant un effet principal de la contiguïté temporelle sur les performances d'apprentissage et de compréhension du pré-test au post-test a été confirmée, la condition synchronisée conduisant à des performances au post-test plus élevées que dans les conditions désynchronisées, indépendamment du retard, qu'il soit de 2s. ou 6s. Nous avons également trouvé un effet d'asymétrie marginal, **H1b** : les scores de performances d'apprentissage étaient légèrement plus élevés lorsque les images RV dynamiques étaient présentées avant les mots prononcés que lorsque les images étaient présentées après les mots prononcés. Ceci est cohérent avec les résultats obtenus par Baggett, (1984), mais pas avec ceux de Kahsharem et al. ; (2019), ni de Xie, Mayer, Wang & Zhou, (2018).

5.3.3 Mémoire de travail, capacité spatiale, contiguïté temporelle et performances d'apprentissage au post-test

Afin de tester l'effet de la mémoire de travail verbale (MDTV) et des habiletés spatiales (HS) sur les performances d'apprentissage (H2a) ; et pour étudier si cet effet potentiel pouvait différer en fonction de la longueur des écarts temporels entre les mots prononcés et les images RV (H2b), trois analyses ont été menées.

Premièrement, des régressions simples des scores de MDTV et des scores d'HS ont été effectuées en tant que prédicteurs continus uniques des résultats des post-test factuels et de compréhension, respectivement. La régression a montré un effet direct significatif de la MDTV sur les résultats d'apprentissage, variable post-test, R multiple = 0.18, $R^2 = 0.032$, (ajusté = 0.028), $F(1, 225) = 7.90, p < 0.006$. La régression a également révélé un effet direct significatif de l'HS sur les résultats d'apprentissage, variable post-test, R multiple = 0.24, $R^2 = 0.060$, (ajusté = 0.056), $F(1, 225) = 14.48, p = 0.0002$) et voir Tableau 5.3, en bas.

Deuxièmement, deux ANCOVA ont été réalisées, avec les conditions de contiguïté temporelle comme facteur inter-sujets, MDTV ou HS, respectivement, comme facteurs de covariance

continus et le score au post-test comme variable dépendante. Pour la première ANCOVA qui incluait la performance de MDTV, un effet principal du facteur covariant a été trouvé, $F(1, 221) = 9.98, p < 0.002, \eta_p^2 = 0.043$; avec un effet résiduel significatif des conditions de contiguïté temporelle, $F(4, 221) = 4.66, p < 0.002, \eta_p^2 = 0.078$. La deuxième ANCOVA, qui incluait la performance HS, a montré un effet principal du facteur covariant, $F(1, 221) = 12.15, p = 0.0006, \eta_p^2 = 0.052$ et un effet significatif des conditions de contiguïté temporelle, $F(4, 221) = 3.56, p = 0.008, \eta_p^2 = 0.060$.

Troisièmement, afin de comprendre plus directement les effets de la MDTV et des HS dans chacune des conditions du facteur de contiguïté temporelle, des régressions simples ont été effectuées sur les effets de la MDTV et des HS respectivement, sur les scores pré-test et post-test au sein de chaque groupe de contiguïté temporelle. Les résultats de ces analyses sont présentés dans le Tableau 5.3.

Table 5.3

Résultats des régressions simples des scores de mémoire de travail verbale et d'aptitudes spatiales sur les performances pré-test et post-test pour chaque condition de contiguïté temporelle.

	Verbal Working Memory		Spatial Abilities	
	Pre-test	Post-test	Pre-test	Post-test
-6	$R = .21, R^2$ (adj.) = .020; $F(1,39) = 1.83, p$ = .18	**$R = .34, R^2$ (adj.) = .12; F (1,39) = 5.28, p = .026	$R = .12, R^2$ (adj.) = .0010; $F(1,39) = 0.58, p$ = .45	**$R = .34, R^2$ (adj.) = .096; $F(1,39) = 5.27,$ $p = .027$
-2	$R = .09, R^2$ (adj.) = .007; $F(1,41) = 0.31, p$ = .58	$R = .032, R^2$ (adj.) = .023; $F(1,41) = 0.04, p$ = .84	$R = .05, R^2$ (adj.) = .021; $F(1,41) = 0.11, p$ = .74	$R = .22, R^2$ (adj.) = .026; $F(1,41) = 2.15, p$ = .15
0	$R = .25, R^2$ (adj.) = .070; $F(1,38) = 2.70, p$ = .11	**$R = .43, R^2$ (adj.) = .17; $F(1,38) = 8.95,$ $p = .005$	$R = .12, R^2$ (adj.) = .010; $F(1,38) = 0.60, p$ = .44	$R = .24, R^2$ (adj.) = .032; $F(1,38) = 2.29, p$ = .14
+2	$R = .076, R^2$ (adj.) = .02;	$R = .08,$ 37.50% (adj.) =	$R = .24, R^2$ (adj.) = .039;	$R = .13, R^2$ (adj.) = .006;

	F(1,42)= 0.24, p = .62	.016; F(1,42)= 0.28, p = .59	F(1,42)=2.75, p = .104	F(1,42)= 0.73, p = .40
	R = .15, R^2	R = .19, R^2	R = .039, R^2	*R = .25, R^2
+6	(adj.) = .008; $F(1,57)$ = 1.46, p = .23	(adj.) = .020; $F(1,57)$ =2.19, p = .14	(adj.) = .016; $F(1,57)$ = 0.09, p = .77	(adj.) = .045; $F(1,57)$= 3.74, p = .057
All conditions	R = .004, R^2 (adj.) = .004; $F(1,225)$ = 0.0045, p = .94	**R = .18, R^2 (adj.) = .030; $F(1,225)$= 7.90, p = .0054	R = .10, R^2 (adj.) = .007; $F(1,225)$ = 2.53, p = .11	**R = .24, R^2 (adj.) = .060; $F(1,225)$ = 14.48, p = .0002

Pour résumer, nous avons trouvé des effets significatifs de la MDTV et des HS sur les scores d'apprentissage post-test uniquement. Cet effet était particulièrement significatif pour la MDTV dans les conditions extrêmement négatives -6s. et dans les conditions synchronisées 0. L'effet de l'AS était significatif pour les conditions extrêmes -6 et +6 (marginal). Les capacités de MDTV et des HS peuvent modérer l'effet des conditions de contiguïté temporelle. Les processus sous-jacents de l'effet modérateur pourraient être différents dans les conditions extrêmes d'écart temporel - 6 et dans la condition temporelle simultanée 0. Ce relevé sera discuté plus tard et des explications hypothétiques seront fournies.

5.3.4 Test de mémoire épisodique de correspondance texte-image

Pour le test de correspondance texte-image, les résultats de la proportion de choix respectivement -6, -2, 0, +2, +6 secondes de correspondance texte-image dans chaque condition de contiguïté temporelle sont présentés dans le Tableau 5.4 ci-dessous, et pour une meilleure illustration visuelle dans la Figure 5.5.

Tableau 5.4

Proportion en pourcentages (Moyennes et SD) de respectivement -6, -2, 0, +2, +6 secondes choix de correspondance texte image dans chaque condition de contiguïté temporelle.

Cond. De contiguïté temporelle. Choix de correspondance % *	-6	-2	0	+2	+6	Le total
-6 % au choix	34.14 (26.07)	21,70 (21,37)	11,25 (23,68)	8,71 (11,64)	15,81 (20,15)	18.06 (22.54)
-2 Choix %	35,36 (24,20)	29.45 (21.77)	27.08 (20.90)	22,24 (14,79)	25,98 (16,45)	27,82 (19,90)
0 Choix %	62,29 (14,70)	68,50 (14,58)	77,19 (13.10)	70.07 (6,97)	64,69 (9,97)	68.22 (12.91)
+2 Choix %	38.21 (16.34)	36.04 (20.55)	27,91 (13.80)	42,80 (17,75)	49,15 (17,63)	39,72 (18.73)
+6 % de choix	33,73 (19,89)	32,55 (19,56)	24.16 (24.44)	46.21 (22.69)	50.28 (22.20)	38,54 (2374)

* : proportion de choix par rapport au nombre maximum d'images dans chaque condition présentée dans le post-test de correspondance texte-image, c'est-à-dire que pour chaque condition temporelle, les pourcentages sont supérieurs à 100 %.

Quatre indicateurs ont été utilisés pour l'analyse des données du Tableau 5.4, comme décrit dans la partie méthode ci-dessus : (i) la proportion (%) de choix des items texte-image synchronisés du couple présenté, (ii) la proportion (%) de choix des items texte-image de la condition à laquelle l'apprenant avait été affecté (iii) la proportion (%) de choix des items texte-image d'une autre condition avec un décalage temporel, différent de celui auquel l'apprenant avait été assigné et (iv) la proportion (%) de choix des items texte-image autres que les items texte-image synchronisés. Les mesures dépendantes correspondaient à quatre types de réponses : (i) le choix attendu c'est-à-dire l'image correspondant à la condition appropriée, (ii) la réponse texte-image intégrée qui correspondait à la condition narration-animation simultanée, 0 ; (iii) le choix inattendu, par exemple l'image correspondant à une autre condition ayant un décalage temporel ; (iv) le type de réponse désynchronisé, voir Figures 5.5 (a,b,c,d). Quatre ANOVA à un facteur ont été réalisées, avec la condition de contiguïté temporelle comme facteur inter-sujets et chaque indicateur comme variable dépendante.

Premièrement, l'ANOVA concernant la réponse intégrée texte-image (choix correspondant à la version synchronisée, 0, texte-image), avec des conditions de contiguïté temporelle comme facteur inter-sujets, et la proportion de choix intégrés texte-image comme variable dépendante, a montré un effet principal des conditions temporelles, voir aussi la Figure 6a, $F(4, 223) = 9.57$, $p < 0.00001$, $\eta_p^2 = 0.15$). Lorsque la plupart des éléments texte-image synchronisés ont été choisis (de 62 % dans la condition -6 à plus de 77 % dans la condition 0), les participants ont choisi l'élément intégré texte-image dans les conditions retardées moins fréquemment que dans la condition 0 (présentation synchronisée), respectivement pour les délais -6 et +6, $F(1, 222) = 25.98$, $p < 0.00001$, $d = 0.84$; pour les délais -2 et +2, $F(1, 222) = 19.01$, $p = .0001$, $d = .88$. Le délai -6 a été choisi moins fréquemment que le délai -2, $F(1, 222) = 5.59$, $p < .02$, $d = 0.42$; et le délai +6 a été choisi moins fréquemment que le délai +2, $F(1, 222) = 5.098$, $p < 0.03$, $d = 0.62$

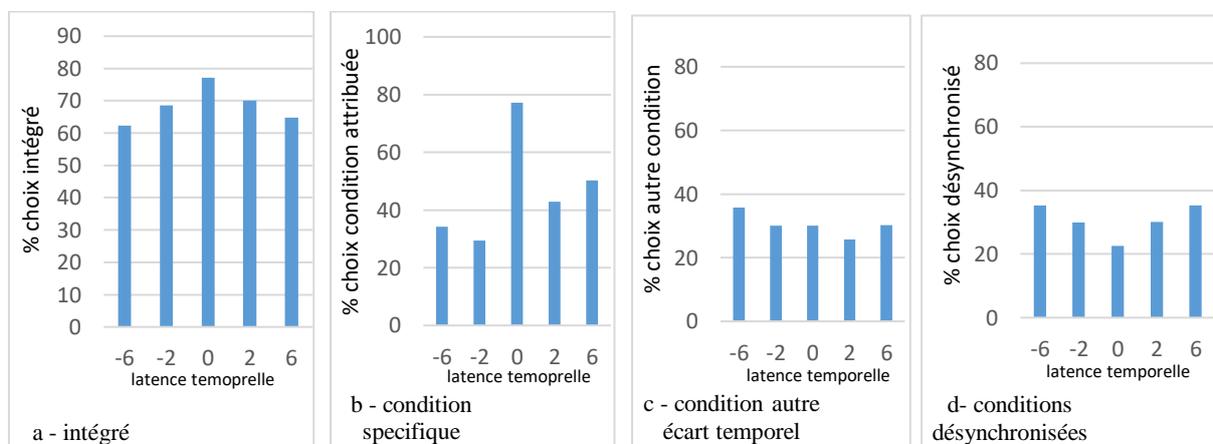


Figure 5.5. Proportion (%) de (a) choix de l'élément texte-image intégré (b) choix de l'élément texte-image attendu et (c) choix d'éléments texte-image désynchronisés correspondant à un décalage autre que celui de la condition assignée au participant (d) toutes les conditions désynchronisées.

La seconde ANOVA a été réalisée sur le choix attendu, c'est-à-dire les items texte-image correspondants à la condition d'apprentissage du participant (par exemple -6, -2, 0, +2 ou +6), avec la contiguïté temporelle comme facteur inter-sujets et les éléments texte-image du choix de la paire comme variable dépendante. Les résultats ont révélé un fort effet significatif de la contiguïté temporelle, voir Figure 5.5b, $F(4, 222) = 33.45$, $p < 0.0001$, $\eta_p^2 = 0.37$). Sans surprise, les participants de la condition 0, (synchronisée) ont choisi les éléments texte-image intégrés beaucoup plus fréquemment que les participants des autres conditions, avec le groupe synchronisé supérieur aux groupes -6 et -2, $F(1, 222) = 129.1$, $p < .0001$, $d = 1.98$) et aux groupes +2 et +6, $F(1, 222) = 62.21$, $p < 0.0001$, $d = 1.78$), respectivement. De plus, un

pourcentage important de choix correspondant à la condition du participant peut être noté dans les conditions différées (-2, -6, +2, +6), Figure 6b. Ce constat révèle que les participants ont perçu, avec au moins une certaine conscience, le décalage narration-animation lors de la phase d'apprentissage. De plus, la conscience du décalage temporel était significativement plus élevée dans les conditions +2s. et +6s. que dans les conditions -6s. et -2s., $F(1, 222) = 23.02, p < 0.001, d = 0,66$.

La troisième ANOVA a été réalisée sur le choix non attendu, en excluant les paires d'items texte-image intégrés ; par exemple les éléments texte-image correspondant à des conditions retardées autres que la condition d'apprentissage du participant (par exemple -6, -2, +2 ou +6), avec la contiguïté temporelle comme facteur inter-sujets et les éléments texte-image du choix de la paire comme variable dépendante. Les résultats ont montré un effet des conditions de contiguïté temporelle, $F(4, 222) = 3.024, p < 0.02, \eta_p^2 = 0,05$. Cependant, cet effet était principalement dû au fait que la proportion de choix inattendus était plus élevée dans la condition -6 que dans les autres conditions (Figure 6c). Les comparaisons réalisées ont montré une différence entre les conditions -6 et -2 ($F(1, 222) = 3.81, p = 0.02, d = 0.41$) ; -6 et 0, ($F(1, 222) = 3.63, p < 0.06, d = 0.37$, marginal) ; -6 et +2, ($F(1, 222) = 12.04, p < 0.001, d = 0.85$) ; -6 et +6, ($F(1, 222) = 4.07, p < 0.5, d = 0.43$). Cependant, il n'y avait pas de différence entre -2, 0, +2 et +6.

La dernière ANOVA a été réalisée sur tous les choix désynchronisés (voir Figure 5.5d), avec la contiguïté temporelle comme facteur inter-sujets et les éléments texte-image du choix de la paire comme variable dépendante. Les résultats ont montré un effet principal significatif des conditions de contiguïté temporelle, $F(4, 222) = 9.16, p < 0.0001, \eta_p^2 = 0.14$. Par rapport à la condition 0 synchronisée : les participants aux conditions -6 et -2 et aux conditions +2 et +6 ont choisi plus souvent une condition désynchronisée, respectivement $F(4, 222) = 20.94, p < 0,0001, d = 0,77$; $F(4, 222) = 22,19, p < 0,0001, d = 0,91$. Il y avait aussi des différences entre : -6 et -2 ($F(1, 222) = 4.72, p = 0.031, d = 0.41$) ; -2 et 0 ($F(1, 222) = 8.51, p < 0.01, d = 0.56$) ; +6 et +2, $F(1, 222) = 5.40, p < .03, d = 0.60$; +2 et 0, $F(1, 222) = 8.81, p < 0.01, d = 0.70$. Cette hiérarchie significative claire entre les différentes conditions désynchronisées suggère que les élèves n'ont pas choisi d'éléments au hasard au sein des paires, mais avaient tendance à être conscients des délais présents entre le texte et les images.

En somme, les résultats de cette tâche de mémoire épisodique dans laquelle les participants devaient choisir, pour chaque paire, quelle image de la paire ils pensaient avoir vue, ou se souvenaient avoir vue, lorsqu'ils ont entendu la phrase pendant la leçon de RV, suggèrent que les apprenants n'ont pas répondu au hasard. Au contraire, les résultats ont indiqué que deux principaux facteurs significatifs ont influencé leurs choix dans cette tâche de correspondance texte-image. Premièrement, un fort effet du traitement d'intégration cognitive du texte et de l'image pendant la phase d'apprentissage a conduit à un choix plus fréquent de l'item texte-image intégré au sein des paires dans toutes les conditions, suggérant un processus d'intégration cognitif très actif. D'autre part, dans les conditions différées, le niveau de choix du couple texte-image était significativement plus faible. Ces résultats démontrent que les apprenants étaient, dans une certaine mesure, conscients des décalages au cours de la phase d'apprentissage, en particulier dans les conditions +2 et +6. De plus, nous avons trouvé des corrélations entre la proportion du choix attendu de la condition de l'apprenant et le score au post-test qui, bien que peu élevées, étaient significatives, r (Bravais-Pearson, $N = 227$) = 0.16, $p < .05$; et aussi entre la proportion de choix d'items textes-images intégrés et les scores au post-test, r (Bravais-Pearson, $N = 227$) = 0.17, $p < 0.05$). Les régressions de choix sur les scores au post-test étaient également significatives, respectivement $R^2 = 0.24$ R^2 ajusté = 0.020, $F(1, 225) = 5.72$, $p < 0.02$; et $R^2 = 0.029$, R^2 ajusté = 0.025, $F(1, 225) = 6.85$, $p < 0.01$. Ce point sera discuté plus tard. Enfin, nous avons trouvé une corrélation négative faible mais significative entre la mesure de la mémoire de travail verbale et le choix de conditions d'écart temporel autres que la propre condition du sujet, r (Bravais-Pearson, $N = 227$) = -.15, $p < .05$; et une corrélation faible mais significative entre la mesure de la mémoire de travail verbale et le choix de la condition 0 intégrée, r (Bravais-Pearson, $N = 227$) = .13, $p < .05$.

5.4 Discussion et conclusion

L'objectif de cette étude était d'étendre et d'approfondir les résultats des recherches antérieures sur l'effet de la contiguïté temporelle sur l'apprentissage de contenus complexes à partir d'une présentation multimédia, plus particulièrement dans le contexte de la réalité virtuelle. Beaucoup moins d'études ont été réalisées sur le principe de la contiguïté temporelle et ses limites que sur le principe de contiguïté spatiale (Mayer, 2009, 2014, 2006 Ginn., Khacharem et al, 2020). Cette étude a tenté de répondre à plusieurs questions nouvelles. Elle fait partie d'un projet plus vaste, qui vise à créer un simulateur de forêt interactif en RV immersive, afin d'en apprendre davantage sur les écosystèmes forestiers dans des

établissements d'enseignement secondaire. La synchronisation de la narration et de l'animation peut être d'une importance cruciale en RV immersive, en particulier parce que lors des activités de navigation et d'exploration, les mouvements de la tête dirigent la dynamique visuelle par le regard. Dans la présente étude, nous avons d'abord utilisé une vidéo de 12 min en RV non immersive (sur écran d'ordinateur) sur le processus de décomposition des sols forestiers et de ses effets sur le développement des arbres, des plantes et des animaux. L'expérience a eu lieu en milieu scolaire dans le cadre d'un programme d'enseignement des sciences de la vie et de la terre destiné à des élèves de 12 ans, qui incluait l'apprentissage du fonctionnement de l'écosystème forestier. Tester les principes multimédias en dehors du laboratoire peut être pertinent. Contrairement à la plupart des études précédentes sur la contiguïté temporelle qui testaient des présentations successives séquentielles d'informations verbales et picturales, nous avons utilisé une présentation continue dans laquelle l'animation se déroulait continuellement tandis que la narration se déroulait en parallèle, avec différentes latences possibles entre le texte et l'animation. Nous avons utilisé de nouvelles latences temporelles plus courtes entre la narration et la diffusion de l'animation que dans les études précédentes, elles étaient respectivement de -6, -2, 0, +2 et +6 secondes. Afin d'examiner les processus cognitifs impliqués dans l'intégration de la narration et de l'animation tout au long de la visualisation de la vidéo en RV dans les différentes conditions temporelles et dans le contexte des différences individuelles, nous avons mesuré la capacité de mémoire de travail verbale ainsi que les habiletés spatiales des élèves. En plus des résultats d'apprentissage de rappel et de compréhension, nous avons ajouté un test de mémoire épisodique afin d'évaluer si (et si oui, dans quelle mesure) les apprenants étaient conscients de l'écart présent entre la narration et les images animées correspondantes. Dans cette dernière alternative, l'effet potentiellement fort du processus d'intégration cognitive entre la narration et l'animation lors de la compréhension multimédia (voir le modèle CTML, Mayer, 2009, 2014 et le modèle intégratif de Schnotz, 2005) peut « supprimer » ou nuire à la mémorisation précise de la présentation de la condition de latence temporelle pendant la phase d'apprentissage (voir aussi Arndt, Schüler, & Scheiter, 2019).

Deux types de résultats d'apprentissage ont été utilisés : un post-test de connaissances et de compréhension factuelle et un post-test de mémoire épisodique. La mémoire de travail verbale et les capacités spatiales des apprenants ont également été mesurées.

Quatre hypothèses ont été formulées.

Concernant l'hypothèse 1a, nous avons trouvé un principal effet significatif de la contiguïté temporelle sur le rappel factuel et les performances de compréhension. La condition synchronisée a donné lieu à de meilleures performances en post-test que les conditions désynchronisées, quelle que soit la latence temporelle (2s. Ou 6s). Même un délai de deux secondes entre la narration et l'animation semblait potentiellement préjudiciable aux performances d'apprentissage des jeunes élèves. Dans le cas d'un document multimédia-RV-assez long (12 minutes), les processus impliqués dans le développement progressif, continu et croissant de l'intégration cognitive narration-animation et l'organisation de l'information en mémoire de travail (Mayer, 2009, 2014), pourraient être perturbés même quand l'écart entre la narration et les visualisations dynamiques est faible. Un effet d'asymétrie marginal, H1b, a également été trouvé : les performances d'apprentissage étaient légèrement meilleures lorsque les images dynamiques étaient présentées avant la narration que lorsqu'elles étaient présentées après la narration. Ce résultat est cohérent avec les conclusions de Baggett (1984), qui a montré la même tendance. Cependant, cette tendance est différente des résultats décrits par Khacharem et al. ; 2019, et aussi Xie, et al., (2018) dans leurs études. Ces différences ne représentent pas nécessairement une contradiction. Cette tendance pourrait révéler l'existence de mécanismes d'intégration texte-image de substitution intéressants en mémoire de travail, dépendamment du type de matériel d'apprentissage et de la tâche requise. Dans le cas présent d'une présentation continue (comme dans l'étude de Baggett, 1984) au lieu d'une présentation successive (comme dans la plupart des études antérieures) de la narration et des animations, quand la narration est diffusée avant les images correspondantes, les apprenants peuvent souvent être à l'écoute des mots prononcés alors qu'ils regardent effectivement des images qui ne correspondent pas, ou ne correspondent pas assez précisément, aux mots et/ou qui peuvent également être sans rapport, ou incompatibles avec les mots prononcés qu'ils sont en train de traiter à ce moment-là. Comme indiqué ci-dessus, la conséquence potentielle peut être que les apprenants s'engagent dans une activité de recherche visuelle inutile afin de trouver des informations visuelles liées aux mots qu'ils viennent d'entendre. Cette activité pourrait augmenter la charge extrinsèque (Sweller, 2010) et nuire aux processus d'intégration. Elle peut également entraîner la perte de l'information parlée voir de de potentielles erreurs d'interprétation. (Baggett, 1984, et dans le processus de déclin temporel en mémoire de travail décrit dans le modèle élaboré par Barrouillet, Bernardin et Camos, (2004). En revanche, lorsque les images dynamiques étaient présentées avant la narration, ou mieux, simultanément (version synchronisée), on peut supposer que la narration a fourni des informations sur l'interprétation des images vues immédiatement avant (ou en même temps), et a donc facilité l'intégration des images dans la

narration. La narration peut alors rendre les images plus pertinentes. Cette interprétation semble également en accord avec le développement récent du modèle de traitement multimédia texte-image élaboré par Schnotz (2005). Des études empiriques récentes menées par Schnotz & Wagner (2018) ont apporté la preuve que, dans le traitement multimédia, le texte « prime » (sur les images), en particulier lorsque celui-ci est écrit. Selon Schnotz et Wagner (2018), on peut supposer que « le *traitement conjoint du texte et des images* » se déroule de manière asymétrique. Texte et images « *remplissent des fonctions différentes mais complémentaires* ». Le traitement conjoint est censé commencer par la construction du modèle général, axé sur la cohérence mentale. Au cours du processus de compréhension, le modèle mental est « *adapté aux exigences de la tâche par l'élaboration de modèles mentaux adaptatifs* ». Les auteurs ont émis l'hypothèse selon laquelle la construction de modèles mentaux est d'abord plus axée sur le texte que sur les images, et que l'élaboration ultérieure de modèles mentaux adaptatifs est davantage axée sur les images. Si nous regardons nos résultats à la lumière du modèle de Schnotz et Wagner (2018), lorsque l'information verbale était présentée avant les images correspondantes, les apprenants ont pu être induits en erreur si les images visionnées lors de l'occurrence de cette information verbale n'étaient pas utiles, ou ne correspondaient pas avec suffisamment de précision à la narration. Cependant, en dépit d'une tendance à l'asymétrie, et contrairement à l'état synchronisé contigu, nos résultats ont également montré que les décalages positifs +2 et +6 ont eu un effet négatif sur la compréhension. Lorsque l'explication verbale était donnée après les images pertinentes correspondantes dans l'animation, d'autres, les images suivantes, et de nouvelles phrases, peuvent avoir remplacé les images et les mots précédents qui eux étaient pertinents avec l'information verbale précédente. La distance temporelle entre les mots prononcés et les images dynamiques correspondantes peut également avoir imposé une forte demande de ressources cognitives en mémoire de travail. La qualité du traitement d'intégration cognitive peut donc avoir été affectée. Cependant, de telles spéculations sur le traitement cognitif sous-jacent des conditions de contiguïté temporelle devraient être étudiées de manière plus empirique, en utilisant des méthodes Online telles que des études de suivi oculaire.

En ce qui concerne l'hypothèse 2, comme prévu, des effets significatifs sur l'empan de la mémoire de travail verbale et les habiletés spatiales sur les performances d'apprentissage ont été trouvés, avec des corrélations significatives: les élèves ayant une MDTV élevée et de fortes HS ont mieux appris que les élèves ayant une faible MDTV ou de faibles HS (**H2a**). Ce pattern de résultats peut sembler évident ; Cependant, lorsque les facteurs MDTV et HS étaient

contrôlés, l'effet de contiguïté temporelle a subsisté et cela démontre que les différences individuelles ne compensent pas vraiment l'effet négatif des écarts temporels, en particulier l'écart extrême de -6sc. Un niveau élevé de demande imposé par les latences temporelles entre la narration et l'animation peut avoir induit une diminution des ressources disponibles dédiées au contrôle cognitif, conduisant à une augmentation de la quantité d'informations non pertinentes traitées (Lavie, 2005; Lavie et al., 2004). Les régressions effectuées sur les groupes de chaque condition (-6, -2, 0, +2, 6) ont montré que l'effet de la MDTV et des HS semble varier en fonction de la longueur des intervalles temporels entre la narration et les images, Tableau 5.3, **H2b**. Tout d'abord, dans le plus difficile, la condition extrême -6, les apprenants ayant des performances élevées de MDTV ont eu des performances plus élevées au post-test. Les régressions ne sont pas significatives et faibles pour les conditions -2 et 2. Pour +6 la régression n'a pas atteint la significativité ($p = .14$). Il est intéressant de noter qu'étonnamment, les apprenants ayant de bonnes performances en MDTV avaient également tendance à avoir de bonnes performances en post-test dans la condition 0 (synchronisée). En surface, ces résultats peuvent sembler contradictoires, mais nous pensons que ce n'est pas nécessairement le cas. Nous proposons deux explications possibles, qui, à ce stade restent spéculatives. Dans la condition -6, le décalage temporel important entre la narration et les images dynamiques peut avoir augmenté la charge extrinsèque liée à la difficulté de relier l'information verbale à l'information picturale (Kalyuga, 2011, Sweller, 2010). Les apprenants ayant un bon niveau de MDTV auraient été en mesure de mieux gérer cette charge extrinsèque que les apprenants avec un niveau de MDTV inférieur. La condition 0 (synchronisée) pourrait favoriser la charge germane, et les apprenants à forte MDTV pourraient « faire pleinement usage » de leurs capacités de traitement de l'information en la mémoire de travail. Bien sûr, ces explications hautement spéculatives doivent être plus empiriquement validées. Ensuite, des régressions significatives importantes pour les conditions extrêmes -6 et +6 ont également révélé que les apprenants à haute HS ont mieux compensé la difficulté liée à la charge extrinsèque. Il se peut que pour ces apprenants, la plus grande facilité de traitement d'images et des relations qu'elles entretiennent a conduit à libérer des ressources cognitives permettant ainsi le traitement des informations verbales pertinentes associées.

Enfin, une hypothèse plus ouverte et exploratoire, **H3**, a été émise au sujet d'un potentiel effet des latences temporelles sur la capacité de reconnaître, en post-test, les paires texte-image qui ont été présentées au cours de la phase d'apprentissage. Une telle reconnaissance pourrait être difficile compte tenu de l'aspect transitoire de la présentation audiovisuelle dynamique de la

leçon multimédia en RV. Les résultats ont montré deux principaux ensembles complémentaires de traitement. Tout d'abord, quelle que soit la condition de contiguïté temporelle, le pourcentage élevé de choix de la paire avec la correspondance texte-image intégrée a révélé un processus d'intégration texte-image fort et significatif, en particulier dans le cas d'une présentation multimédia narration-animation dans un format RV riche. Ainsi, conformément à l'hypothèse récemment émise par Arndt, Schüller & Scheiter (2019), nos résultats ont montré que l'intégration texte image se produit aussi bien lorsque les apprenants doivent lire des passages de texte plus longs combinés avec des images, que dans la plupart des recherches empiriques. En second lieu, et dans une mesure moindre que dans les processus d'intégration précédents, les apprenants de la condition temporelle désynchronisée étaient conscients d'un écart entre la narration et l'animation RV. Ils n'étaient pas (toujours) capables d'estimer précisément l'écart temporel exact, mais ils ont choisi plus souvent l'item du couple texte-image désynchronisé que les apprenants de la condition synchronisée (cela montre aussi qu'ils n'ont pas choisi au hasard). De plus, l'efficacité de ces deux types de traitements peut avoir eu des conséquences sur le niveau de performance en post-test. Dans une certaine mesure, il semble qu'un niveau élevé de conscience de l'écart temporel soit plus susceptible d'être associé à un score élevé en post-test. Le choix de l'élément intégré des paires texte-image est également plus susceptible d'être associé à des scores élevés en post-test. Bien entendu, l'une des limitations de ce type de tâche de choix de correspondance texte-image était que nous ne pouvions pas exclure l'occurrence de confusion entre les différents éléments dans les conditions texte-image désynchronisées. Les apprenants n'avaient probablement pas de souvenirs précis des écarts temporels dont ils ont fait l'expérience pendant la phase d'apprentissage. Cependant, nous pensons que les résultats actuels dans ce post-test épisodique ne sont pas influencés par une mauvaise compréhension de la tâche, qui consistait à choisir le texte que les élèves écoutaient quand l'image animée en RV a été présentée dans la tâche d'apprentissage, et non le texte qui était le plus sémantiquement lié aux images animées. Il est à noter que nous avons vérifié la bonne compréhension de la tâche à plusieurs reprises lors de l'administration du post test

Conclusion, travaux futurs.

Dans cette étude, un matériel RV riche et non immersif a montré que même un petit écart temporel entre la narration et l'animation nuisait aux performances d'apprentissage. Des recherches supplémentaires sont nécessaires pour obtenir une meilleure compréhension du mécanisme cognitif sous-jacent de l'intégration texte-image et des processus de compréhension

dans les longues présentations multimédias éducatives. L'analyse des mouvements oculaires pourrait s'avérer un outil utile dans les études à venir (Xie, Mayer et al., 2018). Dans un avenir proche, une correspondance précise et limitée dans le temps entre le texte parlé (au lieu du texte écrit à l'écran, voir Makransky, 2019) et les images dynamiques pourrait s'avérer importante pour la réalité virtuelle immersive, en raison du degré élevé d'interactivité, comme par exemple le contrôle par l'utilisateur de la dynamique de l'environnement. Cette question est vraiment importante car un nombre croissant d'études récentes (par exemple, Makransky & al., 2019, 2020 ; Araiza-Alba, Keane, Matthews, Simpson, Strugnell, Chen, & Kaufman, 2021) qui comparent l'apprentissage en RV immersive et RV vidéo PC non immersive, ont révélé un apprentissage plus altéré et une charge cognitive plus importante en condition immersive.

**CHAPITRE 6 - APPRENTISSAGES EN
REALITE VIRTUELLE IMMERSIVE :
ROLE DE DEUX TYPES DE
SIGNALISATION ET DU FEEDBACK.**

6.1 Introduction.

L'objectif de cette étude est de tester deux niveaux de guidage par signalisation visuelle ainsi que le rôle du feedback dans le cadre d'une tâche de décision et d'apprentissage en réalité virtuelle immersive. Avant de développer plus précisément notre recherche, nous allons analyser les études antérieures réalisées dans le cadre d'apprentissages en RVI. Pour commencer, nous allons aborder le caractère contrasté des résultats obtenus pour ce type d'apprentissage puis nous évoquerons la nécessité d'étudier les mécanismes perceptifs en jeu dans les tâches d'apprentissage en RVI. Nous ferons ensuite un point sur les études ayant testé la signalisation en RVI et pour finir, nous aborderons le rôle du feedback en RVI.

6.1.1 Apprendre avec la réalité virtuelle immersive : des résultats très contrastés.

Une partie des travaux réalisés sur les apprentissages en réalité virtuelle montrent des résultats positifs avec un gain de connaissances accru en faveur de la présentation en RVI en comparaison d'autres dispositifs plus conventionnels.

Par exemple, l'étude de Liu, Wang, Lei, Wang & Ren (2020) avait pour objectif de comparer l'apprentissage traditionnel en classe à partir de supports conventionnels (manuels scolaires) à l'apprentissage en RVI chez de jeunes enfants. Les auteurs ont intégré des sessions d'apprentissage de groupe en RVI à l'aide de casques VR pendant la classe. Les résultats obtenus ont montré des performances supérieures en termes d'acquisition de connaissances académiques et d'engagement en faveur de l'apprentissage en RVI comparé à un apprentissage traditionnel. Cependant, la méthode employée pour cette étude n'était pas rigoureusement expérimentale, ce qui ne permet pas de considérer ces résultats comme étant totalement fiables. Ce même profil de résultats est observé par Saputro & Setyawan (2020) qui ont comparé les performances d'apprentissage du système solaire entre une présentation traditionnelle et une présentation RVI. Ici, les scores d'acquisition de connaissances en RVI surpassaient les scores d'acquisition de connaissances du mode traditionnel. Cependant, la méthodologie employée n'était pas strictement expérimentale et le nombre de participants était très faible (N=22). Le même profil de résultat est observé dans l'étude de Maresky & al. (2019) sur l'apprentissage de l'anatomie cardiaque. De nouveau, les résultats obtenus par les participants ayant bénéficié d'une présentation RVI surpassent les performances des apprenants ayant étudié à l'aide d'un support conventionnel. Néanmoins, cette recherche a été menée à des fins d'enseignement et

n'avait pas pour but d'analyser les processus cognitifs liés à l'apprentissage en RVI. De plus, le nombre de participants était de 42 pour trois groupes expérimentaux.

Si la méta-analyse d'Hamilton & al. (2021) indique que dans la majorité des cas la RVI montre des résultats significativement plus élevés en termes d'acquisitions de connaissances comparé aux supports traditionnels, les résultats qu'ils ont obtenus sont exprimés sous forme de pourcentages et n'ont pas fait l'objet d'un traitement statistique rigoureux.

On observe des résultats un peu plus nuancés par les travaux de Brewer & al. (2021) qui avaient pour objectif de comparer le type de visualisation sur les performances d'apprentissage de l'anatomie du corps humain. Ils ont comparé l'apprentissage conventionnel en laboratoire, l'apprentissage 2D et l'apprentissage RV. Leurs résultats montrent une supériorité des performances d'apprentissages en RVI par rapport à l'apprentissage réalisé à l'aide d'une vidéo 2D. Dans ce cas, l'apprentissage RVI était équivalent à l'apprentissage en laboratoire. Cela montre que l'apprentissage RVI peut parfois se substituer à l'apprentissage traditionnel au sens où il permet un apprentissage équivalent. Dans le cas présent il permettrait donc de limiter le recours à la dissection de corps humains tout en permettant malgré tout un apprentissage similaire. Encore une fois, l'échantillon expérimental sollicité pour cette étude était de petite taille (N=38) représentant un unique groupe évalué dans les trois conditions expérimentales. L'étude de Pulijala & al. (2017) montre que l'apprentissage RV est équivalent à un apprentissage classique concernant l'acquisition de connaissances. De plus, l'avantage observé de la RVI est qu'elle permet d'accroître la confiance en soi dans la pratique d'une intervention chirurgicale maxillo-faciale. Le but ici était de limiter le risque d'erreurs lié au stress en situation d'intervention chirurgicale réelle et non l'acquisition de connaissances en tant que telle.

Cependant, souvent, les supports traditionnels se montrent plus efficaces en termes d'apprentissages que la RVI. C'est par exemple ce que montrent les travaux de Parong & Mayer (2020) à travers une tâche d'apprentissage sur la circulation sanguine. Leurs résultats indiquent que la quantité de ressources cognitives sollicitée dans l'apprentissage en RVI est supérieure à celle sollicitée lors d'un apprentissage traditionnel. Cependant les performances en terme d'acquisition de connaissances ne diffèrent pas entre les deux supports. Les auteurs montrent également que les performances de transfert de connaissances sont supérieures lorsque les apprenants étudiaient avec un support traditionnel plutôt qu'avec la RVI. Nous pouvons également évoquer, l'étude de Makransky, Terkildsen & Mayer (2019) dont l'objectif était de comparer un apprentissage à l'aide d'une simulation sur PC ou en RVI sur le thème de la

sécurité en laboratoire. Les auteurs ont mesuré le mouvement oculaire ainsi que l'activité cérébrale via EEG pendant la phase d'apprentissage dans le but de quantifier objectivement la charge mentale engagée. Leurs résultats montrent que si les sentiments de présence et de motivation des apprenants en RVI sont plus importants qu'en simulation sur écran d'ordinateur, leur charge cognitive est également plus élevée ce qui se traduit par un apprentissage moins efficace. Dans le cas présent, les performances d'apprentissage des apprenants ayant étudié sur un écran d'ordinateur ont surpassé celles des apprenants ayant étudié à l'aide d'un casque de réalité virtuelle. Se pose alors la question de la nécessité de l'immersion pour l'apprentissage en fonction des objectifs et des contraintes de la tâche.

6.1.2 Les mécanismes cognitifs liés aux apprentissages en réalité virtuelle immersive.

Si la réalité virtuelle immersive (RVI) offre un sentiment de présence et de motivation elle peut également distraire l'apprenant et provoquer une surcharge des ressources cognitives ce qui finalement, perturberait sérieusement l'apprentissage (Makransky, 2019). Ces idées sont en accord avec les travaux de Lowe & Boucheix (2008), qui suggèrent qu'un haut degré de réalisme, la richesse de l'information, le nombre d'éléments interactifs, les facteurs de contrôle par l'utilisateur et les effets implicites de saillance perceptive peuvent mener à une forte distraction attentionnelle et à une certaine désorientation.

Par exemple, Baceviciute, Mottelson, Terkildsen & Makransky (2020), ont testé, l'efficacité du principe de modalité (Mayer & Moreno, 1998 ; Mayer, 2014) selon lequel, pour un document multimédia (texte accompagné d'images et/ou d'animation) il est préférable de présenter l'information verbale sous forme orale plutôt qu'écrite, dans un environnement en RVI. De nombreuses études, ainsi que des méta-analyses ont confirmé l'effet positif de l'application de ce principe dans l'apprentissage multimédia. Contrairement aux résultats de ces travaux précédents, Baceviciute & al. (2020) ont montré que dans le cadre d'un apprentissage en réalité virtuelle immersive, la présentation de l'information verbale sous forme de texte -commentaire-écrit permettait un apprentissage plus efficace en comparaison à la présentation de l'information verbale sous forme auditive (Figure 6.1).

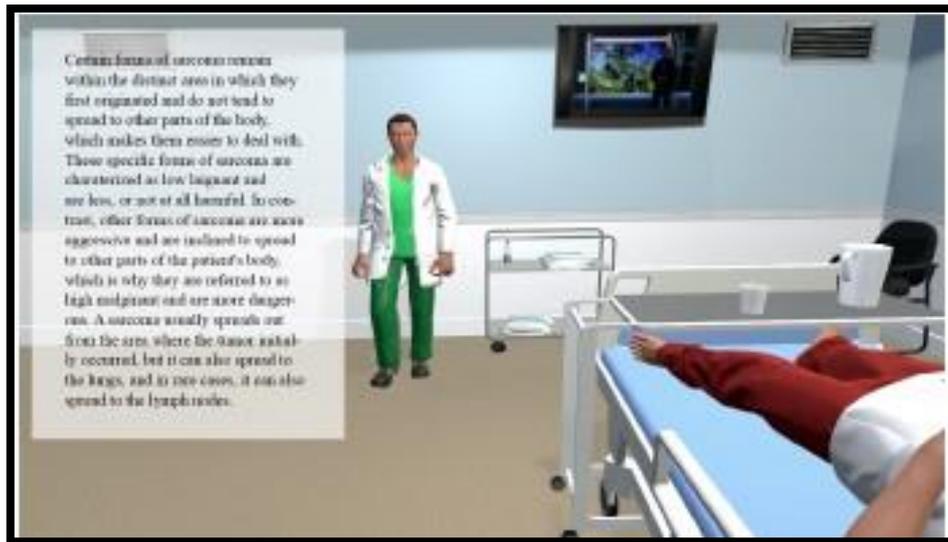


Figure 6.1. Exemple de l'environnement en RVI proposé par Baceviciute & al. (2020).

Il semble donc crucial de comprendre précisément les mécanismes perceptifs et cognitifs impliqués dans les apprentissages en RVI, qui semblent, au moins en partie, différents de ceux décrits dans les modèles classiques de l'apprentissage multimédia (cf. Mayet, 2014 ; Schnotz, 1999). Peu de travaux ont étudié les mécanismes cognitifs liés aux apprentissages à partir de la RVI.

Compte tenu de la charge visuelle, attentionnelle et cognitive induite par la RVI évoquée précédemment, nous proposons l'idée qu'un guidage visuel, spatio-temporel, de l'attention de l'apprenant vers les informations pertinentes par une technique de signalisation accompagnée d'un feedback pourrait soutenir et améliorer l'apprentissage et à la compréhension. Certes, « voir n'est pas forcément comprendre » (De Koning & al, 2010), mais l'utilisation de niveaux pertinents de signalisation, accompagnés de feedback pourrait optimiser la sélection, l'organisation et l'intégration cognitive des informations à apprendre. En particulier, la signalisation visuelle, dans le temps et l'espace de l'environnement dynamique, pourrait alléger la recherche des informations pertinentes, faciliter leur organisation interne et ainsi, favoriser l'intégration, c'est-à-dire la mise en relation des informations entre elles, et enfin, optimiser la construction d'un modèle mental de haute qualité, approprié et complet du contenu à apprendre.

6.1.3 La signalisation ou « cueing » en réalité virtuelle immersive.

Peu d'études ont été réalisées sur le thème du guidage par signalisation visuelle dans les apprentissages en réalité virtuelle, la plupart utilisent un guidage sous forme d'annotations

verbales ou ne proposent pas de protocole expérimental précis permettant une analyse statistique. On peut par exemple citer l'étude de Blome, Diefenbach, Bucher & von Mammen (2017) qui semble montrer un effet positif de la signalisation visuelle par l'utilisation de contrôleurs fantômes dans l'apprentissage des gestes de premier secours en RVI. Cependant, le nombre restreint de participants et l'absence d'analyses statistiques ne nous permet pas de conclure clairement à un avantage du guidage proposé. De plus, aucun lien n'est établi entre les observations réalisées et les processus cognitifs sous-jacents. Nous allons donc nous concentrer sur les quelques études utilisant une méthode expérimentale rigoureuse, faisant le lien entre les apprentissages en réalité virtuelle, la signalisation et les processus cognitifs sous-jacents.

Pour commencer, l'étude de Chen & Toh (2005) avait pour objectif de tester la signalisation visuelle en RVI sur le thème de la circulation routière. Les chercheurs ont proposé une version RVI dans laquelle l'apprenant bénéficiait d'indicateurs visuels de signalisation (un traceur indiquant en temps réel la position du véhicule sur une carte ainsi que des flèches directionnelles), un mode RVI sans ces éléments de guidage ainsi qu'un mode d'apprentissage traditionnel (sans RV). Leurs résultats montrent que la présence d'indicateurs visuels aidant l'apprenant à se repérer sur une carte offraient de meilleures performances d'apprentissage.

Albus, Vogt & Seufert (2021) ont quant à eux évalué l'effet de la signalisation sous forme d'annotations verbales en RVI dans une tâche d'apprentissage sur le thème de la désalinisation de l'eau de mer. Les auteurs ont observé un bénéfice de la modalité avec ajout d'annotations écrites à la tâche de rétention, mais les performances entre les deux types d'apprentissage, avec ou sans annotations verbales, sont équivalentes pour la tâche de transfert.

Ces études portent sur l'ajout de texte en RVI, en revanche, le guidage par signalement visuel, ou « cueing », (de Koning & al., 2010) à l'aide de couleurs ou encore de mise en lumière de certains éléments pertinents de l'environnement n'a pas encore été étudié dans ce contexte. Les recherches antérieures portant sur les documents multimédias plus classiques, montrent que le guidage visuel par signaux peut améliorer considérablement la sélection et la mise en relation des informations pertinentes. C'est par exemple le cas des travaux de Scheiter & Eitel (2010) et de Désiron, Bétrancourt & de Vries (2021) qui ont révélé que l'utilisation de légendes sur une illustration schématique respectant un code couleur reproduit dans le texte adjacent permettait d'extraire plus rapidement l'information pertinente et de créer un modèle mental plus cohérent du processus étudié que dans le cas d'une présentation classique (un texte accompagné d'une image sans légendes ni code couleur). Ainsi, les signaux qui indiquent la disposition du lien de causalité entre différents éléments permettent un meilleur transfert des connaissances

(Rey, 2010). De plus, l'utilisation de techniques permettant d'augmenter la saillance perceptive telle que la présentation d'une partie de l'image en couleur, peut permettre une meilleure rétention de l'information (Jamet & al., 2008), mais cette observation n'est pas systématique (Ozcelick & al., 2010 ; Doolittle & Altstaedter, 2009). Il est aussi important de rappeler que le guidage proposé doit être cohérent avec le matériel proposé (Boucheix & Lowe, 2010). Dans le cas de notre étude, nous avons utilisé deux types de guidages : un guidage sémantique, par l'ajout de points d'informations complémentaires et un guidage par signalement, par le clignotement successif des éléments clés (voir section matériel). S'il semble que l'ajout d'informations sous forme de texte soit bénéfique pour l'apprentissage en RVI, nous n'avons aucune indication quant à l'effet du guidage par signalement visuel.

6.1.4 Le rôle du feedback en réalité virtuelle immersive.

Le feedback consiste à donner à l'apprenant un retour informatif sur sa propre action ou une prise de décision pendant ou après l'apprentissage. Le feedback peut être plus ou moins informatif pour l'apprenant. Il peut s'agir simplement d'un retour sensoriel (comme par exemple une couleur telle que le vert, l'orange ou le rouge, d'un bruit comme un « ding » ou un « gong » etc.) pour indiquer dans quelle mesure la réponse donnée est juste ou fautive. Dans ce cas, le feedback ne fournit pas d'informations explicatives qui permettraient à l'apprenant d'orienter la modification de ses stratégies et/ou réponses futures. Dans d'autres recherches, les feedbacks peuvent fournir des informations visant à aider l'apprenant à modifier ses réponses et ses stratégies. Les rares travaux réalisés sur l'utilisation du feedback en réalité immersive ne se sont pas intéressés aux processus cognitifs sous-jacents à l'utilisation de cette méthode d'apprentissage. L'étude d'Eaves & Breslin (2011) semble indiquer un effet positif du feedback dans l'apprentissage moteur de positions en danse classique. Dans cette étude, l'apprenant était équipé de capteurs de mouvements et devait réaliser des mouvements de danse classique. Il disposait d'un feedback en temps réel. Ce feedback montrait le mouvement réalisé par un expert sur lequel étaient retranscrits des indicateurs (12 points ou 4 points) du mouvement de l'apprenant. Ici, l'apprentissage en lui-même n'est pas réalisé en réalité virtuelle. Seul le feedback, décrit comme étant un feedback RV en temps réel est employé (Figure 6.2). De plus, il s'agit d'un apprentissage procédural moteur, ce qui est très différent des apprentissages classiques et conceptuels que nous étudions.

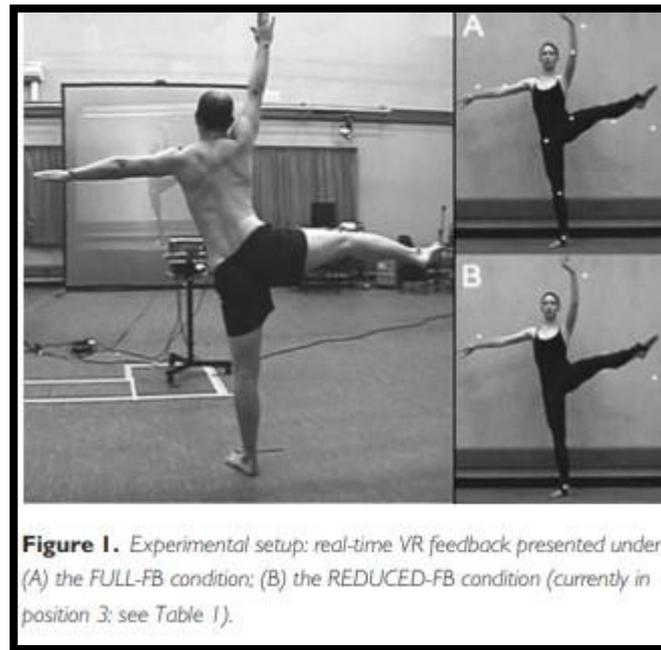


Figure 6.2. Feedback en temps réel du mouvement réalisé par le danseur, selon 12 ou 4 points.

L'étude de Kruglikova & al. (2009) avait pour objectif de mesurer l'impact d'un feedback de type constructif sur le thème de l'endoscopie gastro-intestinale, par l'utilisation d'un simulateur en RV haute-fidélité. Après une session d'entraînement à la manipulation du simulateur à l'aide de vidéos issues de ce simulateur, les apprenants ont réalisé plusieurs essais de simulation d'endoscopie. Deux groupes ont été déterminés au préalable : un groupe recevait un feedback sur sa pratique de l'endoscopie, l'autre groupe ne recevait pas de feedback. Le feedback consistait à illustrer l'état de l'intestin et celui du patient lors de l'acte médical. En cas de mauvaise pratique, l'intestin pouvait par exemple se perforer. Le patient virtuel pouvait quant à lui manifester de l'inconfort pendant l'intervention. Les résultats ont montré un effet positif du feedback constructif sur l'apprentissage comparé à un entraînement seul. Les participants du groupe « feedback » ont surpassé les participants du groupe entraînement seul concernant le temps de réalisation de la procédure et le temps mis pour atteindre l'objectif déterminé par l'acte médical demandé. Enfin, les apprenants du groupe feedback ont montré des performances supérieures au groupe sans feedback en test de rétention et de transfert différé (après un délai de 4 à 6 semaines).

D'autres études ont été réalisées sur le feedback en réalité virtuelle dans de nombreux domaines : l'industrie, la médecine ou encore les sciences de l'éducation mais ces recherches

ne portent pas sur les mécanismes cognitifs sous-jacents à l'apprentissage conceptuel. De plus, dans la plupart des cas la méthodologie n'est pas expérimentale, elle consiste à rendre les interactions entre l'utilisateur et l'environnement virtuel plus réelles. En effet, bien souvent les travaux réalisés portent sur la modélisation de feedbacks haptiques dans le but de rendre l'expérience en RVI plus fidèle à la réalité lorsque l'utilisateur interagit avec son environnement.

Nous nous baserons donc sur les résultats obtenus à ce sujet à partir de travaux réalisés lors de tâches d'apprentissage déclaratifs ou conceptuels comme par exemple l'étude de Mayer & Moreno (2005) qui montre que les apprenants auxquels on fournit un feedback explicatif sont plus à même d'expliquer correctement leur choix et de corriger leurs erreurs que ceux qui ne reçoivent pas de feedback explicatif. L'étude de Mayer et Johnson (2010) montre également que le type de feedback est important. Dans leurs travaux, les auteurs proposent soit un feedback sonore indiquant à l'apprenant si la réponse donnée est juste ou fautive soit, en plus du feedback sonore, un feedback explicatif est fourni. Leurs résultats ont révélé que le fait de fournir un feedback explicatif permettait une compréhension plus complète du dispositif étudié. Dans le cadre de nos travaux, nous avons choisi de proposer un feedback explicatif en fournissant des indications sur les éléments clés correspondant à l'environnement visualisé par l'apprenant (voir méthode).

Questions de recherche et hypothèses.

Le réalisme visuel qu'offre la réalité virtuelle permet une représentation plus fidèle du monde qui nous entoure, d'autant plus lorsqu'il s'agit d'un environnement complexe. La forêt est composée de nombreux éléments comme bien évidemment les arbres, mais aussi les arbustes, les herbacés, ou encore les animaux. Tous ces éléments sont en synergie et une mauvaise gestion de cet écosystème pourrait donner lieu à des conséquences importantes sur leur évolution. La forêt est un environnement qui évolue lentement, à l'échelle de plusieurs années voire plusieurs décennies. Il est donc difficile de se représenter ce qu'une action réalisée aujourd'hui peut avoir comme impact dans les années à venir. Nous nous sommes donc demandés quelle représentation les élèves se faisaient de l'évolution de cet environnement. C'est pourquoi nous leur avons proposé une tâche d'exploration de l'écosystème forestier selon plusieurs paramètres. La tâche de l'apprenant consistait en l'exploration de l'environnement forestier avec comme objectif l'aménagement d'une zone choisie de manière à préserver le plus possible l'écosystème forestier. Cette tâche permettait à la fois l'acquisition de connaissances

académiques et professionnelles. En effet, il s'agissait de reproduire une tâche qu'ils seraient susceptibles de réaliser dans l'exercice de leur futur métier dans le cadre d'une expérimentation.

Un examen approfondi des études antérieures montre que la réalité virtuelle immersive permet un plus grand sentiment de présence dans l'environnement exploré (Makransky & Petersen, 2021 ; Freina & al., 2016 ; Parong & Mayer, 2020 ; Makransky, Terkildsen & Mayer, 2019) ce qui peut être un avantage dans le cadre de notre recherche puisqu'il s'agit pour les apprenants de se représenter fidèlement l'écosystème forestier. Par contre, la complexité des informations présentées entraîne une augmentation de la charge cognitive informationnelle, ce qui peut perturber l'acquisition de connaissances lors de l'apprentissage (Sweller & al., 2011 ; Makransky, Terkildsen & Mayer, 2019, Parong & Mayer, 2020). Une telle augmentation de la charge cognitive en réalité virtuelle renvoie à la notion d'« overwhelming » (Lowe & Ploetzner, 2017) , c'est-à-dire que l'apprenant est submergé par un trop grand nombre d'informations, ce qui nuit à l'apprentissage. Si la motivation des élèves peut être accrue lors de l'exploration, nous pensons qu'il est également possible que la multiplicité des informations visuelles et verbales présentes dans l'environnement provoque une surcharge de leurs ressources cognitives, rendant l'extraction d'informations pertinentes plus difficile.

Nous avons donc choisi de proposer deux types de signalisation visuelle : une signalisation spatio-temporelle par clignotement d'éléments végétaux de la forêt (arbres, plantes herbacées), et une signalisation visuelle par l'implantation de points d'informations concernant des éléments de la forêt peu ou non directement visibles dans l'environnement « naturel » (comme le type de sol par exemple). Ainsi, nous pensons faciliter la détection, la sélection et la mémorisation des informations pertinentes présentes dans cet écosystème complexe (Jamet & al., 2008).

Les apprenants auprès desquels nous avons expérimenté étaient novices dans le type de tâche demandé, il était donc difficile pour eux de se représenter l'impact de leur action au sein de l'environnement forestier. De ce fait, nous avons proposé un feedback à la fois visuel et verbal qui regroupait les éléments clés de chaque zone explorée comme par exemple les différents types d'arbres présents dans l'environnement, le type de sol ou encore la présence d'animaux. Ce feedback était présenté après le premier choix de l'apprenant à la tâche demandée, suite à l'exploration de l'environnement. Les informations présentées lors du feedback permettaient de prendre connaissance de la présence d'éléments qui n'avaient peut-être pas été consultés pendant la phase d'exploration et/ou de créer une représentation plus globale de chaque zone forestière proposée. Ainsi, l'apprenant pouvait déterminer si le type de sol était approprié à

l'aménagement de structures (comme des tables et des bancs) et si la biodiversité était suffisamment modérée pour permettre cet aménagement dans l'objectif de préserver l'écosystème forestier. Nous pensons ainsi favoriser l'émergence d'un apprentissage permettant à l'élève de modifier sa réponse en cas d'erreur et de produire des explications d'une plus grande pertinence (Mayer & Johnson, 2010 ; Moreno & Mayer, 2005)

Hypothèses.

Notre première hypothèse **H1a**, prédit un effet principal de la modalité d'exploration sur le choix réalisé par l'apprenant. Ainsi, nous pensons que les apprenants explorant l'écosystème de manière guidée, quel que soit le type de signalisation, choisiront plus souvent la zone de forêt la plus appropriée que ceux ayant réalisé une exploration libre. Aussi, nous pensons que les élèves ayant bénéficié de la présence de balises informatives auront de meilleures performances que les élèves ayant bénéficié d'une exploration guidée par flashing (**H1b**).

Notre deuxième hypothèse, **H2** prédit un effet principal du feedback. Nous pensons que la présentation du feedback permettra aux élèves de détecter leurs erreurs et de modifier leur choix, ou de confirmer leur premier choix dans le cas où ils ne se soient pas trompés (choix 2 > choix 1). Nous pensons également que le score obtenu après le feedback sera supérieur pour les élèves ayant bénéficié de la signalisation, comparé aux élèves ayant exploré librement l'environnement : Flashing > Balises > libre (**H2a**). Nous prédisons également une interaction entre la signalisation et le feedback, c'est-à-dire entre la signalisation et le deuxième choix réalisé par l'apprenant (**H2b**).

Notre troisième hypothèse, **H3**, prédit que le nombre d'arguments produit pendant l'exploration sera identique quelle que soit la modalité d'exploration. Nous pensons également que la nature des arguments (vrais, faux, partiels) sera identique dans toutes les conditions d'exploration (**H3a**). Cependant, nous pensons que le nombre d'arguments centrés sur la préservation de l'écosystème sera différent selon le type d'exploration : le pourcentage d'arguments centrés sur la préservation de l'écosystème sera plus important dans les conditions Flashing et balises que dans la condition libre (**H3b**).

6.2 Méthode.

6.2.1 Participants.

Quatre-vingt-neuf élèves de classes de première et terminale de Sciences et Technologies de l'Agronomie et du Vivant (STAV) issus d'un lycée français agricole ont participé à cette étude. Leurs enseignements étaient composés d'un tronc commun comprenant les matières générales (français, mathématiques, SVT, anglais, etc.) et de deux options au choix : technologie de la production agricole ou aménagement et valorisation des espaces. Quatre classes ont été sollicitées pour cette expérimentation à raison de deux classes par section : deux classes de premières et deux classes de terminales. L'expérience comprenait deux sessions espacées d'une à trois semaines : (i) une session de pré-tests (ii) une phase d'exploration suivie immédiatement d'une session de post-test. Nous avons retenu tous les élèves ayant participé à la phase expérimentale en RVI (âge moyen : 16.19 ans, 39 filles et 50 garçons, respectivement 43,82% et 56.18%). En revanche, en raison de l'absence de quelques élèves à la première session de prétest (en liaison avec les confinements occasionnés par la crise sanitaire du COVID19), seuls 79 élèves ont complété les tâches des deux sessions (âge moyen = 16.15 ans, 37 filles et 42 garçons, respectivement : 46.84% et 53.16%), et 89 ont pu participer à la deuxième session, partie principale de l'expérimentation après les pré-tests.

Les prétests se sont déroulés en salle de classe et les tests dans deux salles spécifiques dédiées à la réalité virtuelle. Au cours de la première session, les élèves ont répondu à une série de prétests de mémoire de travail, d'habiletés spatiales ainsi qu'à un questionnaire de connaissances préalables portant sur le sujet de l'exploration en RVI à savoir : l'écosystème forestier. Les performances des élèves à ces tests ainsi que leur moyenne générale académique ont été prises en compte, et utilisées pour créer quatre groupes homogènes (avec N = 20 participants par groupe) correspondant aux quatre conditions d'exploration de la forêt en RVI. Une à trois semaines plus tard, lors de la deuxième session, les élèves ont exploré individuellement un espace forestier en réalité virtuelle à l'aide d'un casque RV (HTC Vive). Cette exploration destinée à l'analyse de l'espace forestier était suivie d'une tâche de décision puis d'un feedback et se terminait par une seconde tâche de décision. Enfin, chaque participant remplissait un questionnaire démographique. Chaque session durait environ 60 minutes, et la durée totale de l'expérience était de 120 minutes. La première session se déroulait en classe entière, et la deuxième se déroulait sous forme de passations individuelles dans une salle dédiée à l'expérimentation. A la fin de la première session, les élèves étaient informés qu'ils

participeraient à une deuxième session en réalité virtuelle immersive, mais ils n'ont pas été informés de la nature de la tâche qu'ils devraient réaliser.

6.2.2 Matériel expérimental.

6.2.2.1 Prétests utilisés lors de la session I en vue de l'homogénéisation des groupes expérimentaux.

Lors de la session de prétests, les participants ont complété trois tests, visant à la constitution de groupes expérimentaux homogènes : (i) un test d'orientation spatiale ; (ii) un test de mémoire de travail verbale, et un test de connaissances préalables (iii).

Lorsqu'on est immergé dans un environnement en RV, il peut parfois être difficile de s'orienter dans l'espace, c'est pourquoi nous avons proposé un test d'orientation spatiale aux participants. *Le test d'orientation spatiale (Annexe H)* était une traduction française du test « Perspective Taking/spatial Orientation Test » développé par Hegarty, Kozhevnikov et Waller (2008). Ce test était composé de 12 items organisés de la manière suivante (Figure 6.3) : divers objets sont répartis sur une demi page comme une voiture, une maison, un chat ou une fleur. Sous ces images, une consigne est donnée à l'élève. On lui demande de se situer à la place d'une des images, de regarder une autre image, et à partir de là, ils doivent déterminer où se trouve un troisième élément. Pour répondre, le participant a pour tâche de tracer un trait dans un cadran circulaire situé sous la consigne. L'élève dispose de 5 minutes pour compléter les douze items.

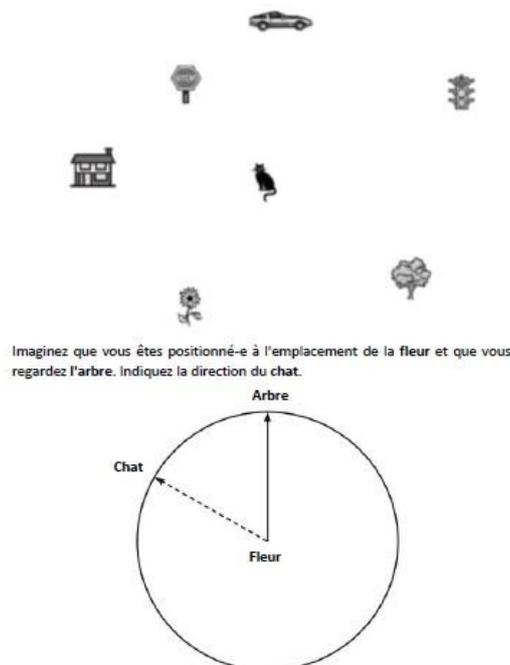


Figure 6.3. Item d'exemple du test d'orientation spatiale.

Le test de mémoire de travail verbale proposé était celui de Danneman & Carpenter (1980), adapté en français par Desmette, Hupet, Schelstraete & Van Der Linden (1995).

Dans ce test (séries 1 & 2 disponibles en Annexe I), les participants doivent écouter une série de phrases présentées oralement. A la fin de chaque série de phrases, ils doivent stocker et rappeler le dernier mot de chacune des phrases de la série. Chaque série comprend 2, 3, 4, 5 ou 6 phrases. Le participant doit alors écouter les phrases et en même temps garder le dernier mot de chaque phrase, actif en mémoire de travail. Les participants devaient écouter la série de phrases, et ce n'est qu'après avoir reçu le signal « go » qu'ils devaient écrire le dernier mot de chaque phrase dans le cahier de réponses. Deux scores ont été calculés à partir des réponses de l'élève : le critère le plus exigeant pour noter les réponses était le nombre de séries de phrases complètement réussies (sans erreurs ni mots manquants, sur 15), et le critère le moins exigeant était le nombre de mots correctement et complètement rappelés dans le bon ordre dans toutes les séries de phrases (max. 40). Les deux scores ont été transformés en pourcentages et nous avons utilisé la moyenne des deux scores pour les analyses de données.

Le prétest de connaissances factuelles était composé de deux questions ouvertes :

- Selon toi, qu'est-ce qu'un écosystème forestier ?
- Selon toi, quelle est la meilleure manière de gérer une forêt ?

Nous avons choisi ce type de question, premièrement parce qu'elles renvoient aux questions habituellement posées par les enseignants aux élèves, et également pour ne pas influencer les réponses données lors de la phase expérimentale. En effet, proposer un QCM englobant la biodiversité, les caractéristiques pédologiques et d'aménagement de l'environnement forestier aurait pu mener les élèves à produire un raisonnement et des réponses différentes lors de l'exploration en RVI. De plus, ce type de tâche aurait un pouvoir plus discriminant qu'un questionnaire de type QCM (Mayer & Fiorella, 2021).

Pour coter ce test, nous nous sommes basés sur les réponses fournies par l'enseignante de SVT à ces questions. Nous avons relevé les 10 critères principaux fournis par l'enseignante : la diversité, l'intervention humaine, les êtres vivants adaptés (adaptation), le système dynamique, le biotope, l'interaction entre les espèces, le cycle de renouvellement, le milieu stable, la pérennité du milieu et le règne. Chaque critère était noté sur 1 de la manière suivante : nous avons accordé 0.5 points si le critère était évoqué par l'élève de manière peu développée ou

imprécise et 1 point s'il employait correctement ce critère. Nous avons alors obtenu un score sur 10 que nous avons traduit en pourcentage pour l'analyse statistique.

Ce test était suivi de questions démographiques qui comprenaient des questions sur leur âge, leur sexe, l'utilisation des ordinateurs et des outils numériques tels que les réseaux sociaux et les jeux vidéo. Nous leur avons également demandé s'ils portaient des lunettes et s'ils avaient connaissances de difficultés d'apprentissage.

6.2.2.2 Matériel d'apprentissage.

Pour cette étude, nous avons utilisé un casque de réalité virtuelle HTC vive (voir Figure 6.4).



Figure 6.4. Casque de réalité virtuelle, manettes et capteurs (boîtiers cubiques) utilisés pour notre expérience.

Pour utiliser ce casque RV, nous avons délimité une aire d'exploration de 9m^2 à l'aide de capteurs (boîtiers cubiques présents sur la Figure 6.4) installés sur des trépieds à une hauteur de 2m. Ces deux capteurs étaient disposés en diagonale dans la salle, chacun marquant le coin d'un carré de 3mètres de côté. Cette surface a été paramétrée et calibrée à l'aide des manettes sur un ordinateur « gaming » sous Windows 10 au moyen du logiciel SteamVR (voir Figure 6.5).

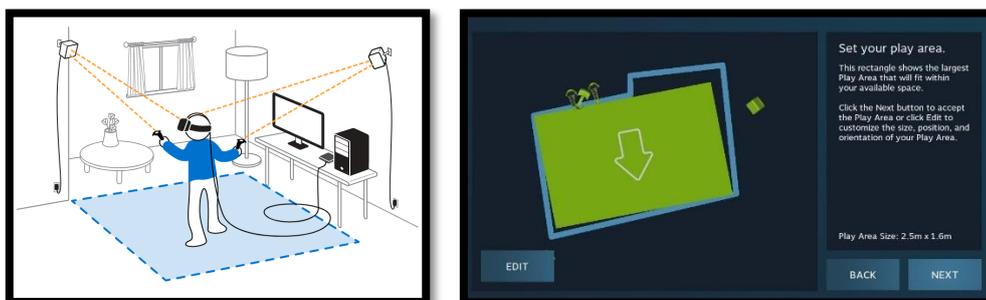


Figure 6.5. Illustration d'une zone d'exploration RV. A gauche : aire délimitée par les capteurs. A droite : aperçu de l'aire calibrée sur le logiciel Steam VR.

Cette aire d'action permet à l'utilisateur de se déplacer sans risque dans une zone dépourvue d'obstacles. Pour indiquer la limite de la zone d'exploration, une grille bleue apparaissait dans le casque lorsque l'utilisateur s'en approchait (Figure 6.6). De cette manière, il pouvait se réorienter de manière à rester dans la zone délimitée en toute sécurité, tout au long de son exploration.

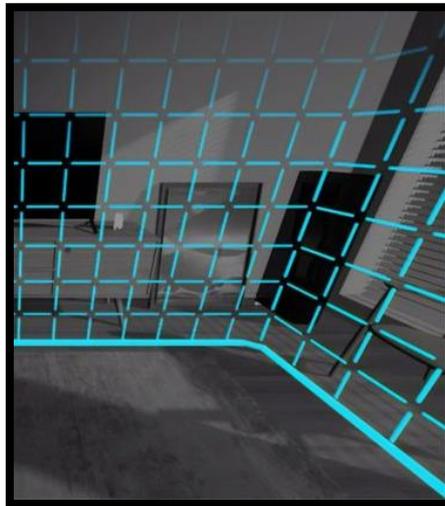


Figure 6.6. Illustration de la grille perçue dans le casque RV lorsque l'utilisateur atteint la limite de la zone déterminée.

L'environnement proposait une reproduction de parcelles de forêts typiques que l'on pourrait retrouver dans un environnement forestier réel en France (Figure 6.7).

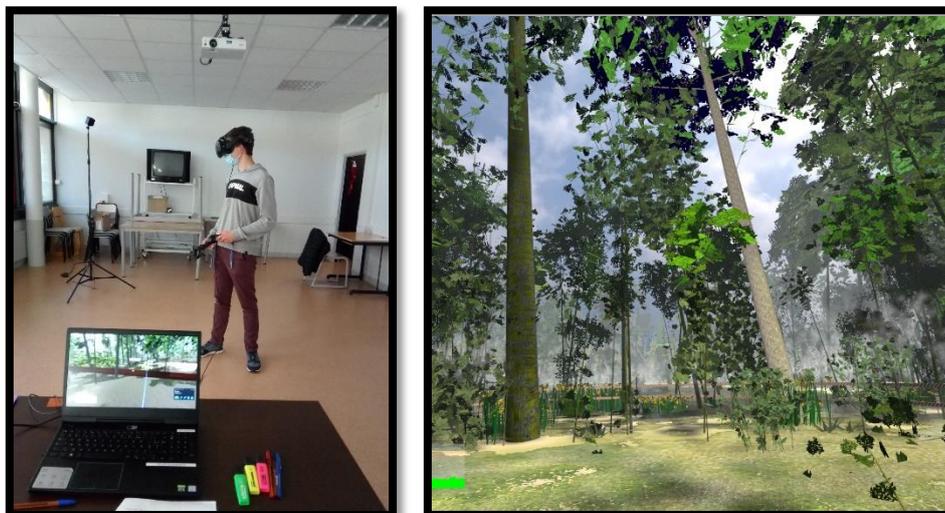


Figure 6.7. Exemple de l'environnement virtuel.

Le but était d'immerger l'apprenant dans cette forêt pour qu'il puisse, à partir des différents éléments proposés (arbres, herbacés, micro-habitats...), proposer un diagnostic de l'espace

forestier autorisant ou non l'installation de structures destinées à l'accueil de promeneurs. Leur diagnostic était construit au cours de la tâche d'exploration et d'analyse des trois zones forestières par l'extraction -le marquage- et la mise en lien d'éléments pertinents, aboutissant à une prise de décision favorable ou non à l'installation d'une zone d'accueil du public, de type observatoire de la forêt avec espace de pique-nique. Cette tâche répondait à nos besoins en termes de recherche sur les processus cognitifs liés aux apprentissages en réalité virtuelle immersive, et également aux besoins pédagogiques des apprenants dans leur pratique de terrain en lien avec leur future activité professionnelle.

Deux expérimentateurs accueillaient les élèves individuellement, chacun dans une salle dédiée spécialement à l'expérimentation en réalité virtuelle au sein du lycée. Chaque salle était équipée d'un ordinateur doté du logiciel expérimental et d'un casque RV HTC Vive.

Le didacticiel : il était composé de trois zones délimitées par des barrières. Chacune des zones comportait une balise et un herbacé. Le but était de permettre à l'élève de se familiariser à l'environnement en RV, aux commandes permises par les manettes, sans être directement immergé dans la forêt.

*L'environnement virtuel*⁷ : il était composé de trois zones au sein d'une forêt virtuelle. La tâche de l'élève était d'analyser ces trois zones en consultant les différents éléments présents dans chaque zone (arbres, arbustes, herbacés) de manière à en extraire les caractéristiques spécifiques. L'élève pouvait se déplacer en se téléportant dans l'environnement à l'aide des manettes. Il pouvait également consulter et marquer, ou démarquer des éléments, qui prenaient alors une couleur rouge (Figure 6.8).

⁷ Environnement virtuel en RVI développé par Louis Rapet et conçu par Rémi Malin, sous la supervision de Jean-Luc Martinez ; AMVALOR, Institut Image, Chalon sur Saône.

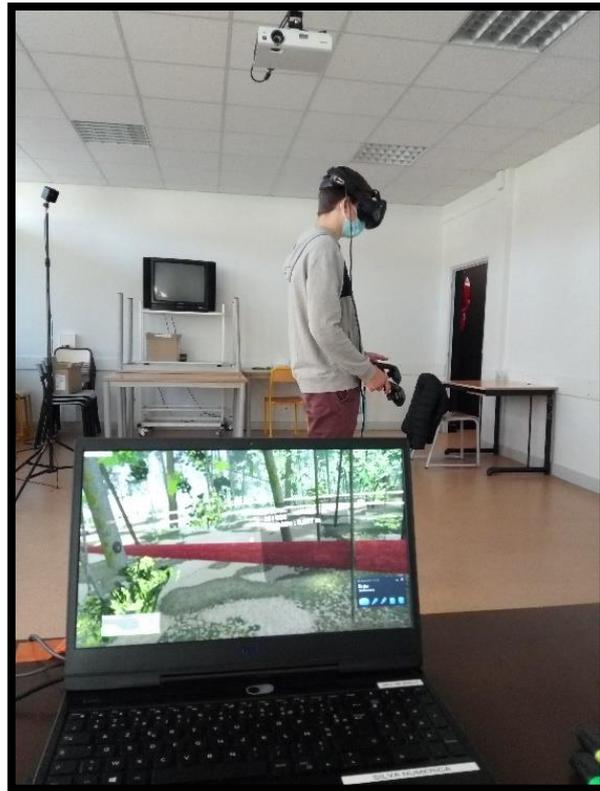


Figure 6.8. Exemple d'élément marqué (rouge), à l'écran pendant l'exploration.

Chaque élément consulté, marqué ou démarqué était enregistré automatiquement par l'ordinateur. Concernant la disposition des différentes zones, une zone humide était implantée proche d'un plan d'eau et était composée d'éléments spécifiques à un sol marécageux comme par exemple des iris, des fougères, des hêtres et des ormes. Une seconde zone était éloignée du plan d'eau et présentait un sol stable avec la présence modérée de biodiversité : de l'aubépine, de l'euphorbe, des hêtres, des érables et une majorité de chênes. La troisième zone était située entre les deux autres zones et présentait un sol stable et une biodiversité riche : de l'aubépine, de l'euphorbe, des érables, des frênes, des charmes, des hêtres, des chênes et des épicéas. On pouvait retrouver des micro habitats dans les trois zones : des champignons polypores, des cavités de pic, des troncs morts au sol etc. Ainsi, le niveau de biodiversité était malgré tout élevé pour chacune des zones, évitant que la tâche ne soit trop simple. Cette diversité impliquait une analyse approfondie des éléments présents dans chaque zone, afin d'élaborer une décision raisonnée suite à l'exploration complète de l'environnement.

Chaque zone était délimitée par des barrières, et seul les éléments situés à l'intérieur de ces trois zones étaient consultables. Chaque zone était composée d'environ 30 arbres, 15 arbustes et 20 herbacés dont la densité (y compris visuelle) était similaire dans le but de ne pas influencer

la réponse de l'élève. Chaque zone était générée par un logiciel reproduisant l'équilibre forestier réel spécifique au type de zone créé.

A tout moment de l'exploration des zones de la forêt, l'élève avait la possibilité d'accéder à une carte via la gâchette de la manette afin de pouvoir se repérer dans l'environnement virtuel (Figure 6.9). Une flèche (en gris dans la Figure 6.8) indiquait la position et l'orientation de l'élève en temps réel, l'aidant ainsi à se repérer et à se déplacer dans l'environnement. Chaque zone portait un numéro permettant à l'élève de se repérer et de communiquer son choix par la suite. Trois types de numérotations ont été programmés et contrebalancés, tout comme l'ordre d'exploration des trois zones.

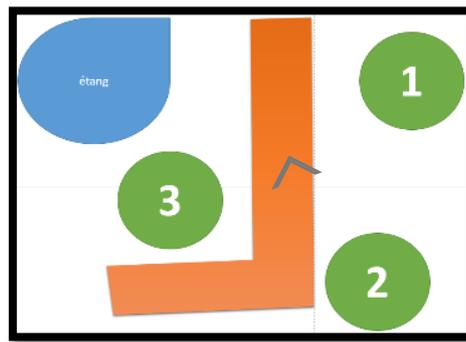


Figure 6.9. Reproduction de la carte accessible dans l'environnement virtuel, comportant un point d'eau (en bleu), les trois zones (en vert) et un chemin (en orange).

Les zones étaient toutes trois composées d'arbres, d'arbustes et d'herbacés. Chaque élément pouvait être visé et consulté à l'aide des manettes. Une fois l'élément visé avec la manette, une infobulle donnait alors les informations suivantes : type d'élément (Arbre/arbuste/herbacée) ainsi que le nom de l'élément correspondant. Pour les arbres étaient donnés : l'essence (le type d'arbre, comme par exemple chêne, hêtre, aulne etc.), l'âge, le diamètre de l'arbre, la présence éventuelle de branches mortes dans le houppier, de micro habitats, de cavités de pics, de champignons polypore ou encore de cavité de pied.

Dans le mode signalisation par flashing (Figure 6.10), des éléments clignotaient successivement ; ils devenaient plus sombres puis plus clairs jusqu'à ce que l'élève les ait consultés. Une fois consulté, l'élément arrêta de clignoter et l'élève devait chercher l'élément clignotant suivant.

Dans le mode signalisation par balise : trois balises étaient disposées à des endroits spécifiques de chaque zone (Figure 6.10). Ces balises fournissaient des informations complémentaires cruciales sur l'environnement, et qui n'étaient pas directement visibles, comme par exemple :

sol stable, sol humide, présence de grenouilles, de libellules et de couleuvres etc. Les informations complètes fournies par ces balises sont disponibles en Annexe J.



Figure 6.10. Balise et exemple de clignotement (à droite).

Deux niveaux de signalisation ont été créés, correspondant à deux facteurs permettant de créer quatre groupes expérimentaux selon un plan 2x2, Tableau 6.1:

Tableau 6.1.

Nombre d'élèves par groupe expérimental selon le type de signalisation.

	Balises présentes : B+	Balises absentes : B-
Flashing présent : F+	n=23	n=21
Flashing absent : F-	n= 22	n= 23

Les verbalisations concomitantes des élèves étaient enregistrées à l'aide d'un dictaphone pendant toute la phase expérimentale. Ces verbalisations concomitantes ont été retranscrites et les arguments donnés pour chaque élément marqué ont été répertoriés. Nous avons relevé le nombre total d'éléments marqués par l'élève au cours de son exploration. Pour chaque argument donné, nous avons attribué un point à la (ou les) catégorie(s) qui lui correspondaient : vrai, faux, partiel, absence d'argument. Un argument était considéré comme vrai lorsque l'explication correspondant à l'élément était correcte et comme faux dans le cas inverse. Un argument était considéré comme partiel lorsque l'élève fournissait un argument incomplet ou peu précis. On notait une absence d'argument lorsque l'élève ne savait pas justifier son choix lors du marquage de l'élément, par exemple : « *je sais que c'est important, mais je ne sais pas expliquer pourquoi* ». Nous avons également relevé le nombre d'arguments donnés en faveur de la préservation de l'écosystème et le nombre d'arguments donnés en faveur du bien être humain

dans l'environnement. Chaque catégorie a été transformée en pourcentage pour l'analyse statistique.

6.2.2.3 Matériel du questionnaire final.

Le *post-test* final (Figure 6.11) était un questionnaire concernant l'appréciation de l'expérience vécue en réalité virtuelle (la totalité du test est disponible en Annexe K). Il était composé de 10 affirmations que le participant devait évaluer en sélectionnant un chiffre sur une échelle de 1 à 5. Enfin, un espace était dédié aux remarques éventuelles des élèves : « Remarques éventuelles sur l'utilisation du programme et de la réalité virtuelle : qualités/défauts éventuels ; autres : ».

1- L'interface graphique est réaliste et facile à interpréter.

1 2 3 4 5
Pas du tout Légèrement Modérément Beaucoup Absolument

2- L'environnement virtuel pris dans sa globalité m'a permis de me sentir immergé dans cette expérience.

1 2 3 4 5
Pas du tout Légèrement Modérément Beaucoup Absolument

Figure 6.11. Exemple d'items au post-test RV.

Le feedback (Figure 6.12) : il était composé de trois pages au format A4 à raison d'une page par zone de forêt. Chaque page était composée d'une capture d'écran de la zone spécifique et d'une description synthétique regroupant les caractéristiques principales relatives à cette zone (entre 46 et 50 mots). Le matériel complet est disponible en Annexe L. Ce feedback était présenté suite à la première décision donnée par l'apprenant dans le but de lui permettre de confirmer, ou de modifier son choix. Le court texte explicatif présenté conjointement à l'image permettait à l'apprenant de se remémorer l'exploration et de prendre connaissance de la présence d'éléments qu'il n'avait peut-être pas remarqués lors de son exploration. L'apprenant pouvait alors comparer les différentes zones entre elles et interpréter le rôle de nouveaux éléments potentiels avec ceux relevés pendant l'exploration RVI.

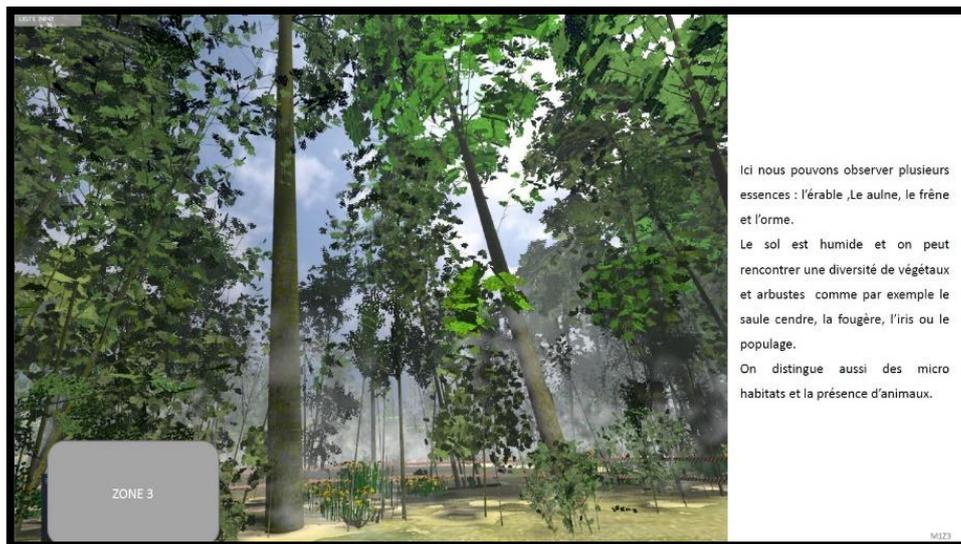


Figure 6.12. Exemple de feedback donné à l'élève

6.2.3 Procédure.

La première séance a duré 60 minutes. Les participants ont effectué les trois prétests : un test de mémoire de travail verbale, un test d'orientation spatiale et un test de connaissances préalables. Ils devaient également remplir un questionnaire démographique qui comprenait des questions sur leur âge, leur sexe, l'utilisation des ordinateurs et des outils numériques tels que les réseaux sociaux et les jeux vidéo. On leur a également demandé s'ils portaient des lunettes et s'ils avaient connaissances de difficultés d'apprentissage. Les tests étaient réalisés en classe entière à l'aide de livrets individuels spécifiques, sous la supervision de trois expérimentateurs et d'un enseignant. A la fin de la séance, les élèves ont été informés qu'ils seraient invités à une deuxième session au cours de laquelle différentes activités leur seraient proposées.

La deuxième séance a eu lieu une à trois semaines après la première séance et a également duré 60 minutes. Les élèves étaient reçus individuellement par l'un des deux expérimentateurs. Les sessions des deux expérimentateurs se déroulaient en parallèle. L'expérimentateur expliquait le déroulement général de l'expérimentation à l'élève puis l'équipait du casque de réalité virtuelle (HTC Vive). Pour commencer, un didacticiel était proposé dans lequel trois zones épurées étaient délimitées par des barrières. Chacune des zones comportait un arbuste et une balise. On indiquait les commandes des manettes (Annexe M) aux élèves à qui on demandait ensuite : de consulter la carte, de se téléporter, de consulter un élément, de le marquer, puis de le démarquer. Quand l'apprenant était confiant dans l'utilisation du matériel, on lui demandait de se téléporter dans une zone spécifique du didacticiel. Cette étape durait entre 5 et 10 minutes selon l'aisance

du participant dans la manipulation du matériel. On lui donnait ensuite les consignes (disponibles en Annexe N) : l'élève devait explorer les trois zones et déterminer, à la fin de l'exploration dans laquelle des trois zones il choisirait d'implanter une aire de pique-nique en respectant le plus possible l'écosystème forestier. La tâche de l'élève était d'explorer chaque zone et de marquer les éléments qui lui semblaient importants. A chaque fois, il devait justifier son choix. Les élèves disposaient de 6 min pour explorer chaque zone. Une fois les 6 minutes écoulées, l'expérimentateur leur indiquait dans quelle zone se rendre.

Critères de codage de la décision.

Après avoir exploré chacune des zones, l'élève retirait le casque de réalité virtuelle. On lui demandait alors de dire dans quelle zone il choisissait d'implanter l'aire de pique-nique et d'expliquer verbalement son choix. Il s'installait ensuite à une table où on lui présentait le feedback (Annexe L).

L'expérimentateur fournissait une série d'images accompagnées du texte explicatif en lui indiquant la zone choisie. Le participant pouvait alors consulter les images/textes du feedback (Figure 6.12 et Annexe L). L'élève avait ensuite le choix de confirmer son choix, ou de changer d'avis. Dans les deux cas, il devait justifier son raisonnement. Pour finir, on lui présentait le test RV afin de recueillir son ressenti sur l'expérience à laquelle il venait de participer (Annexe K). L'expérimentateur restait disponible pour répondre aux questions des élèves à la fin de l'expérimentation, et leur indiquait de ne pas communiquer à propos de l'expérience auprès de leurs camarades de classe.

Critères du codage de décision.

Après l'exploration de l'environnement RVI, l'élève nous indiquait la zone choisie pour l'installation de l'aire de pique-nique. Nous avons attribué un score compris entre 0 et 1 à cette décision selon les caractéristiques suivantes :

1 point était attribué lorsque le choix de la zone ainsi que la justification correspondante étaient corrects.

0.5 points étaient attribués lorsque le choix était correct et que la justification associée était partiellement correcte.

0 points quand la zone choisie n'était pas la bonne ou lorsque la zone choisie était la bonne, mais que la justification produite par l'élève était totalement fautive.

Critères de codage de données de justification verbale.

Pour les besoins de notre analyse de données, nous avons relevé le nombre total d'arguments donnés par les élèves au cours de l'exploration de la forêt virtuelle, ainsi que le nombre d'arguments donnés lors de leur premier et second choix. Pour chaque argument, nous avons déterminé s'il était : juste, faux, partiel ou absent. Nous avons également relevé le nombre d'arguments en référence à la protection de l'écosystème et le nombre d'arguments en faveur du bien-être humain au sein de l'environnement forestier. Ainsi, chaque argument correspondait à un point dans l'une des catégories énoncées pour : l'exploration des zones, le premier choix (décision transmise immédiatement après l'exploration) et le deuxième choix (décision transmise après consultation du feedback). Chaque catégorie d'argument a été transformée en pourcentage par rapport au nombre total d'arguments donnés par l'apprenant dans la catégorie correspondante.

Par exemple :

« *Il y a des bactéries et des champignons donc ce n'est pas spécialement bien pour le corps humain, c'est pas très hygiénique.* » Ici, l'argument a été coté comme faux et également en faveur du bien être humain dans l'écosystème forestier.

« *Je ne sais pas si c'est favorable ou pas mais il y a des micro habitats quoi, donc si après la présence d'humain et tout ça enlève tout heu... c'est important de le préserver.* ». Cet argument a été coté comme étant partiel, en faveur de la protection de l'écosystème forestier.

« *Le houx ça sert pas à grand-chose donc on l'enlève.* » cet argument a été coté comme faux (le houx est une espèce protégée).

Balise sol stable : « *bah la structure du sol est bonne. Il n'y a pas de risque niveau effondrement du sol, ou de tassement* ». Ici l'argument a été coté comme vrai.

Analyse des données.

Des ANOVAs intergroupe et à mesure répétées ont été réalisées pour analyser les données. Les modalités de guidage, balise et flashing, constituaient deux facteurs inter-sujets. Dans les différentes ANOVAs, les variables dépendantes étaient les suivantes : (i) exactitude de la réponse donnée en premier choix (avant le feedback) ; (ii) exactitude de la réponse donnée en second choix (après le feedback). Comme la variable dépendante exactitude des décisions

des choix 1 et 2 variaient de 0 à 1, en complément, et à des fins de vérification, nous avons confirmé les résultats principaux des ANOVAs par des ANOVAs non paramétriques, par rang, de Kruskal-Wallis. Pour toutes les analyses tout au long de cette étude, nous avons utilisé $p < 0.05$ comme critère de significativité. Les valeurs de l'êta partiel au carré (η_p^2) et du d de Cohen sont fournies en tant que mesures de taille d'effet pour tous les effets principaux, les interactions et les comparaisons post-hoc. Les valeurs η_p^2 de 0,01, 0,06 et 0,14 représentent respectivement des tailles d'effet petite, modérée et grande (Clark-Carter, 1997), et les valeurs d de 0,20, 0,50, 0,80 et 1,3 représentent respectivement les tailles d'effet petite, moyenne, grande et très grande.

6.3 Résultats.

Nous présenterons dans un premier temps les résultats obtenus aux prétests. Nous présenterons ensuite les scores obtenus concernant le premier choix réalisé suite à l'exploration puis nous analyserons l'effet du feedback. Pour finir, nous étudierons le type d'arguments produits par les élèves pendant leur exploration.

6.3.1 Contrôle d'homogénéité du groupe.

Afin de vérifier l'homogénéité des groupes dans chaque condition, des ANOVAs à un facteur ont été effectuées, avec la condition d'exploration de l'environnement comme facteur inter-sujet, et les scores au test de mémoire de travail verbale, au test d'orientation spatiale et au test de connaissances préalables comme variables dépendantes. Les performances au test de mémoire de travail verbale, au test d'orientation spatiale, au test de connaissances préalables ainsi que les moyennes scolaires des élèves sont présentés dans le tableau 6.2.

Tableau 6.2.

Moyennes et écarts types obtenus au prétest de mémoire de travail verbale, d'orientation spatiale, de connaissances préalables et moyennes générales scolaires pour chaque modalité.

	Aucune signalisation (F- B-) N = 19	Balises seules (F- B+) N = 20	Flashing + balises (F+ B+) N = 22	Flashing seul (F+ B-) N = 18
MdT verbale, nb total de mots rappelés /40	33.58 (3.73)	33.21 (3.91)	34.45 (3.43)	34.40 (4.31)
Orientation spatiale , score /12	6.63 (2.71)	7.05 (2.41)	6.41 (3.40)	6.45 (3.65)
Connaissances préalables %	15.22 (6.82)	18.41 (8.64)	18.48 (10.27)	15.00 (5.00)
Moyennes générales scolaires /20	12.62 (1.78)	11.94 (1.56)	11.71 (1.94)	12.18 (1.40)

Nous avons réalisé des ANOVAs à un facteur avec comme facteur inter-sujet la modalité expérimentale et comme variable dépendante, respectivement, le score de mots rappelés sur 40 au test de mémoire de travail verbale, le score sur 12 au test d'orientation spatiale, le pourcentage de bonnes réponses au test de connaissances préalables et les moyennes générales scolaires.

L'ANOVA n'a montré aucun effet des conditions d'exploration de l'environnement sur les scores de mémoire de travail : $F(3, 76) = .50942$, $p = .67697$, $\eta_p^2 = 0.019$. Aucune des comparaisons post-hoc entre les conditions n'était significative. De même, l'ANOVA n'a montré aucun effet des conditions d'exploration de l'environnement sur les scores au test d'orientation spatiale : $F(3, 76) = .18$, $p = .91$, $\eta_p^2 = 0.007$. Aucune des comparaisons post-hoc entre les conditions n'était significative. Également, L'ANOVA n'a montré aucun effet des conditions d'exploration de l'environnement sur les scores de connaissances préalables : $F(3, 85) = 1.29$, $p = .28$, $\eta_p^2 = 0.04$. Aucune comparaison post-hoc entre les conditions n'était significative. Pour finir, l'ANOVA n'a montré aucun effet des conditions d'exploration de l'environnement sur les moyennes scolaires : $F(3, 85) = 1.23$, $p = .31$, $\eta_p^2 = 0.04$. Aucune

comparaison post-hoc entre les conditions n'était significative. En somme, les groupes expérimentaux ont été considérés comme homogènes en ce qui concerne les tests de mémoire de travail verbale, d'orientation spatiale, de connaissances préalables et concernant les moyennes scolaires.

Analyse de l'effet des niveaux de signalisation sur la sélection et le marquage des éléments forestiers durant l'analyse des zones.

Afin d'analyser les effets des deux types de signalisation (flashing et balises) sur la sélection et le marquage des éléments des zones forestières nous avons réalisé des ANOVAs à deux facteurs avec comme facteur intergroupes, respectivement balise et flashing. Les nombres moyens d'éléments observés (visés avec les manettes) et marqués sont présentés tableau 6.3.

Tableau 6.3.

Nombre moyens d'éléments observés et marqués pour l'ensemble des trois zones selon chaque modalité d'exploration de l'environnement forestier.

	Aucune signalisation (F- B-) n = 23	Balises seules (F- B+) n = 22	Flashing + balises (F+ B+) n = 23	Flashing seul (F+ B-) n = 21
Nombre moyen d'éléments observé (visés) pour les 3 zones.	83.37 (70.42)	85.51 (60.73)	50.89 (29.25)	35.98 (13.87)
Nombre moyen d'éléments marqués pour les 3 zones.	5.80 (2.66)	5.66 (3.91)	6.26 (2.28)	6.85 (4.62)

Une ANOVA factorielle incluant les deux facteurs intergroupes, respectivement balise et flashing effectuée sur le nombre moyen d'éléments observés a montré un effet significatif du flashing, $F(1,85) = 13.09, p < .001, \eta_p^2 = 0.15$, mais l'effet de la présence de balises n'était pas significatif, $F(1,85) = 0.57, p = .45, \eta_p^2 = 0.007$. Il apparaît donc que la signalisation, en particulier la présence de flashing, dirige l'attention des élèves vers les éléments signalés pertinents pour la tâche et diminue le nombre des éléments observés par les élèves. En revanche,

l'ANOVA factorielle réalisée sur les éléments finalement marqués, ne révèle pas d'effet significatif, ni de la présence de balises ($F(1,85) = 0.22, p = .63, \eta_p^2 = 0.003$), ni de la présence de flashing ($F(1,85) = 1.64, p = .20, \eta_p^2 = 0.02$).

Effet des conditions de signaling et du feedback.

Dans le but de tester les conditions de signalisation et du feedback nous avons réalisé une ANOVA à mesures répétées à trois facteurs, deux facteurs intergroupes, respectivement le facteur flashing et le facteur balise et un facteur intragroupe, avant et après feedback. Les variables dépendantes étaient les scores de choix 1 et choix 2 avant et après feedback. Les scores obtenus au choix 1 avant feedback et au choix 2, après feedback sont présentés dans le Tableau 6.4.

Tableau 6.4.

Scores obtenus (moyennes et écarts types) pour le premier choix et le deuxième choix (après feedback) selon les groupes expérimentaux.

	Aucune signalisation (F- B-) n = 23	Balises seules (F- B+) n = 22	Flashing + balises (F+ B+) n = 23	Flashing seul (F+ B-) n = 21
Scores /1 pour le premier choix.	0.24 (0.33)	0.82 (0.29)	0.70 (0.33)	0.43 (0.46)
Scores /1 après le feedback.	0.57 (0.38)	0.77 (0.30)	0.85 (0.27)	0.83 (0.33)

Nous avons analysé ces résultats avec une ANOVA à mesures répétées à trois facteurs, deux facteurs intergroupes, respectivement le facteur flashing et le facteur balise et un facteur intragroupe, avant et après feedback. Les variables dépendantes étaient les scores de choix 1 et choix 2 avant et après feedback (Figure 6.12).

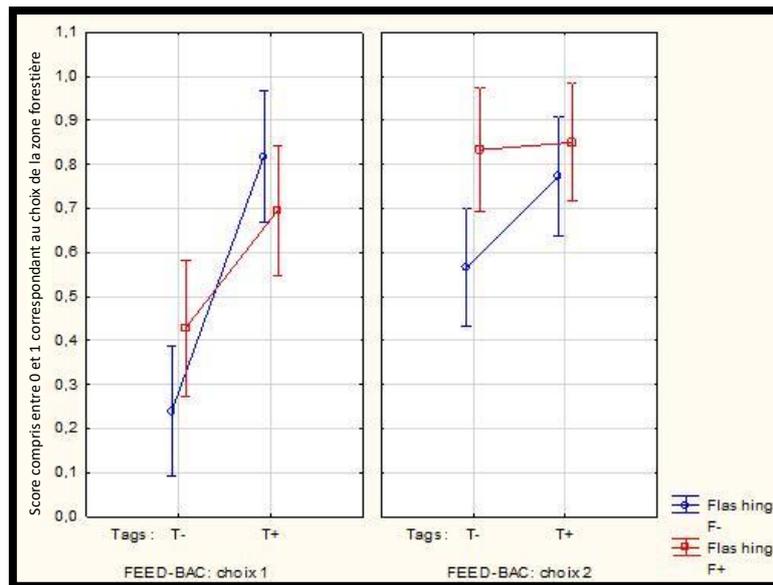


Figure 6.12. Effet des conditions de signalisation sur le choix réalisé avant et après le feedback.

Cette ANOVA a montré un effet principal de la présence de balises sur les scores de choix corrects $F(1,85) = 17.82, p < .0001, \eta_p^2 = 0.17$, mais peu d'effet général de la présence de flashing $F(1,85) = 2.64, p = .10, \eta_p^2 = 0.03$. L'effet principal du feedback sur les scores de choix était très significatif, $F(1, 85) = 38.16, p < .00001, \eta_p^2 = 0.31$. L'interaction entre flashing et balises était tout juste significative $F(1,85) = 3.94, p = .05, \eta_p^2 = 0.04$. Elle montre qu'en l'absence de balises, le flashing améliore les performances de décision, mais qu'en présence des balises, le flashing ne modifie pas les performances de décision.

De plus, les interactions entre feedback et balises, d'une part $F(1,85) = 20.93, p < .0001, \eta_p^2 = 0.19$; et entre feedback et flashing d'autres part, $F(1,85) = 4.20, p = .043, \eta_p^2 = 0.05$ étaient significatives.

L'interaction entre feedback et présence de balises révèle que l'augmentation des performances de décision après le feedback était beaucoup plus forte lorsqu'il n'y avait pas de balise ($F(1,85) = 57.13, p < .0001, d$ de Cohen = 0.92) que lorsque ces balises étaient présentes ($F(1,85) = 1.30, p = .26, d$ de Cohen = 0.17) dans l'environnement forestier lors de la phase d'exploration des 3 zones, Figure 6.13.

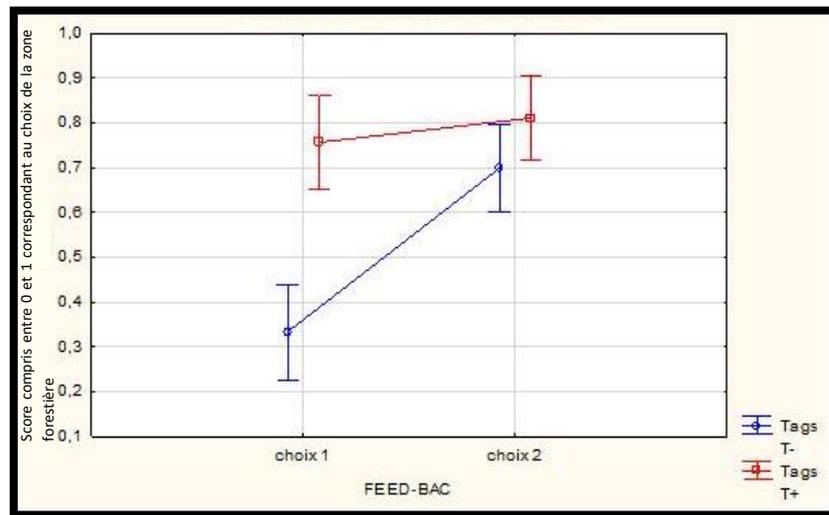


Figure 6.13. Interaction entre le choix réalisé avant et après le feedback et la présence de balises.

L'interaction entre feedback et flashing indique que la présence de flashing pendant l'analyse des zones forestières augmente plus fortement l'effet positif du feedback ($F(1,85) = 33.43, p < .00001, d$ de Cohen = 0.75) que l'absence de flashing ($F(1,85) = 8.63, p < .005, d$ de Cohen = 0.38), Figure 6.14.

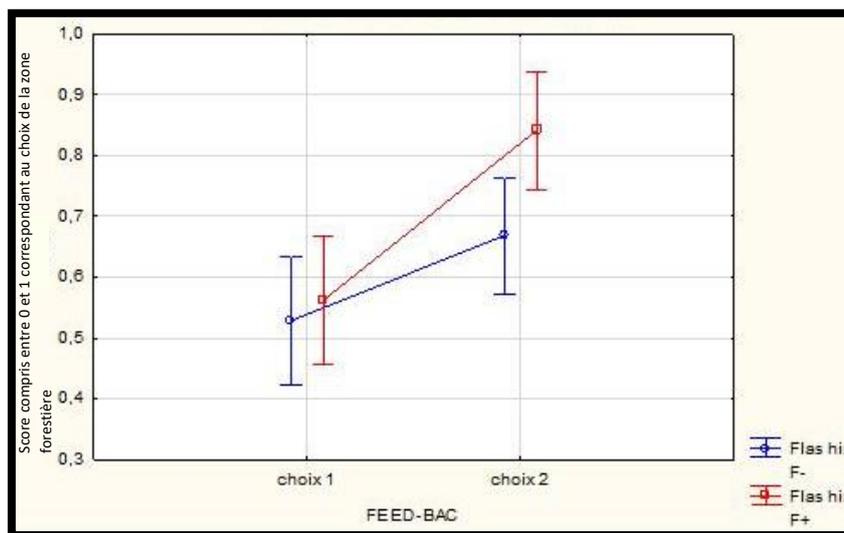


Figure 6.14. Interaction entre le choix réalisé avant et après le feedback et la présence de flashing.

Pour le choix 1, avant le feedback, il n'y avait pas de différence entre la présence et l'absence de balises ($F(1,85) = 0.19, p = .66, d$ de Cohen = 1.20). En revanche pour le choix 2 après le feedback, les scores de décision étaient significativement plus élevés dans la condition balises présentes que dans la condition balises absentes ($F(1,85) = 6.39, p < .01, d$ de Cohen = 0.35),

voir Figure 6.14. Ce résultat révèle que la présence de flashing, qui a eu un effet sur la quantité des éléments sélectionnés pendant l'analyse des zones forestières (voir Tableau 6.3), pourrait, non seulement diriger efficacement l'attention de l'apprenant vers des informations moins nombreuses mais pertinentes, mais cette technique de signalisation peut également faciliter, au moment du feedback et du choix 2, l'activation et la récupération en mémoire de concepts pertinents pour la tâche. Ainsi, on pourrait supposer qu'au moment du traitement des informations délivrées par le feedback, le flashing a facilité la mémorisation d'une sélection d'éléments forestiers pertinents qui peuvent ainsi être plus aisément mis en relation avec les informations fournies par le feedback. Les mécanismes d'inférence et d'élaboration conduisant à une meilleure décision après le feedback semblent être facilités par la récupération en mémoire d'informations pertinentes sélectionnées préalablement.

Deux ANOVA, non paramétriques par rang, de Kruskal-Wallis ont été conduites pour compléter les analyses précédentes. Il s'agissait d'ANOVA pour groupes indépendants multiples correspondants aux 4 groupes expérimentaux (F-B-, F+B-, F-B+, F+B+) et effectuées sur les variables dépendantes choix 1 et choix 2 respectivement.

La première analyse portait sur la variable dépendante choix 1. Elle montrait des différences significatives entre les groupes, $H(3, N = 89) = 26.035, p < .00001$; rangs moyens des groupes: F-B- = 27.26 ; F+B- = 38,47 ; F-B+ = 60.90 ; F+B+ = 53.48. Les comparaisons multiples entre les groupes sont présentées dans le Tableau 6.5 ci-dessous.

Tableau 6.5.

Comparaisons multiples des scores obtenus avant le feedback (choix 1) entre les 4 groupes expérimentaux.

Comparaisons multiples Valeurs p (bilatérales) ; choix 1 Variable indépendante (de groupement) : modalité Test de Kruskal-Wallis : $H(3, N = 89) = 26.03468 p = .0000$				
Choix	F- B-	F- B+	F+ B+	F+ B-
1	R:27,261	R:60,909	R:53,478	R:38,476
F- B-		0,000076	0,003475	0,902185
F- B+	0,000076		1,000000	0,026557
F+ B+	0,003475	1,000000		0,326235
F+ B-	0,902185	0,026557	0,3262350	

La seconde analyse portait sur la variable dépendante choix 2 et a également révélé des différences significatives entre les groupes, $H(3, N = 89) = 10.74, p < .02$; rangs moyens des groupes: F-B- = 32.50 ; F+B- = 51.26 ; F-B+ = 45.31 ; F+B+ = 51.47. Les comparaisons multiples entre les groupes sont présentées dans le Tableau 6.6 ci-dessous.

Tableau 6.6.

Comparaisons multiples des scores obtenus après le feedback (choix 2) entre les groupes expérimentaux.

Comparaisons multiples valeurs p (bilatérales) ; choix 2 Variable indépendante (de groupement) : modalité Test de Kruskal-Wallis : $H(3, N= 89) = 10,74055 p = ,0132$				
choix	F- B-	F- B+	F+ B+	F+ B-
2	R:32,500	R:45,318	R:51,478	R:51,262
F- B-		0,577058	0,076423	0,096765
F- B+	0,577058		1,000000	1,000000
F+ B+	0,076423	1,000000		1,000000
F+ B-	0,096765	1,000000	1,000000	

Les résultats de ces analyses non-paramétriques corroborent et confirment globalement les ANOVAs précédentes, présentées précédemment.

Analyse des arguments produits par l'élève.

Pour commencer, nous allons réaliser des calculs de corrélation R de Spearman, entre les arguments de justification (nombre d'arguments, pourcentage d'arguments vrais, pourcentage d'arguments faux et absents, type d'argument biologique vs. humain) et les scores respectivement au choix 1 et au choix 2. Ensuite, nous allons analyser le nombre d'arguments produits par les élèves pendant l'exploration selon les modalités expérimentales. Nous analyserons ensuite la nature des arguments produits pendant l'exploration en fonction des modalités d'exploration de l'environnement en RVI. Les scores du nombre d'arguments total, vrais, faux, partiels, absents en référence à la biodiversité et en référence à l'humain sont présentés dans le tableau 6. 7.

Tableau 6.7.

Nombre total d'arguments et pourcentage des arguments vrais, faux, partiels, absents, en référence à la biodiversité et en référence à l'humain produits par les élèves pendant l'exploration de l'environnement virtuel.

	Aucune signalisation (F- B-) N = 23	Balises seules (F- B+) N = 22	Flashing + balises (F+ B+) N = 23	Flashing seul (F+ B-) N = 21
Total arguments	19.35 (7.40)	17.64 (9.02)	22.17 (8.96)	19.71 (9.17)
% Arguments vrais	16.45 (20.91)	18.78 (14.44)	20.46 (17.55)	19.84 (20.41)
% Arguments faux	6.76 (9.98)	8.87 (8.61)	9.05 (7.46)	3.34 (5.36)
% Arguments partiels	50.14 (18.96)	49.09 (14.93)	45.07 (16.27)	46.15 (18.59)
% absence d'argument	27.37 (22.00)	21.21 (15.87)	24.81 (16.55)	29.58 (19.26)
% Arguments biodiversité	22.87 (20.44)	24.31 (18.19)	34.96 (22.97)	33.25 (26.21)
% Arguments humain	5.30 (9.78)	13.14 (16.24)	10.03 (13.95)	4.16 (8.11)

Nous avons tout d'abord effectué des calculs de corrélation, R de Spearman, entre les arguments de justification (nombre d'arguments, pourcentage d'arguments vrais, pourcentage d'arguments faux et absents, type d'argument –biologique vs. Humain-) et les scores respectivement au choix 1 et au choix 2. Ces calculs ont mis en évidence des corrélations significatives (i) positives entre le score au choix 1 et le pourcentage d'arguments vrais, $R = .38$, $p < .05$; le score au choix 2 et le pourcentage d'arguments biologiques, $R = .24$, $p < .05$ et négatives entre le choix 1 et le pourcentage d'arguments absents, $R = -.30$, $p < .05$; et entre le choix 2 et le pourcentage d'arguments absents, $R = -.26$, $p < .05$. Ce pattern de corrélation atteste de la cohérence des données expérimentales concernant la relation entre les scores de décision et les justifications fournies.

Ensuite, de façon à analyser les effets éventuels des facteurs expérimentaux sur la qualité des justifications pendant l'exploration de l'environnement, nous avons réalisé des ANOVA factorielles avec deux facteurs intergroupes, le flashing d'une part et les balises d'autre part ; la variable dépendante était le type d'arguments fournis par les élèves.

Les ANOVA factorielles ayant pour facteurs catégoriels le type de signalisation « flashing » et « balises » n'ont montré aucune différence significative pour les variables dépendantes nombre d'arguments produits : $F(1, 85) = 1.78, p = .18, \eta_p^2 = 0.02$, cependant le nombre d'arguments est marginalement plus élevé en présence de flashing et de balises qu'en présence de balises sans flashing ($F(1, 85) = 3.09, p = .08$). En ce qui concerne le pourcentage d'arguments vrais, aucun effet significatif des facteurs balises et flashing n'a pu être mise en évidence, respectivement, $F(1, 85) = .14, p = .70, \eta_p^2 = 0.001$, et $F(1, 85) = .41, p = .52, \eta_p^2 = 0.004$. Pour le pourcentage d'arguments partiels, aucun des deux facteurs, balise et flashing ne montre d'effet significatif, respectivement, $F(1, 85) = .08, p = .77, \eta_p^2 = 0.0009$ et $F(1, 85) = 1.19, p = .27, \eta_p^2 = 0.001$. Pour le pourcentage d'arguments absents les résultats sont également non significatifs, respectivement pour le facteur balise, $F(1, 85) = 1.91, p = .17, \eta_p^2 = 0.02$ et pour le facteur flashing, $F(1, 85) = .54, p = .46, \eta_p^2 = 0.006$. Aucune autre comparaison post-hoc entre les conditions n'était significative.

L'ANOVA factorielle ayant pour facteurs catégoriels le type de signalisation « flashing » et « balises » et pour variable dépendante le pourcentage d'arguments faux a montré un effet significatif de la présence de balises, $F(1, 85) = 6.53, p = .012, \eta_p^2 = 0.07$, mais pas d'effet de la présence de flashing, : $F(1, 85) = 1.38, p = .24, \eta_p^2 = 0.02$.

L'ANOVA factorielle ayant pour facteurs intragroupes le type de signalisation « flashing » et « balises » et pour variable dépendante le pourcentage d'arguments en faveur de la préservation de la biodiversité a montré un effet significatif de la présence du flashing, $F(1, 85) = 5.02, p = <.03, \eta_p^2 = 0.05$. Ainsi, la présence de flashing a entraîné une augmentation du nombre des justifications en faveur de la préservation de la biodiversité. En revanche, il n'y avait pas d'effet de la présence de balises sur le pourcentage de réponses en faveur de la préservation de la biodiversité, $F(1, 85) = .11, p = .73, \eta_p^2 = 0.001$. L'interaction entre les deux facteurs, balise et flashing n'était pas significative.

Corolairement, L'ANOVA factorielle ayant pour facteurs catégoriels le type de signalisation « flashing » et « balises » et pour variable dépendante le pourcentage d'arguments en faveur du bien être humain dans l'environnement forestier a révélé un effet de la présence de balises sur

les justifications relatives au bien être humain, $F(1, 85) = 6.69, p = .01, \eta_p^2 = 0.07$. En revanche, il n'y avait pas d'effet de la présence de flashing, $F(1, 85) = .64, p = .42, \eta_p^2 = 0.007$, ni d'interaction entre les deux facteurs, balise et flashing.

Enfin, nous avons souhaité analyser les relations éventuelles entre les connaissances préalables, d'une part, les moyennes générales des élèves d'autre part, ainsi que les performances de décision aux choix 1 et 2.

D'abord, des analyses de régression simples ont été conduites. Les deux premières ont testé l'effet de la mesure des connaissances préalables sur les scores en choix 1 et en choix 2. Pour le choix 1 ainsi que pour le choix 2, il n'y avait pas d'effet des connaissances préalables sur les scores de décision, respectivement pour le choix 1, $R^2 = .003$, ajusté, $-.008, F(1, 87) = .31, p = .57$; et pour le choix 2, $R^2 = .001$, ajusté, $-.0009, F(1, 87) = .17, p = .68$. Ensuite deux analyses de régression ont testé l'effet de la moyenne générale obtenue au cours de l'année par les élèves, incluant toutes les disciplines de la spécialité. Pour le choix 1, comme pour le choix 2, le niveau moyenne générale, ne prédisait pas significativement les performances de décision, respectivement pour le choix 1, $R = .17, R^2 = .030$, Ajusté, $R = 0.19, F(1, 87) = 2.74, p = .10$; et pour le choix 2, $R = .15, R^2 = 0.022$, Ajusté $R = 0.011, F(1, 87) = 1.99, p = .16$.

6.4 Discussion et conclusion.

L'objectif de cette étude était de tester deux types de guidage par signalisation visuelle, ainsi que le rôle du feedback dans le cadre d'une tâche d'exploration et de décision en réalité virtuelle immersive. Les processus cognitifs impliqués dans les apprentissages en RVI sont encore peu étudiés, et les résultats obtenus sont mitigés. Si certaines études semblent montrer un bénéfice des apprentissages en RVI, la méthodologie qu'elles utilisent n'est pas expérimentale et ne permet donc pas de dégager de conclusions fiables quant aux résultats obtenus (par exemple : Maresky & al. 2019 ; Liu, Wang & al. 2020 ; Saputro & Setyawan ; 2020). D'autres indiquent que les performances d'apprentissage sont identiques à celles observées lors d'apprentissages à partir de dispositifs conventionnels (Pulijala, 2017 ; Araiza-Alba, Keanel & al., 2021). Pour finir, des études plus récentes respectant une méthodologie expérimentale rigoureuse soulignent quant à elles un effet supérieur des supports traditionnels comme les vidéos 2D ou les présentations PowerPoint sur l'apprentissage comparé à un dispositif de RVI (Makransky, Terkildsen & Mayer, 2019 ; Parong & Mayer, 2020). Ces résultats s'expliqueraient par le fait que la charge cognitive sollicitée pendant l'apprentissage serait plus importante que lors d'un apprentissage conventionnel. Il est également possible que

la motivation puisse interférer avec l'acquisition de connaissances, au sens où l'apprenant serait centré sur l'exploration elle-même plus que sur la tâche d'apprentissage (Makransky & Petersen, 2021). Ces résultats nous ont mené à penser que le guidage par signalisation dans le cadre d'un apprentissage en RVI pourrait permettre à l'apprenant de rester centré sur la tâche à réaliser. Cependant, peu d'études ont étudié la signalisation dans le cadre des apprentissages en RVI, et celles-ci proposent une signalisation sous forme d'ajout d'annotations verbales (Blome, Diefenbach, Bucher & Von Mammen, 2017 ; Albus, Vogt & Seufert, 2021) mais aucune ne porte sur le rôle de la signalisation visuelle.

Dans le cadre d'une tâche d'exploration et de diagnostic au sein d'un environnement forestier complexe en RVI, nous avons proposé deux types de guidage : un guidage par signalisation visuelle impliquant le clignotement successif des éléments clés (flashing), et un guidage sémantique par l'ajout de points d'informations complémentaires (balises). Dans le but de palier à une éventuelle charge visuelle trop élevée, submergeant le canal visuel de l'apprenant en raison du flux continu d'informations pendant l'exploration du dispositif (cf. « *overwhelming* », Lowe, 2004), nous avons choisi de proposer un feedback informatif. Le feedback n'ayant pas été étudié dans le cadre d'apprentissages en RVI, nous nous sommes appuyés sur les travaux de Mayer & Moreno (2005). En effet, leurs résultats ont montré un bénéfice du feedback lors d'apprentissages multimédia, favorisant la production d'explications plus correctes de la part des apprenants. Nous nous attendions donc à observer le même profil de résultats en RVI.

Signalisation et performances décisionnelles des apprenants.

Concernant les deux types de signalisation que nous avons proposés, nos résultats ont montré que le flashing permettait d'orienter l'attention des élèves vers les éléments pertinents au cours de l'exploration de l'environnement forestier. En effet, les élèves consultaient moins d'éléments que les apprenants bénéficiant de balises. Nous pouvons expliquer ceci par le fait que, le flashing permettait d'extraire les informations essentielles à la compréhension de la parcelle explorée et pouvaient donc réduire le besoin de « tout » consulter dans le but de trouver le lien existant entre les différents éléments. La balise quant à elle donnait des informations complémentaires sur des éléments qui n'étaient pas toujours visibles directement, cela pouvait donc encourager les apprenants à chercher des éléments concrets permettant de faire le lien entre l'environnement exploré et les informations complémentaires fournies. On pourrait alors parler d'intérêt situationnel, accroissant le désir d'en apprendre davantage à propos de l'environnement proposé (Makransky & Petersen, 2021). Ces deux types de signalisation n'ont cependant pas montré d'effet sur le nombre d'éléments marqués par les apprenants. On peut

penser que les éléments marqués et donc, jugés importants ne dépendaient pas du type d'exploration proposé mais plutôt de la capacité de mémorisation de l'apprenant. Compte tenu du nombre moyen d'éléments marqués, on peut se demander si les apprenants choisissaient de ne sélectionner que le nombre d'informations qu'ils étaient susceptibles de pouvoir mémoriser. En effet, le nombre moyen d'éléments marqués pour les trois zones est compris entre 5.66 et 6.85, ce qui pourrait correspondre à l'empan 7 plus ou moins 2 déterminé par Miller (1956). De ce fait, nous pouvons nous demander si dans le cadre de la tâche proposée, les apprenants étaient capables volontairement ou non, de préserver leurs ressources mnésiques dans le but de créer une sorte de schématisation caractéristique de chaque parcelle explorée pour ensuite proposer un choix cohérent. L'activité de marquage ferait alors office d'indicateur de diversité pour chaque parcelle explorée. Bien évidemment, cette interprétation est spéculative et mériterait d'être étudiée de manière plus précise et objective.

Nos résultats ont également montré un effet des deux types de signalisation en tâche de décision (**H1**), plus précisément lors du choix émis par les apprenants après avoir consulté le feedback. Si le feedback a permis aux apprenants de modifier leur choix au profit d'une réponse plus correcte, (Mayer & Moreno, 2005), confirmant notre hypothèse **H2**, nous avons pu observer que le flashing participait à améliorer ces performances de décision dans la condition où l'environnement était dépourvu de balises (**H2a**). Cependant, en présence de balises, le flashing ne modifiait pas les performances de décision. Les résultats que nous avons obtenus montrent également que les performances de décision après le feedback étaient beaucoup plus importantes en l'absence de balises, même si ces balises favorisaient le choix de la bonne réponse après le feedback. Nous avons également pu noter que la signalisation par flashing permettait d'améliorer les scores de décision des élèves après le feedback. (**H2b**). Le Flashing, en plus de diriger l'attention de l'élève sur les éléments pertinents pendant l'exploration des zones forestières, permettrait de faciliter la récupération des concepts pertinents en mémoire lors du choix 2. Le Flashing favoriserait l'activation de mécanismes d'inférence et d'élaboration conduisant à une prise de décision plus efficace après avoir consulté le feedback.

La pertinence de l'argumentation des élèves.

Nos résultats ont montré que, quel que soit le type de signalisation, le nombre d'arguments produit par les élèves pendant l'exploration était identique (**H3**). Cependant, nous avons observé un effet marginal du flashing sur le nombre d'arguments produits. Aussi, nous avons observé un nombre plus important d'arguments faux produits par les apprenants lorsque l'environnement exploré comportait des balises. Nous pouvons expliquer cela par le fait que les

apprenants étaient novices et pouvaient parfois interpréter les informations fournies de manière erronée. Par exemple, la présence de bactéries dans le sol forestier pouvait être interprétée comme quelque chose de néfaste pour l'être humain alors qu'en réalité, il s'agissait d'un élément constitutif indispensable à la richesse et au renouvellement du sol. Les résultats obtenus ont également montré que le flashing avait entraîné une augmentation du nombre des justifications en faveur de la préservation de la biodiversité. Concernant les balises, elles semblent favoriser la production d'arguments en faveur du bien être humain au sein de l'environnement forestier. Ces résultats peuvent s'expliquer par le fait que le flashing permettait d'étudier une biodiversité visible, comme par exemple les champignons polypores ou encore les cavités de pieds modélisés sur certains arbres. A l'inverse, les balises fournissaient des informations sur des éléments qui n'étaient pas modélisés et qui, pour des apprenants novices pouvaient donner lieu à des interprétations erronées. On peut par exemple prendre l'exemple de la balise indiquant la présence d'animaux de type grenouilles, libellules et couleuvre. Dans ce cas, les apprenants pouvaient interpréter la présence de couleuvres comme un danger pour l'humain plutôt qu'une espèce protégée. Bien évidemment, pour confirmer ces interprétations il serait intéressant de réaliser une analyse plus fine des éléments qualitatifs dont nous disposons.

Conclusion et travaux futurs.

Les résultats que nous avons obtenus ont montré que les deux types de guidage que nous avons proposés accroissaient les performances en tâche de choix suite à l'exploration de l'environnement (**H1**) et que le feedback permettait d'autant plus d'accroître les performances à cette tâche (**H2**). Les résultats nous montrent également que si les balises favorisaient le choix de réponses correcte, l'interprétation par les élèves des informations qu'elles fournissaient pouvaient être erronées et conduire à la production d'arguments incorrects. Il serait donc intéressant de comparer les performances d'apprenants novices avec celles d'apprenants plus expérimentés dans le but de mieux comprendre le rôle du guidage sémantique par l'ajout d'informations complémentaires dans le cas où certains éléments ne puissent pas être modélisés dans l'environnement proposé. Concernant le Flashing, il serait intéressant de réaliser des mesures physiologiques pendant l'exploration comme par exemple le mouvement des yeux, dans le but de mieux comprendre les stratégies employées par les apprenants pour extraire et relier les informations pertinentes de l'environnement en fonction du type de signalisation proposé.

CHAPITRE 7 : DISCUSSION GENERALE

Dans ce chapitre, nous ferons un bref rappel des objectifs de nos travaux puis nous ferons la synthèse des principaux résultats que nous avons obtenus au cours de nos recherches, desquels nous exposerons les limites. Enfin, nous proposerons quelques perspectives d'application de nos recherches.

7.1 Objectif de nos travaux.

L'objectif de cette thèse était d'étudier et d'évaluer les processus cognitifs dans le cadre d'apprentissages en réalité virtuelle au sein du projet E-fran, **SILVA NUMERICA**. Ce projet étant lié principalement à l'enseignement, nous avons associé la recherche en psychologie cognitive à une démarche d'ingénierie pédagogique. La démarche de conception de notre matériel a consisté à répondre à la fois aux besoins des enseignants, et à nos besoins en termes de recherche expérimentale. L'enjeu était donc double : étudier les processus cognitifs liés à l'apprentissage dans le but de fournir des recommandations permettant d'optimiser le simulateur, et répondre aux besoins pédagogiques des enseignants. Nous avons dans un premier temps rencontré les équipes pédagogiques dans le but de recueillir leurs besoins, pour ensuite tenter d'y répondre tout en respectant les exigences méthodologiques liées à la démarche expérimentale que nous souhaitions appliquer.

Concernant les équipes de collège, les enseignants ont transmis la difficulté de la part des élèves à comprendre la chaîne alimentaire responsable de la décomposition du sol. C'est pourquoi nous avons sélectionné ce thème lors des expériences 1 et 2. Si les modalités expérimentales que nous avons établies n'étaient pas destinées à un usage pédagogique (apprentissage libre ou latence temporelle entre la présentation de la narration et de l'image), le matériel initial lui (présentation cohérente et pertinente) pourrait tout à fait être réutilisé à des fins pédagogiques et être intégré aux enseignements de SVT de collège.

Concernant les équipes de lycée, leurs besoins relevaient du fait que les élèves avaient parfois des difficultés à extraire et à relier les éléments pertinents au sein de l'environnement forestier de manière à en identifier les caractéristiques principales. C'est pourquoi nous avons choisi, à partir du simulateur existant, de développer un environnement en réalité virtuelle immersive. Le but était de comprendre d'une part comment les élèves exploraient la forêt avec ce type d'outil, et d'autre part d'analyser comment nous pouvions leur permettre d'extraire les informations pertinentes de l'environnement forestier. Le matériel a été conçu de manière à ce que les apprenants puissent créer un modèle mental cohérent relatif aux configurations

spécifiques des principales stations⁸ forestières qu'ils pourraient rencontrer dans l'exercice de leurs fonctions professionnelles. Le matériel que nous avons conçu nous a permis de mieux comprendre les mécanismes cognitifs liés à l'apprentissage en réalité virtuelle immersive, et il pourrait également servir de support pédagogique aux enseignants. En effet, nous avons déterminé trois zones forestières représentant des stations spécifiques :

- Un sol humide : sol sur lequel l'implantation de structures telles qu'un observatoire ou une aire de pique-nique n'est pas possible (ou du moins pas recommandée). Station qui abrite une grande biodiversité, dont des espèces protégées comme les couleuvres ou les grenouilles que le forestier se doit de préserver.
- Un sol stable abritant une grande diversité d'essences⁹, de végétation et de micro habitats. Cette zone rendait l'implantation de structures de type observatoire ou aire de pique-nique complexe du fait de la variété des éléments qui la composait, le but étant de préserver cette grande diversité.
- Un sol stable dans lequel peu d'espèces différentes étaient représentées. Dans cette zone il convenait de préserver la biodiversité, mais l'implantation d'une structure de type observatoire ou aire de pique-nique était plus appropriée. En effet, la présence d'une diversité d'essences et de végétation moindre est moins complexe à aménager.

A l'avenir, cet environnement virtuel pourrait être utilisé par les enseignants de manière à accompagner les élèves dans l'identification des éléments spécifiques à une station donnée, par exemple : quelles sont les spécificités d'une zone humide ? pourquoi ne pas implanter de structures sur ce type de sol ? Dans chaque station, quelles sont les espèces protégées ? etc.

7.2 Les apprentissages en réalité virtuelle : les apports de nos travaux.

Comme nous avons pu le constater au cours de nos travaux, les études portant sur l'analyse des mécanismes cognitifs dans le cadre d'apprentissages en réalité virtuelle sont peu nombreuses. Aussi, les résultats obtenus sont mitigés : parfois la réalité virtuelle semble être une plus-value pour l'apprentissage (Maresky & al. 2019 ; Liu, Wang & al. 2020 ; Saputro & Setyawan, 2020), parfois elle n'offre aucun bénéfice (Pulijala, 2017 ; Araiza-Alba, Keanel &

⁸ La notion de station forestière désigne « une étendue de terrain homogène sur les plans du climat, du relief, du sol et de la végétation spontanée » (<https://www.cnpf.fr/n/les-stations-forestieres/n:1030>)

⁹ L'essence désigne une espèce d'arbres. <https://www.onf.fr/onf/raconte-moi-la-foret/la-langue-des-bois/+7f3::lessence-lespece-dun-arbre.html>

al., 2021). D'autres études montrent qu'au contraire, la réalité virtuelle entraverait l'apprentissage comparée aux supports traditionnels (Makransky, Terkildsen & Mayer, 2019 ; Parong & Mayer, 2020). C'est pourquoi nous avons choisi d'étudier les mécanismes cognitifs relatifs aux apprentissages en réalité virtuelle, dans une tentative d'étendre nos connaissances et de compléter les conclusions émises lors des recherches antérieures. Les trois études que nous avons réalisées nous ont permis de répondre à plusieurs questions nouvelles sur les mécanismes cognitifs en jeu dans les apprentissages en réalité virtuelle.

7.2.1 Le principe de cohérence sémantique en réalité virtuelle non immersive.

L'objectif de cette première étude était de tester le principe de guidage de la cohérence sémantique dans le cadre d'une leçon en réalité virtuelle non immersive. Il s'agissait de déterminer si un guidage sémantique fixe selon une chaîne causale délivrée automatiquement par le système, était plus efficace pour la compréhension d'un phénomène complexe en réalité virtuelle non immersive, qu'une présentation libre.

Les résultats que nous avons obtenus ont montré que les deux types de présentation, fixe et libre, de la leçon permettaient un gain de connaissances au questionnaire de post-test. En outre, les participants qui avaient bénéficié de la leçon fixe et cohérente ont obtenu des scores significativement supérieurs aux participants ayant étudié la leçon en mode libre. Dans le but d'optimiser les apprentissages en RV, il serait donc nécessaire de proposer un dispositif dans lequel les informations pertinentes seraient présentées de manière cohérente selon une hiérarchie fixe.

Cependant, malgré le gain de connaissances apparent, nous n'avons pas observé d'effet plafond après l'étude de la leçon, pourtant, les élèves possédaient des connaissances préalables sur le thème étudié. Nous pouvons donc nous demander si les performances des apprenants sont liées à la complexité du matériel proposé ou de celle du processus étudié. La leçon proposée était d'une durée de 12 minutes et présentait de nombreux éléments interconnectés. Nous avons pris soin de présenter simultanément de petits groupes d'éléments, cependant, si le rôle de l'un des groupes présentés n'était pas compris et intégré, cela pouvait impacter la compréhension du reste de la leçon. Il est possible que le défilement de l'ensemble des éléments présentés finisse, à un moment donné, par submerger le calepin-visuospatial (Baddeley & Hitch ; 1974) en mémoire de travail des apprenants (Lowe, 2004), empêchant ainsi la compréhension de la totalité de la leçon.

Deux possibilités d'expérimentation pourraient alors être envisagées. La première serait de ne présenter qu'une séquence, ou un petit groupe de séquences de la leçon à l'apprenant, et de laisser un temps d'intégration associé à un contrôle de connaissances avant de proposer la ou les séquences suivantes. Ainsi, il serait possible de déterminer une durée d'exposition optimale aux dispositifs de réalité virtuelle pour les enfants. La seconde possibilité serait de donner comme consigne à l'apprenant d'appuyer sur une touche de clavier lorsqu'il ne comprend pas ou plus les informations présentées. De cette manière, nous pourrions tenter d'identifier le moment où l'intégration de la chaîne causale du processus présenté est rompue. Cela nous permettrait éventuellement de déterminer à quel moment le calepin visuo-spatial de l'apprenant serait surchargé, et d'en évaluer l'impact sur la compréhension de la totalité du processus. De plus, enregistrer des mesures de type EEG pendant l'apprentissage en RV (Makransky, Terkildsen & Mayer ; 2019) permettrait d'expliquer plus objectivement les processus cognitifs en jeu pendant le traitement des informations présentées.

7.2.2 La contiguïté temporelle en réalité virtuelle.

L'objectif de cette étude était de tester et d'approfondir nos connaissances à propos du principe de contiguïté temporelle dans le cadre de l'apprentissage multimédia (Mayer & Anderson, 1991, 1992 ; Mayer & Sims, 1994 ; Mayer, Moreno, Boire & Vagge, 1999 ; Michas & Berry, 2000 ; Mayer, 2009, 2014, Fiorella & Mayer, 2017) au moyen d'un dispositif de réalité virtuelle non immersive.

Les résultats que nous avons obtenus ont montré qu'une latence entre le son et l'image correspondante, même très courte (deux secondes), semblait préjudiciable pour les processus d'organisation et d'intégration des informations en mémoire de travail. De plus, nos résultats montrent une asymétrie des performances en post-test : les performances des apprenants étaient légèrement meilleures lorsque les images étaient diffusées avant la narration plutôt que lorsqu'elles étaient présentées après. Cela pourrait s'expliquer par le fait que, lorsque les apprenants entendent les explications verbales, ils explorent les images de manière à trouver l'élément correspondant aux informations qu'ils viennent d'entendre. Cette tentative de trouver une correspondance entre la narration et l'image visualisée pourrait augmenter la charge extrinsèque (Sweller, 2010) et nuire aux processus d'intégration. Cependant, quel que soit le type de latence, les performances d'apprentissages étaient affectées.

Nous avons également pu observer un effet de la mémoire de travail et des habiletés spatiales sur les performances d'apprentissage. Les apprenants dont la mémoire de travail verbale était

élevée et les apprenants dont les habiletés spatiales étaient fortes ont obtenus de meilleures performances d'apprentissage. Cependant, cela n'a pas permis de compenser les effets néfastes de la latence temporelle entre la narration et l'image, d'autant plus lorsque la latence est élevée (-6 et +6 secondes). Pour finir, les résultats obtenus au post-test d'intégration texte-image ont révélé que les apprenants étaient conscients du décalage présent entre le son et l'image pendant la leçon.

La contiguïté temporelle ayant été peu étudiée, il serait intéressant que de nouvelles études soient réalisées selon une méthodologie similaire à la nôtre (comme dans l'étude de Baggett, 1984) de manière à vérifier et peut-être consolider les résultats que nous avons obtenus. En outre, nos observations mériteraient d'être approfondies, notamment par l'analyse du mouvement oculaire et/ ou par des mesures EEG (Makrasky, Terkildsen & Mayer ; 2019) pendant l'exploration de l'environnement virtuel afin de comprendre avec plus de précision les mécanismes cognitifs en jeu dans la compréhension, l'intégration et l'apprentissage lors de la présence d'incongruités entre les informations verbales et imagées en réalité virtuelle. Poursuivre ces travaux serait d'autant plus important que les processus cognitifs en jeu dans ce type de tâche pourraient expliquer les difficultés d'apprentissage observées en réalité virtuelle immersive (Makrasky & al., 2019, 2020 ; Araiza-Alba, Keane, Matthews, Simpson, Strugnell, Chen, & Kaufman, 2021). En effet, en réalité virtuelle immersive, l'apprenant contrôle lui-même le flux d'images visualisé, par le mouvement de la tête et des yeux. Dans ce cas de figure, et compte tenu de la richesse des environnements virtuels, il serait tout à fait probable que l'apprenant consulte des éléments non pertinents avec la tâche d'apprentissage engagée (Makrasky, 2019), ce qui en toute logique, entraverait l'apprentissage.

7.2.3 Rôle de la signalisation et du feedback en RVI.

L'objectif de cette étude était de tester deux niveaux de guidage par signalisation ainsi que le rôle du feedback dans le cadre d'une tâche d'apprentissage et de décision en réalité virtuelle immersive.

L'apprentissage en RVI, par sa richesse visuelle, solliciterait une quantité de ressources cognitives plus importante que lors d'apprentissages multimédias non immersifs, ce qui empêcherait un apprentissage efficient (Makrasky, Terkildsen & Mayer, 2019 ; Parong & Mayer, 2020). C'est pourquoi nous avons choisi de proposer un guidage par signalisation visuelle, par le clignotement successif des éléments clés (flashing) ainsi qu'un guidage sémantique par l'ajout de points d'informations complémentaires (balises). Nous avons

également proposé un feedback informatif. Notre but était de déterminer si l'intégration d'un guidage sémantique et/ou d'un guidage par signalisation favorisait l'extraction d'éléments pertinents d'un environnement forestier riche et complexe en RVI et si le feedback permettait de modifier une décision erronée au profit d'un raisonnement plus synthétique et abouti. Les résultats que nous avons obtenus ont révélé que la signalisation par flashing permettait d'orienter l'attention des apprenants vers les éléments pertinents au cours de leur exploration, et également d'améliorer les performances de choix après le feedback (Mayer & Moreno, 2005). Concernant le guidage sémantique, les balises comme le flashing favorisaient le choix d'une réponse correcte après le feedback. Cependant, les performances de décision après le feedback étaient beaucoup plus importantes en l'absence de balises. De plus, le nombre d'arguments faux produits au cours de l'exploration était plus important lorsque les apprenants avaient la possibilité de consulter des balises informatives. Ces résultats s'expliquent par le fait que le contenu informatif des balises, contrairement au flashing, ferait l'objet d'interprétations pas toujours correctes de la part des apprenants ; alors que le flashing permettrait l'activation de mécanismes d'inférence et d'élaboration conduisant à une prise de décision plus efficace après avoir consulté le feedback. Quant au feedback, les résultats que nous avons obtenus ont révélé que les performances de décision des élèves étaient significativement supérieures après le feedback (Mayer & Moreno ; 2005). Le feedback aurait alors permis aux apprenants de modifier leur raisonnement et de modifier leur décision dans le cas où ils s'étaient trompés lors du premier choix.

Les apprenants ayant participé à cette étude étaient novices et ne disposaient pas tous des mêmes connaissances préalables à propos de l'environnement forestier. En effet, certains se destinaient à l'exploitation forestière alors que d'autres se destinaient par exemple à l'exploitation agricole ou encore à l'assistantat vétérinaire. Il serait donc intéressant de répliquer notre étude auprès d'élèves ayant les mêmes objectifs professionnels et donc, les mêmes connaissances de l'environnement étudié. Le fait que les apprenants soient novices nous a permis de constater que le guidage sémantique pouvait amener à produire un raisonnement erroné. Par exemple « *ici il y a des bactéries, ça n'est pas très hygiénique pour l'homme* » alors qu'en réalité, les bactéries sont un gage de la qualité du sol. Il serait également intéressant de proposer une nouvelle exploration de l'environnement RVI après le feedback afin d'évaluer l'évolution en termes de stratégies d'exploration de la part des élèves. Cette nouvelle étape expérimentale nous permettrait également d'étudier plus précisément la plus-value du feedback sur les processus

cognitifs dédiés à la sélection d'informations pertinentes dans un environnement complexe en RVI.

Pour finir, les résultats que nous avons obtenus montrent que la RVI pourrait tout à fait représenter une plus-value pour les apprentissages à condition d'en faire un usage pertinent et raisonné. La réalité virtuelle doit pouvoir répondre à un besoin réel. Par exemple, dans le cadre de notre projet, il n'est pas possible de visualiser l'évolution de la forêt suite à la coupe d'un certain nombre d'arbres en situation réelle de terrain. Cela nécessite de se projeter dans le temps sur plusieurs années. Dans ce cas, il est pertinent d'utiliser la RVI qui permet aux apprenants de visualiser une évolution de la forêt fidèle à la réalité et répondant aux actions qui y sont menées. En revanche, proposer un dispositif d'apprentissage en RVI alors que l'immersion n'est pas nécessaire, peut entraîner un apprentissage moins efficient (Exemple : Brian, Forest, Epstein & Brown, 2018 ; Makransky, 2019).

7.2.4 Conclusion.

Les trois études que nous avons réalisées ont révélé des aspects positifs de l'apprentissage en réalité virtuelle, qu'elle soit ou non immersive. Le principe de cohérence sémantique permet de favoriser l'apprentissage en RV non immersive, même si le processus de décomposition de la matière organique reste complexe à intégrer pour les jeunes apprenants. Néanmoins, il est tout à fait possible d'utiliser la leçon hiérarchisée, cohérente et synchronisée que nous avons conçue dans le but de compléter un enseignement en classe ou de terrain. Concernant la contiguïté temporelle, nos résultats ont montré qu'un décalage même de courte durée entre le son et l'image pouvait perturber l'intégration et la manipulation des informations perçues en mémoire de travail, ce qui perturbe l'apprentissage. Concernant le matériel développé en RVI, il pourrait tout à fait être utilisé auprès d'apprenants novices dans le cadre de l'apprentissage des caractéristiques relatives aux différentes stations forestières, sous la supervision des enseignants qui pourraient corriger en temps réel les interprétations erronées des élèves.

7.3 Recommandations.

Les résultats des études que nous avons réalisées nous portent à dire que, dans le cadre d'apprentissages multimédias, il est indispensable que les concepts et les images présentés soient parfaitement synchronisés. Les résultats obtenus lors de nos deux premières études nous mènent à penser que la meilleure configuration possible est l'alignement temporel entre l'information visuelle et l'information verbale orale. C'est par exemple le cas de la version

synchronisée de la leçon que nous avons proposée sur la décomposition de la matière organique. Lorsqu'un élément doit être décrit, il est zoomé au milieu de l'écran et le reste tout au long de la diffusion de la narration correspondante. Ainsi, aucune confusion n'est possible : l'apprenant regarde ce qu'il est supposé regarder lorsque la narration est diffusée. En fait, comme décrit par le modèle APM (Lowe & Boucheix, 2008), l'information animée est dans un premier temps décomposée en unités individuelles. Cette décomposition semble être étroitement liée avec le décours temporel de l'information verbale correspondante (orale). De ce fait, si la narration n'est pas parfaitement synchronisée avec l'image, il est possible que l'apprenant ne construise pas correctement la relation entre les deux sources d'information ce qui perturberait l'apprentissage.

La cohérence sémantique est également cruciale pour favoriser un apprentissage efficace. Ainsi, nous recommandons une présentation hiérarchique, organisée et cohérente des informations.

Par contre, lorsque l'apprenant a des connaissances préalables comme c'était le cas lors de notre première expérimentation, on peut se demander si un engagement et une gestion de l'apprentissage en autonomie serait appropriée. Dans ce cas de figure on proposerait un dispositif d'apprentissage sémantiquement cohérent et interactif suscitant plus d'engagement de la part de l'apprenant. On peut imaginer des leçons dans lesquelles on aurait des systèmes de questions, d'interrogations. Il s'agirait de faire interagir l'élève avec le dispositif éducatif tout en lui proposant des feedbacks réguliers pour qu'il puisse corriger ses erreurs.

Concernant les environnements virtuels riches et dynamiques sollicitant fortement les ressources cognitives, il est essentiel de proposer un guidage sémantique par l'ajout d'informations clés pertinentes, ainsi qu'un guidage par signalisation visuelle. Cependant, il faut veiller à ne pas présenter une signalisation trop riche. Le but est d'orienter l'attention de l'apprenant novice au bon endroit de manière à ce qu'il puisse extraire et organiser les informations présentées dans l'environnement numérique de manière plus efficace, pour ensuite constituer un modèle mental cohérent du processus proposé.

Il est également indispensable de fournir un feedback explicatif à l'apprenant pour qu'il puisse comprendre et corriger ses erreurs. Dans un contexte où on proposerait à un apprenant d'agir sur l'environnement forestier et d'en observer les répercussions, nous pensons que le fait de simplement voir évoluer la forêt sur plusieurs années ne suffirait pas à comprendre les causes de cette évolution. L'évolution de l'environnement en lui-même montre des changements visuels ; il est cependant nécessaire d'en comprendre les causes et les conséquences. Afin de

corriger ses erreurs ou d'être en mesure de reproduire une réussite, il est important que l'apprenant bénéficie d'un feedback explicatif en complément du feedback visuel généré par son action sur l'environnement.

Pour finir, les résultats que nous avons obtenus en RVI nous ont montré que la présence de balises informatives présentant des informations complémentaires sur l'environnement, pouvait parfois donner lieu à la production d'arguments incorrects de la part des apprenants. Il est donc nécessaire d'être très vigilant quant à l'interprétation que les apprenants peuvent faire des informations données. Dans ce cas, il est nécessaire de superviser l'apprenant pendant son exploration de manière à ce qu'il n'apprenne pas d'erreurs. Un feedback correctif peut par exemple être réalisé par l'enseignant en temps réel.

REFERENCES

- Albus, P., Vogt, A. & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education* 166 (2021) 104154
- Allcoat, D. & von Mülhnenen, A. (2018). Learning in virtual reality : Effect on performance, émotion and engagement. *Research in Learning Technology*, Vol.26. <http://dx.doi.org/10.25304/rlt.v26.2140>
- Araiza-Alba, P., Keane, T., Matthews, B., Simpson, K., Strugnell, G., Chen, W.S., & Kaufman, J., (2021). The potential of 360-degree virtual reality videos to teach water-safety skills to children, *Computers & Education*, 163, 104096.
- Arndt, J., Schüler, A., & Scheiter, K. (2019). Investigating the Influence of Simultaneous–Versus Sequential–Text-Picture Presentation on Text-Picture Integration. *The Journal of Experimental Education*, 87:1, 116-127, DOI: 10.1080/00220973.2017.1363690
- Baceviciute, S., Mottelson, A., Terkildsen, T., & Makransky, G. (2020). Investigating Representation of Text and Audio in Educational VR using Learning Outcomes and EEG. *Conference on Human Factors in Computing Systems, CHI 2020 April 25–30, 2020, Honolulu, HI, USA.*
- Baddeley, A. D. (1992). Working memory. *Science*, 255, 556-559.
- Baddeley, A. D. (2000). *The episodic buffer: A new component of working memory?* *Trends in Cognitive Science*, 4, 417-423.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). San Diego, CA: Academic Press.

- Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology*, 76(3), 408-417. doi :[10.1037/0022-0663.76.3.408](https://doi.org/10.1037/0022-0663.76.3.408)
- Baggett, P., & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. *Educational Communication and Technology Journal*, 31, 23-32.
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, 133(1), 83–100. <https://doi.org/10.1037/0096-3445.133.1.83>
- Berney, S. & Bétrancourt, M. (2009). When and why does animation enhance learning? A meta-analysis. *In Proceedings of the 13th EARLI biennial conference on Research on Learning and Instruction*, Aug. 28 - Sept. 1, Amsterdam (NL).
- Berney, S. & Bétrancourt, M. (2016). Does animation enhance learning ? A meta-analysis. *Computers & Education*, 101, 150-167. DOI : [10.1016/j.compedu.2016.06.005](https://doi.org/10.1016/j.compedu.2016.06.005)
- Berney, S., Bétrancourt, M., Molinari, G. & Hoyek, N. (2015). How Spatial Abilities and Dynamic Visualizations Interplay When Learning Functional Anatomy With 3D Anatomical Models. *Anatomical Sciences Education*, 2015, vol. 8, no. 5, p. 452-462
- Berney, S., Groff, J. & Gagnière, L. (2005). L'animations 3D, un soutien à l'apprentissage d'informations spatiales ?. 7ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2015), Jun 2015, Agadir, Maroc. pp.65-70. [ffhal-01405933f](https://hal.archives-ouvertes.fr/hal-01405933f)
- Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 287e296). New York: Cambridge University Press.
- Bétrancourt, M., Bisseret, A. & Faure, A. (2001). Effect of the sequential display of pictures on the user's mental representation : an experimental study. In J.-F. Rouet, J. J. Levonen and

- A. Biardeau (eds.), *Multimedia learning: cognitive and instructional issues*. (pp 112-118),
- Boucheix, J.-M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, 19, 112–127.
- Boucheix, J.-M. & Forestier, C. (2017). Reducing the transience effect of animations does not (always) lead to better performance in children learning a complex hand procedure. *Computers in Human Behavior*, 69, 358-370.
- Boucheix, J.-M., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education*, 10(4), 369-388.
- Boucheix, J.-M. & Lowe, R.K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animation. *Learning and instruction*, 123-135.
- Boucheix, J.-M. & Lowe, R.K., Putri, D.K., Groff, J. (2013). Cueing animations: Dynamic signaling aids information extraction and comprehension. *Learning and Instruction*. 25, 71-94.
- Boucheix, J.-M. & Rouet, J.-F. (2011) « Les animations interactives multimédias sont-elles efficaces pour l'apprentissage ? », *Revue française de pédagogie* [En ligne], 160 | juillet-septembre 2007, mis en ligne le 01 septembre 2011, consulté le 23 mars 2021. URL : <http://journals.openedition.org/rfp/832> ; DOI : <https://doi.org/10.4000/rfp.832>
- Bransford, J.D., Brown, A.L., Cocking, R.R., 1999. How People Learn: Brain, Mind, Experience, and School. *National Academy Press*, Washington, DC.
- Bransford, J.D. & Johnson, M.K. (1972). Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of verbal learning and verbal behavior* 11, 717-726.

- Brewer-Deluce, D., Bak, A.B., Simms, A.J., Sinha, S., Mitchell, J.P., Shin, D., Saraco, J. & Wainman, B.C. (2021). Virtual reality bell-ringer : the development and testing of stereoscopic application for human gross anatomy. *Anatomical Sciences Education*. vol.14, I3, p.330-341.
- Brian, A., Forest, M., Epstein, M. & Brown, D.E. (2018). Comparing the performance of an immersive virtual reality and traditional desktop cultural game. In Proceedings of the 2nd International Conference on Computer-Human Interaction Research and Applications. CHIRA 2018), pages 54-61 DOI: 10.5220/0006922800540061
- Burkhardt, J.-M., « Réalité virtuelle et ergonomie : quelques apports réciproques », *Le travail humain* 2003/1 (Vol. 66), p. 65-91. DOI 10.3917/th.661.0065
- Carroll, J.B. (1993). Human cognitive abilities. A survey of factor-analytic studies. *Cambridge:Cambridge University Press*.
- Castro-Alonso, J.C., Ayres, P. & Paas, F. (2016). Comparing apples and oranges? A critical look at research on learning from statics versus animations. *Computers & Education*, 102, 234-243.
- Castro-Alonso, J.C., Wong, M., Adesope, O.O., Ayres, P. & Paas, F. (2019). Gender Imbalance in Instructional Dynamic Versus Static Visualizations: a Meta-analysis. *Educational Psychology Rreview*. <https://doi.org/10.1007/s10648-019-09469-1>
- Castro-Alonso, J. C. (2019). Overview of visuospatial processing for education in health and natural sciences. In J. C. Castro-Alonso (Ed.), *Visuospatial processing for education in health and natural sciences* (pp. 1–21). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-20969-8_1.
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8,293-33 2.
- Chandler, P. & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233-246.

- Chanquoy, L., Tricot, A. & Sweller, J. (2007). La charge cognitive. Théorie et applications (pp. 131-188). Paris, France : Armand Colin.
- Chen, C.J.& Toh, S.C. (2005). Are learning styles relevant to virtual reality ? *Journal of research on technology in education*.
- Cho, B.H., Lee, J.M., Ku, J.H., Jang, D.P., Kim, J.S., Kim, I.Y., Lee, J.H.,& Kim, S.I. (2002). Attention enhancement system using virtual reality and EEG biofeedback. Proceedings of the IEEE Virtual Reality 2002 (VR.02). DOI: 10.1109/VR.2002.996518 .
- Chung, W-L., & Hu, C-F. (2007). Morphological awareness and learning to read Chinese. *Reading and Writting*, 20,441-461.
- Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split attention effect : Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load ? *Computers in Human Behavior*, 25,315-324.
- Clark-Carter, D. (1997). *Doing Quantitative Psychological Research: From Design to Report*. Hove: Psychology Press.
- Clark, R.E. & Feldon,D.F. (2014). 6 Ten common but questionable principles of multimedia learning. *The Cambridge handbook of multimedia learning*, 151-152.
- Clarys, D. (2001). Psychologie de la mémoire humaine : de nouvelles avancées théoriques et méthodologiques. *L'année Psychologique*, 3-4, pp. 495-519.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Prog Brain Res*. 2008 ; 169: 323–338. Doi:10.1016/S0079-6123(07)00020-9

- Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Dehue, N. & van de Leemput, C.(2014) What does germane load mean ? An empirical contribution to the cognitive load theory.Frontiers in psychology. DOI : 10.3389/fpsyg.2014.01099.
- Desmette, D., Hupet, M., Schelstraete, M-A. & Van Der Linden, M. (1995). Adaptation en langue française du « Reading Span Test » de Daneman et Carpenter (1980). L'année psychologique. 1995 vol. 95, n°3. pp. 459-482; Doi : <https://doi.org/10.3406/psy.1995.28842>
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: seeing is understanding? *Learning and Instruction*, 20(2), 111e122.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011). Improved effectiveness of cueing by self-explanations when learning from a complex animation. *Applied Cognitive Psychology*, Appl. Cognit. Psychol. 25: 183–194 (2011)
- Désison, J.C., Bétrancourt, M. & de Vries, E. (2021). Cross Representational Signaling and Cohesion Support Inferential Comprehension of Text–Picture Documents. *Frontiers in Psychology*. Volume 11, Article 592509. Doi: 10.3389/fpsyg.2020.592509
- Doolittle, P.E., Altstaedter, L.L., (2009). The Effect of Working Memory Capacity on Multimedia Learning: Does Attentional Control Result in Improved Performance? *Journal of research in innovative teaching*. 2, pp.7-23.
- Eaves, D., Breslin, G., van Schalk, P., Robinson, E. & Spears, I.R. (2011). The Short-Term Effects of Real-Time Virtual Reality Feedback on Motor Learning in Dance. *Presence*, Vol. 20, No. 1, February 2011, 62–77. By the Massachusetts Institute of Technology.

- Fabris, C.P., Rathner, J.A., Fong, A.Y., & Sevigny, C.P. (2019). Virtual Reality in Higher Education. *International Journal of Innovation in Science and Mathematics Education*, 27(8), 69-80.
- Fiorella, L., Mayer, R.E., (2017). Spontaneous Spatial Strategy Use in Learning from Scientific Text. *Contemporary Educational Psychology*, doi: <http://dx.doi.org/10.1016/j.cedpsych.2017.01.002>
- Fischer, S.C., Hitkey, D.T., Pellegrino, J.W., Law, D.J. (1994). Strategic processing in dynamic spatial reasoning tasks. *Learning and Individual Differences*. Vol.6, No.1.165-105.
- Freina, L., Bottino R., Tavella, M. (2016), *From e-learning to VR-learning: an example of learning in an immersive virtual world*, *Journal of e-Learning and Knowledge Society*, v.12, n.2, 101-113. ISSN: 1826-6223, e-ISSN:1971-8829
- Freina, L. & Canessa, A. (2015). Immersive vs Desktop Virtual Reality in Game Based Learning. 9th European Conference On Games Based Learning (ECGBL), Steinkjer, Norway.
- Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives. *Paper presented at the Conference eLearning and Software for Education*. (1), 133-141
- Gibson, J. J. (1947). Army air forces aviation psychology program, report no. 7: Motion picture testing and research. Washington, DC: U.S. Government Printing Office.
- Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, 16, 511-525.
- Grassini, S., Laumann, K. & Skogstad, M.S. (2020). The use of virtual reality alone does not promote training performance (but sense of presence does). *Frontiers in Psychology*, volume 11, 1743. DOI : 10.3389/fpsyg.2020.01743

- Hamilton, D., McKechnie, J., Edgerton, E. & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education : a systematic literature review of quantitative learning outcomes and experimental design. *Journal of computer in education*. 8, 1-32. <https://doi.org/10.1007/s40692-020-00169-2>
- 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*. 1,92-102.
- Harp, S.F., Mayer, R.E., (1998). How Seductive Details Do Their Damage: A Theory of Cognitive Interest in Science Learning. *Journal of Educational Psychology*. 3, 414-434.
- Hegarty, M. (1992). Mental animation : Inferring motion from static displays of mechanical systems. *Journal of experimental psychology : Learning, Memory and Cognition*, **18 (5)**, 1084–1102.
- Hegarty, M. & Just, M.A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language*. 32,717-742.
- Hegarty, M., & Kozhevnikov, M. (1999). Spatial abilities, working memory and mechanical reasoning. In J.Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design* (pp. 221–241). Sydney, Australia:University of Sydney
- Hegarty, M., Kozhevnikov, M. & Waller, D. (2008). Perspective Taking/Spatial Orientation Test. For electronic copies of those publications and this test package, see: <http://www.psych.ucsb.edu/~hegarty/>
- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspectivetaking spatial abilities. *Intelligence*, 32, 175-191.
- Höffler, T. (2010). Spatial Ability: Its Influence on Learning with Visualizations—a Meta-Analytic Review. *Educational Psychology Review*, 22:245–269 DOI 10.1007/s10648-010-9126-7

- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: a meta-analysis. *Learning and Instruction*, 17, 722e738.
- Hunt, E., Pellegrino, J. W., Frick, R. W., Farr, S. A., & Alderton, D. L. (1988). The ability to reason about movement in the visual field. *Intelligence*, 12, 77e100.
- Jamet, E. (2006). Peut-on concevoir des documents électroniques plus efficaces ? L'exemple des diaporamas. *Revue européenne de psychologie appliquée*.58, 185-198.
- Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, 18, 135-145. doi:10.1016/j.learninstruc.2007.01.011.
- Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see : Guiding students attention via a model's eye movements fosters learning. *Learning and instruction*, 25,62-70.
- Johnson, D.S., Perlow, R. & Pieper, K.F. (1993). Differences in task performance as a function of type of feedback : Learning-oriented versus performance_oriented feedback. *Journal of Applied Social Psychology*, 23,4,pp. 303-320.
- Kalyuga, S., (2001). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, 23:1–1 9 DOI 10.1007/s10648-010-9150-7
- Kalyuga, S., Chandler, P., Sweller, J. (2000) Incorporating learner Experience into the design of multimedia instruction. *Journal of educational psychology*. 92.1,126-136.
- Kalyuga, S., Chandler, P., Sweller, J. (1998). Levels of Expertise and Instructional Design. *Human Factors*.40, 1-17.
- Kalyuga, S., Sweller, J. (2005). Rapid Dynamic Assessment of Expertise to Improve the Efficiency of Adaptive E-learning. *Educational Technology Research and Development*. 53, 3. 1042-1629.

- Kester, L., Kirschner, P.A., & van Merriënboer, J.J.G. (2005). The management of cognitive load during complex cognitive skill acquisition by mean of computer-simulated problem solving. *British Journal of Educational Psychology*, 75, 71-85.
- Khacharem A, Trabelsi K, Engel FA, Sperlich B and Kalyuga S (2020) The Effects of Temporal Contiguity and Expertise on Acquisition of Tactical Movements. *Front. Psychol.* 11:413. doi: 10.3389/fpsyg.2020.00413
- Kirschner. P. A., Sweller J., & Clark R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75-86
- Kirschner, P. & Clark, R.E. (2007). Why minimal guidance during instruction does not work : A reply to commentaries. *Educational Psychologist*, 42 (2), 115-121.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object-manipulation and perspective-taking spatial abilities. *Memory & Cognition*, 29, 745-756.
- Krassmann, A.L., Melo, M., Peixoto, B., Pinto, D., Bessa, M. & Bercht,, M. (2020). Learning in virtual reality : investigating the effects of immersive tendencies and sense of presence. J. Y. C. Chen and G. Fragomeni (Eds.): HCII 2020, LNCS 12191, pp. 270–286, 2020. https://doi.org/10.1007/978-3-030-49698-2_18
- Kulhavy, R.W., & Stock, W.A. (1989). Feedback in Written Instruction : The Place of Response Certitude. *Educational psychology Review*, 1(4), 279-308.
- Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, 9, 2, 75-82.
- 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):339-54 DOI: [10.1037/0096-3445.133.3.339](https://doi.org/10.1037/0096-3445.133.3.339)

- Lehman, S., Schraw, G., McCrudden, M.T., Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, 32, 569-587.
- Liu, R., Wang, L., Lei, J., Wang, Q. & Ren, Y. (2020). Effects of an immersive virtual reality-based classroom on students' performance in science lessons. *British Journal of Educational Technology*. Vol 51, n°6, 2034-2049. Doi:10.1111/bjet.13028
- Lowe, R.K. (2017). 1.3 Challenges for learner processing. In R.K. Lowe & R. Ploetzner (Eds), *Learning from Dynamic Visualization, Innovations in Research and application*. (p.11) Springer Nature.
- Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14, 257-274.
- Lowe, R.K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, 13,157-176.
- Lowe, R.K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, 14,225-224.
- Lowe, R., & Boucheix, J.-M. (2017). Attention Guidance Strategies for supporting learning from dynamics. In R. Lowe & R. Ploetzner (Eds.), *Learning from dynamic visualization — Innovations in research and application*. Berlin: Springer
- Lowe, R.K., & Boucheix, J.-M (2008). Learning from animated diagrams: how are mental models built? In G. Stapleton, J. Howse, & J. Lee (Eds.), *Theory and applications of diagrams* (pp.226-281) Berlin: Springer.
- Lowe, R.K.,& Ploetzner, R. (2017). Learning form dynamic visualization.*Innovation in Research and Application*. DOI 10.1007/978-3-319-56204-9
- Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, 93, 377-389. doi:10.1037/0022- 0663.93.2.377.

- Makransky, G., & Petersen, G. B. (2019a). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education*, 134, 15–30. <https://doi.org/10.1016/j.compedu.2019.02.002>.
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019b). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 225–236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>
- Makransky, G. & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-020-09586-2>
- Makransky, G., Wismer, P., & Mayer, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning*, 35(3), 349-358. <https://doi.org/10.1111/jcal.12335>
- Maresky, H. S., Oikonomou, A., Ali, I., Ditkofsky, N., Pakkal, M., & Ballyk, B. (2019). Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. *Clinical Anatomy*, 32(2), 238-243.
- Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, 81. 240-246.
- Mayer, R. E. (2002). Multimedia Learning. *The annual report of educational psychology in Japan*. 41,27-29.
- Mayer, R. E. (2009). *Multimedia Learning, second edition*, Cambridge : Cambridge University Press. http://dx.doi.org/10.1017/CBO9780_511811678
- Mayer, R. E. (Ed.), (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). New York, NY: Cambridge University Press. [http:// dx.doi.org/10.1017/CBO9781139547369](http://dx.doi.org/10.1017/CBO9781139547369)

- Mayer, R.E. (2014). 1 Introduction to multimedia learning. The Cambridge handbook of multimedia learning, p3.
- Mayer, R. E. & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83, 484-490.
- Mayer, R. E. & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R.E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more : meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of educational psychology*, 88,64-73.
- Mayer, R.E. & Chandler, P. (2001). When learning is just a click away: Does simple interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, 93, 390–397.
- 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.), *Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.015>
- Mayer, R.E., Heiser, J. and Lonn, S. (2001). Cognitive constraints on Multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187–198.
- Mayer, R.E., & Johnson, C.I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research*, 3, 241-265.
- 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*,90,312-320.

- Mayer, R.E., & Moreno, R. (2003) Nine Ways to Reduce Cognitive Load in Multimedia Learning, *Educational Psychologist*, 38:1, 43-52, doi: 10.1207/S15326985EP3801_6
- Mayer, R.E., & Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing Constructivist Learning From Multimedia Communications by Minimizing Cognitive Load. *Journal of Educational Psychology*, 4, 638-643.
- Mayer, R.E. & Sims, K. (1994). For whom is a picture worth a thousand words ? Extensions of a dual-coding theory of multi-media learning. *Journal of Educational Psychology*, **86**, 389–401.
- Mayer, R.E., Steinhoff, K., Bower, G. & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *ETR&D* **43**, 31–41 (1995). <https://doi.org/10.1007/BF02300480>
- Mestre, D. R. (2006). Immersion and Presence. *Le traité de la réalité virtuelle, Paris: Ecole des Mines de Paris*, 309-338
- Meyerhoff, H. S., & Huff, M. (2016). Semantic congruency but not temporal synchrony enhances long-term memory performance for audio-visual scenes. *Memory & Cognition*, 44(3), 390-402. doi :10.3758/s13421-015-0575-6
- Michas, I. C., & Berry, D. C. (2000). Learning a procedural task: effectiveness of multimedia presentations. *Applied Cognitive Psychology*, 14, 555-575.
- Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev.* 1956;63:81–97.
- Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, 56(3), 769–780. <https://doi.org/10.1016/j.compedu.2010.10.020>.

- Moreno, R., & Mayer, R. E. (1998). A Split-Attention effect in Multimedia Learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90, 2, 312-320.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.
- Moreno, R. & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, 97, 117-125.
- Moreno, R., & Mayer, R. E. (2005). Role of Guidance, Reflection, and Interactivity in an Agent-Based Multimedia Game. *Journal of Educational Psychology*, 1, 117-128.
- Najjar, L. J. (1998). Principles of educational multimedia user interface design. Human Factors, 40(2), 311-323.
- Naumann, J., Richter, T., Flender, J., Christmann, U., & Groeben, N. (2007). *Journal of educational psychology*. 4, 791-807.
- Owens, P. & Sweller, J. (2008). Cognitive load theory and music instruction. *Educational Psychology*, 29-45, <https://doi.org/10.1080/01443410701369146>
- Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye tracking. *Computer in Human Behavior*, 26, 110-117. doi:10.1016/j.chb.2009.09.001.
- Paas, F. & Sweller, J. (2012). An Evolutionary Upgrade of Cognitive Load Theory: Using the Human Motor System and Collaboration to Support the Learning of Complex Cognitive Tasks. *Educational Psychology Review*, 24, 27-45. DOI 10.1007/s10648-011-9179-2
- Paivio, A. (1986). *Mental representations : A dual-coding approach*. Oxford, England : Oxford University Press.

- Paivio, A. (2006). Dual coding theory and education. *Educational psychology review*. DOI: 10.1007/BF01320076
- Park, B., Moreno, R., Seufert, T., Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*. 27,5-10.
- Pellegrino, J. W., & Hunt, E. B. (1989). Computer-controlled assessment of static and dynamic spatial reasoning. In R. F. Dillon & J. W. Pellegrino (Eds.), *Testing: Theoretical and applied perspectives* (pp. 174–198). New York: Praeger. x c
- Picard, A., Seigneuric, A., de Guibert, M., Hakima, M., & Potier, D. (2008). La MTVE : une épreuve de mémoire de travail verbale chez l'enfant adaptée du Listening Span Task de Siegel et Ryan (1989). *L'année psychologique*. 108, 2, 273-308
- Ploetzner, R., Berney, S. & Bétrancourt, M. (2020). A review of learning demands in instructional animations: The educational effectiveness of animations unfolds if the features of change need to be learned. *Journal of Computer Assisted Learning WILEY*. DOI: 10.1111/jcal.12476
- Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, 28, 781–794. <https://doi.org/10.1016/j.chb.2011.12.001>
- Porte, L., Boucheix, J.M., Bétrancourt, M., Lowe, R.K., Ceddia, M., & Bard, P. (2016). Dynamic spatial abilities and learning from animation. In Désiron, J., Berney, S., Bétrancourt, M., & Tabbers, H. (Eds.), *EARLI SIG 2 Conference: Comprehension of Text and Graphics, Learning from Text and Graphics in a World of Diversity* (pp. 145-148). Geneva, July 11-13.: UNiversity of Geneva.
- Porte, L., Boucheix, J.-M., Bétrancourt, M., Lowe, R. K., Argon, S. & Bard, P. (2018, August). Dynamic spatial abilities in transformation task and learning from biological animation. Poster presented at EARLI SIG2, Comprehension of Text and Graphics. Freiburg, August 27-31. : University of Freiburg, Germany.

- Porte, L., Boucheix, J.-M., Lowe, R. K., Bétrancourt, M., & Bard, P. (2019, Juillet). *Effet des habiletés de raisonnement spatial dynamique dans l'utilisation de jeux vidéo et sur l'apprentissage d'animations complexes*. Paper presented at EPIQUE 2019, 10^{ème} Colloque de Psychologie Ergonomique de l'Association pour la Recherche en Psychologie Ergonomique (ARPEGE), Lyon, France.
- Pulijala, Y., Ma, M., Pears, M., Peebles, D.& Ayoub, A. (2017). Effectiveness of immersive virtual reality in surgical training : a randomized control trial. *Journal of oral and maxillofacial surgery*, 76 (5), p1065-1072. <https://doi.org/10.1016/j.joms.2017.10.002>
- Robertson, G., Card, S.K. and Mackinlay, J. (1993) Three views of virtual reality: nonimmersive virtual reality, *Computer*, 26(2), p.81.
- Roy, E., Bakr, M.M., & George,R. (2017). The need for virtual reality simulators in dental éducation : a review. *The Saudi Dental Journal*, King Saud University. 1013-9052 <http://dx.doi.org/10.1016/j.sdentj.2017.02.001>
- Sanchez, C.A., & Wiley, J. (2014). The role of dynamic spatial abilities in geosciences text comprehension. *Learning and Instruction*, 31, 33-45.
- Saputro, S.D. & Setyawan, A. (2020). The effectiveness Use of Virtual Reality Media in Physics Education of Solar System Toward Cognitive Learning Outcomes. *Jurnal Pendidikan Indonesia*, Vol.9 N^o3. DOI: 10.23887/jpi-undiksha.v9i3.23105
- Schmidt, H. G., Loyens, S. M. M., van Gog, T., & Paas, F. (2007). Problembased learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42, 91–97.
- Schneider, E. & Boucheix, J.-M. (2007). « How to improve complex mechanical system's comprehension with animations? ». *Proceedings of the 12th Biennial Conference for Research on Learning and Instruction (EARLI)* ; Budapest : August 28-September 1, 2007.

- Schnotz, W. (2014). "Integrated model of text and picture comprehension," in *The Cambridge Handbook of Multimedia Learning*, 2nd Edn, ed. R. E. Mayer (New York, NY: Cambridge University Press), 72–103. doi: 10.1017/CBO9780511816819.005
- Schnotz, W. (1999). Towards an Integrated View of Learning From Text and Visual Displays, *Educational Psychology Review*, 14, 101-120.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In Mayer, R. E. (Ed.), (2005). *The Cambridge handbook of multimedia learning*. Pp. 49-70. New York, NY: Cambridge University Press
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 72–103). Cambridge University Press.
- Schnotz, W., Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141-156. Doi:10.1016/S0959-4752(02)00017-8
- Schnotz, W. & Rasch, T. (2005). Enabling, Facilitating, and Inhibiting Effects of Animations in Multimedia Learning: Why Reduction of Cognitive Load Can Have Negative Results on Learning. *Educational Technology Research and Development*. Vol. 53, No.3, pp. 47–58 ISSN 1042–1629
- Schnotz, W., & Wagner, I. (2018). Construction and elaboration of mental models through strategic conjoint processing of text and pictures. *Journal of Educational Psychology*, 110, 850–863. <http://dx.doi.org/10.1037/edu0000246>
- Shah, P., & Myaké, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of experimental Psychology*, 125, 4-27.

- Scheiter, K., & Eitel, A. (2010). The Effects of Signals on Learning from Text and Diagrams: How Looking at Diagrams Earlier and More Frequently Improves Understanding. Conference Paper · August 2010 DOI: 10.1007/978-3-642-14600-8_26 · Source: DBLP
- Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, 30, 679-701. doi:10.1007/s10648-018-9435-9
- Schüler, A., Arndt, J., Scheiter, K., (2019). Does text-picture integration also occur with longer text Segments? *Applied cognitive Psychology*, 33, 1137- 1146, <https://doi.org/10.1002/acp.3558>
- Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2012). Explaining the modality effect in multimedia learning : Is it due to a lack of temporal contiguity with written text and pictures ? *Learning and Instruction*, 22(2), 92-102.
- Schwan, S., Dutz, S., & Dreger, F. (2018). Multimedia in the wild: Testing the validity of multimedia learning principles in an art exhibition. *Learning and Instruction*, 55,148–157.
- Stull, A.T., & Mayer, R.E., (2007). Learning by Doing Versus Learning by Viewing: Three Experimental Comparisons of Learner Generated Versus Author-Provided Graphic Organizers. *Journal of educational psychology*. 4, 808-820.
- Sung, E., & Mayer, R.E., (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*. 28, 1618-1625.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational psychology Review*, 22,123-138.

- Sweller, J. (2012). Humain cognitive architecture: why some instructional procedures work and others do not. In K. Harris, S. Graham & T. Urdan (Eds.), *APA educational psychology handbook* (vol. 1, pp. 295-325). Washington, DC : American Psychological Association.
- Sweller, J., Ayres, P., & Kalyuga, S. (2003). The Expertise Reversal Effect. *Educational psychologist*, 38(1), 23-31.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176-192.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251e296.
- Teranishi, S. & Yamagishi, Y. (2017). Educational Effects of a Virtual Reality Simulation System for constructing Self-Built PCs. In J. Johnston (Ed.), *Proceedings of EdMedia 2017* (pp. 1214-1222). Washington, DC: Association for the Advancement of Computing in Education (AACE). Retrieved September 26, 2021 from <https://www.learntechlib.org/p/178439>.
- Thomas, R., Hooper, E., 1991. Simulations: an opportunity we are missing. *J. Res. Comput. Educ.* 23, 497–513.
- Tricot, A (1998). Charge cognitive et apprentissage. Une présentation des travaux de John Sweller. *Revue de Psychologie de l'Éducation*, 1, 37-64.
- Van Buskirk, W. (2011). *Investigating The Optimal Presentation Of Feedback In Simulation-based Training An Application Of The Cognitive Theory Of Multimedia Learning*. Electronic Theses and Dissertations, 2004-2019. 1986. <https://stars.library.ucf.edu/etd/1986>

Wan Ahmad Jaafar Wan Yahaya & Awatif Ahmad (2017) The effectiveness of signaling principle in virtual reality courseware towards achievement of transfer learning among students with different spatial ability. *AIP Conference Proceedings* 1891, 020144 (2017); <https://doi.org/10.1063/1.5005477>

Wittrock, M.C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 485-498.

Xie, H., Mayer, R. E., Wang, F., & Zhou, Z. (2018). Coordinating visual and auditory cueing in multimedia learning. *Journal of Educational Psychology*, 111(2), 235-255. <https://doi.org/10.1037/edu0000285>

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.

Ressources internet :

<http://aqnp.ca/laneuropsychologie/lesfonctionscoGNitives/#Fonctionsvisuospatiales>

<http://www.edu.gov.on.ca/fre/literacynumeracy/SpatialReasoningFr.pdf> National Research Council p6.

<https://www.biodiversiteetbati.fr/Files/Other/FT%20BPU/FT13-StratificationVegetale.pdf>

Ressources Silva Numerica :

Tutoriel d'utilisation du logiciel par Christophe Corne-Artibiocom :

https://ent.artibiocom.com/ressources/silvanumerica/Guide_Eve/co/INTERFACE_GENERALE.html

vidéo de démonstration : <https://www.youtube.com/watch?v=9SNxvjjaxcU>

ANNEXES

A	Test d'habiletés spatiales, Shah & Miyaké (1996)	P 222
B	QCM de pré :post-test de connaissances factuelles	P 228
C	Images issues de la leçon : la décomposition de la matière organique.	P 233
D	Pictogrammes présentés avant la leçon en RV non immersive	P 234
E	Script de la narration de la leçon en RV non immersive	P 235
F	Test de mémoire de travail verbale MTVE	P 239
G	Test de correspondances texte-image	P 242
H	Test d'orientation spatiale	P 251
I	Test de mémoire de travail verbale Daneman & Carpenter.	P 265
J	Contenu des balises	P 268
K	Questionnaire RV	P 269
L	Contenu du feedback pour l'expérience en RVI	P 271
M	Consignes et commandes du didacticiel en RVI	P 272
N	Consignes de l'expérimentation en RVI	P 273

Annexe A.

Le test d'habileté spatiale, Shah et Miyaké (1996)

jeux de visualisation dans l'espace

Nom: _____
Prénom: _____
Classe: _____
Ecole: _____

Ne rien inscrire dans le tableau

Subtest	Modèle	Empan	Total
Lettres			
Objets			
Formes			
Total			

P

1

P

P

P

P

P

Ne tournez pas la page avant le signal

B

2

B

B

B

B

B

Ne tournez pas la page avant le signal

3

K

K

K

K

K

K

Ne tournez pas la page avant le signal

4

G

G

G

G

G

G

Ne tournez pas la page avant le signal

5

F

F

F

F

F

F

Ne tournez pas la page avant le signal

R

6



Ne tournez pas la page avant le signal



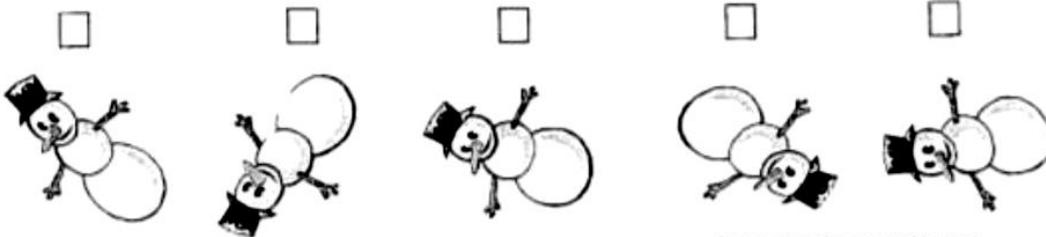
7



Ne tournez pas la page avant le signal



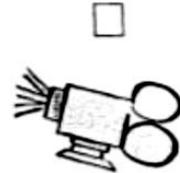
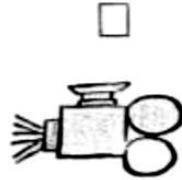
8



Ne tournez pas la page avant le signal



9



Ne tournez pas la page avant le signal



10



Ne tournez pas la page avant le signal



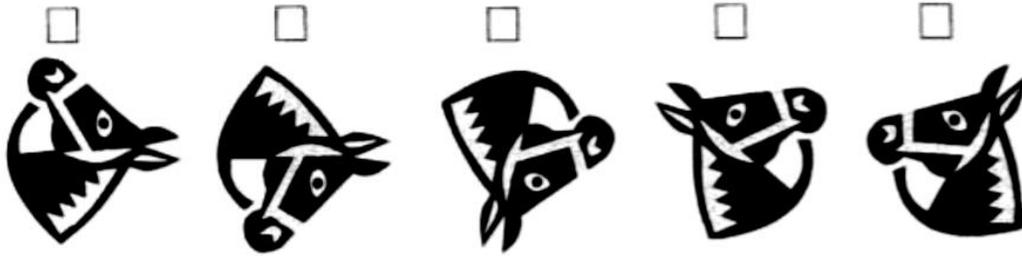
11



Ne tournez pas la page avant le signal



12



Ne tournez pas la page avant le signal



13



Ne tournez pas la page avant le signal



14



Ne tournez pas la page avant le signal



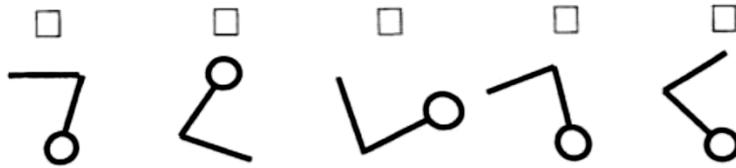
15



Ne tournez pas la page avant le signal



16



Ne tournez pas la page avant le signal



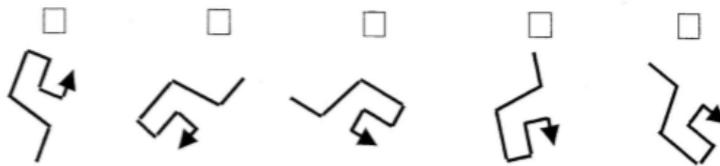
17



Ne tournez pas la page avant le signal



18



Ne tournez pas la page avant le signal

Annexe B.

Pré/post-test de connaissances factuelles (QCM) proposé aux élèves de collège.



PR1



Collège Rameau, Dijon

Mars 2019.

Date :

Nom :

prénom :

CLASSE :

SEXE :

Masculin/ Féminin

AGE :



Bonjour !

Que sais-tu de la décomposition de la matière organique ? Cela m'intéresse beaucoup alors je vais te poser quelques questions à ce sujet. Ne t'en fais pas, ceci n'est pas un contrôle tu ne seras pas noté 😊

Je vais te demander de lire chacune des affirmations du tableau et de répondre par VRAI ou par FAUX. Coche ou entoure la réponse de ton choix.

Essaies de répondre à un maximum de questions par VRAI ou par FAUX, mais si vraiment tu ne connais pas la bonne réponse, alors coche ou entoure NSP pour me dire que tu ne sais pas.

Tu peux commencer dès que tu es prêt ! Prends le temps dont tu as besoin 😊

1	Sur les feuilles mortes on peut voir de petits fils blancs : il s'agit d'un champignon appelé mycélium.	VRAI	FAUX	NSP
2	Les bactéries du sol transforment le caca des animaux en matière minérale.	VRAI	FAUX	NSP
3	Le collembole, le cloporte et l'oribate fragmentent la matière organique.	VRAI	FAUX	NSP
4	Le caca des animaux du sol est transformé en humus par les bactéries.	VRAI	FAUX	NSP
5	Le caca des petits animaux du sol sert de nourriture aux prédateurs.	VRAI	FAUX	NSP
6	Le ver de terre mange les géophiles et pseudoscorpions.	VRAI	FAUX	NSP
7	Le humus est une couche de terre humide qui permet aux végétaux de se nourrir.	VRAI	FAUX	NSP
8	Les bactéries et les champignons ont besoin d'eau et d'air pour se développer.	VRAI	FAUX	NSP
9	Le humus se transforme en matière minérale quand l'eau qu'il contient s'évapore.	VRAI	FAUX	NSP
10	Le géophile et le pseudoscorpion mangent des brindilles.	VRAI	FAUX	NSP
11	Les petits animaux qui fragmentent la matière organique sont mangés par des prédateurs.	VRAI	FAUX	NSP

12	Le ver de terre est le seul animal à manger les feuilles de la litière.	VRAI	FAUX	NSP
13	les micro-organismes mangent des bactéries et des champignons.	VRAI	FAUX	NSP
14	Certains animaux découpent la matière organique en petits morceaux pour la manger.	VRAI	FAUX	NSP
15	Les micro-organismes sont les êtres vivants qui mangent la matière organique en premier.	VRAI	FAUX	NSP
16	La matière minérale sert de nourriture aux vers de terre.	VRAI	FAUX	NSP
17	Le ver de terre digère la matière minérale du sol.	VRAI	FAUX	NSP
18	Les bactéries mangent les petits animaux qui fragmentent la matière organique.	VRAI	FAUX	NSP
19	Les végétaux se nourrissent de la matière minérale du sol avec leurs racines.	VRAI	FAUX	NSP
20	Les êtres vivants et ce qu'ils rejettent sont de la matière organique.	VRAI	FAUX	NSP
21	Le pseudo-scorpion et le géophile sont des prédateurs.	VRAI	FAUX	NSP
22	Les roches et les excréments de bactéries sont de la matière organique.	VRAI	FAUX	NSP

23	Le humus est composé de poils de plumes, de brindilles et de feuilles.	VRAI	FAUX	NSP
24	Le humus permet aux petits animaux de se développer.	VRAI	FAUX	NSP
25	Le caca des animaux sert de nourriture aux bactéries du sol.	VRAI	FAUX	NSP
26	Les animaux qui mangent la matière organique participent à la création du humus.	VRAI	FAUX	NSP
27	Les champignons et les bactéries fragmentent la matière organique.	VRAI	FAUX	NSP
28	Les prédateurs sont des animaux qui mangent des champignons.	VRAI	FAUX	NSP
29	Le collembole, le cloporte et l'oribate sont les premier à se nourrir de la matière organique.	VRAI	FAUX	NSP
30	La matière organique est composée de roches et de terre.	VRAI	FAUX	NSP
31	Les bactéries du sol rendent les animaux de la litière malades.	VRAI	FAUX	NSP
32	Le humus, c'est de la matière organique digérée et rejetée par les animaux du sol.	VRAI	FAUX	NSP
33	Le ver de terre digère la terre et les feuilles de la litière.	VRAI	FAUX	NSP

34	Les êtres vivants et ce qu'ils rejettent sont de la matière organique.	VRAI	FAUX	NSP
35	Les micro-organismes prémâchent la matière organique pour que d'autres êtres vivants puissent la manger ensuite.	VRAI	FAUX	NSP
36	La litière du sol est composée de feuilles, de brindilles, de poils et de plumes.	VRAI	FAUX	NSP

Merci beaucoup pour ta participation !



Annexe C.

Images issues de la vidéo sur la décomposition de la matière organique.



B1 : présentation de la forêt.



B2 : promenade autour des arbres.



B3 : zoom sur les bactéries.



B4 : les couches du sol.



B5 : les éléments de la litière.



B6 : zoom sur l'humus.

Annexe D.

Pictogrammes présentés en début de vidéo.

Bonjour ! avant de commencer la leçon, regarde les images ci-dessous.



quand tu vois un œil comme celui-ci, c'est que les éléments présentés sont visibles à l'œil nu.



quand tu vois une loupe comme celle-ci, c'est que les éléments présentés nécessitent qu'on s'approche vraiment pour les observer.



quand tu vois un microscope comme celui-ci, c'est que les éléments présentés sont invisibles à l'œil nu.

Annexe E.

Script de la narration : la décomposition de la matière organique.

La décomposition de la matière organique

Bonjour! aujourd'hui, nous allons parler de la décomposition de la matière organique.

Partons en forêt à la découverte du sol !!

Pour commencer, regardons la surface du sol.

Qu'y-a-t-il sous tes pieds?

Ce qu'on voit ici, c'est ce que l'on appelle la litière.

Elle est composée de différents éléments que tu connais bien!

Comme on peut le voir, il s'agit d'une feuille. Les arbres perdent leurs feuilles, celles-ci viennent alors recouvrir le sol de nos forêts.

Les branches et les brindilles des arbres peuvent casser et tomber, on en retrouve alors sur le sol.

Les animaux perdent leurs poils. On en retrouve donc par terre.

De même, Les oiseaux perdent leurs plumes, on peut donc aussi en retrouver sur sol.

Ces éléments ont un point commun : ils proviennent tous d'êtres vivants !

Les branches et les feuilles sont produites par les arbres et plus généralement par les végétaux.

Les poils et les plumes sont produits par les animaux.

Ils font partie de ce que l'on appelle la matière organique.

Sur le sol, les êtres vivants et ce qu'ils rejettent sont de la matière organique. Les animaux comme leurs cadavres sont eux aussi de la matière organique.

Cette matière organique se décompose, mais comment ?

« Voyons quels sont les acteurs de la décomposition du sol.

Quand on se promène en forêt on a l'habitude de voir la litière à la surface du sol.

Elle est composée de matière organique et cette matière organique est faite de débris produits par les êtres vivants.

Mais on peut se poser la question suivante : pourquoi l'épaisseur de la litière n'augmente pas avec le temps ?

Ici, on peut voir les différentes couches du sol. La première est une fine couche de litière. »

Voici les feuilles de la litière. Certaines sont trouées : C'est parce qu'elles se transforment, elles se décomposent. Voyons comment se déroule cette décomposition.

La litière sert de nourriture à de nombreux êtres vivants. Comme ils ne mangent pas tous les mêmes éléments de la litière au même moment, nous allons les observer ensemble.

Les premiers êtres vivants qui participent à la décomposition de la matière organique sont les micro-organismes.

On les appelle micro-organismes parce qu'ils sont très petits. D'ailleurs, nous avons besoin d'un microscope pour les observer! Voyons à quoi ressemblent ces microorganismes.

« Sur cette feuille on peut voir de petits fils blancs. Ce sont des filaments de champignon! On appelle cela le mycélium. On l'observe à l'aide d'un microscope.

Pour se développer, le mycélium a besoin d'eau et d'air. Il se nourrit de matière organique morte, c'est pourquoi on peut l'observer sur les feuilles mortes.

Il est indispensable à la décomposition de la matière organique car il se nourrit de ce que la plupart des autres êtres vivants ne peuvent pas manger.

« Dans la litière, on trouve aussi des bactéries. Pour se développer, elles ont besoin d'humidité et d'air. Elles se nourrissent de matière organique.

Les bactéries mangent ce que les autres animaux ne peuvent pas digérer. Elles prémâchent la litière pour permettre aux petits animaux peuplant le sol de se nourrir.

Une fois que les microorganismes ont prémâché la litière, les animaux décomposeurs entrent en jeu.

Certains de ces animaux fragmentent la litière. Mais qu'est-ce que cela signifie: fragmenter?

Cela veut dire qu'ils vont découper la matière organique en petits morceaux pour pouvoir la manger. »

Les oribates sont de petits animaux qui mesurent entre 0,6 et 1,5 mm et vivent dans la litière. Ils mangent la matière organique qu'ils découpent en petits morceaux ainsi que les champignons.

Les collemboles sont de petits animaux qui mesurent moins de 3 mm de long! Ils mangent les végétaux en décomposition et les champignons du sol. On les retrouve dans la litière mais aussi plus profondément dans le sol. Leur rôle est de fragmenter la matière organique. »

Le cloporte mesure 0,5 à 2cm de long. On peut le rencontrer dans la litière. Cet animal ne digère que les végétaux morts c'est-à-dire les détritux végétaux.

Les petits animaux que nous avons découverts peuvent eux aussi être mangés. Voyons ensemble qui sont leurs prédateurs. »

Tout d'abord, il y a le pseudoscorpion. Il mesure entre 2 et 5mm de long. Il vit dans la litière. Il mange d'autres animaux comme les collemboles et les oribates.

Il y a aussi le Géophile. Il mesure 4cm de long et vit à la fois dans la litière et dans le humus. Le humus est la couche du sol situé sous la litière. Carnivore, il chasse les cloportes et les vers de terre.

Si nous regardons plus profondément dans le sol, nous pourrions observer le ver de terre.

Le ver de terre est appelé lombric. Il mesure environ 12cm de long et vit dans la terre. Il aime l'humidité, c'est pourquoi on peut parfois en voir par temps de pluie.

Il sort de terre pour se nourrir des feuilles de la litière. En creusant la terre pour entrer ou sortir du sol, il crée des galeries. Celles-ci permettent d'aérer le sol ce qui est très important! En effet, un sol bien aéré est un sol riche car cela permet aux champignons et aux bactéries de se développer. Le ver de terre digère la terre ainsi que les déchets de la litière. Il les rejette sous forme d'excréments sous terre ou sur terre.

On peut parfois voir ceci sur le sol, n'est pas simplement un petit tas de terre. Il s'agit en fait d'excréments, c'est-à-dire de caca de ver de terre! On l'appelle turricule parce qu'il ressemble à une petite tour ». C'est le résultat du travail du ver de terre.

Observons maintenant le résultat du travail de tous ces êtres vivants.

Le travail de tous les animaux que nous venons de voir participe à la formation de ce que l'on appelle le humus.

Cette couche de terre noire et humide située sous la litière, c'est de l'humus. Elle est le fruit du travail de tous les êtres vivants que nous venons d'étudier!

Ils mangent, digèrent puis rejettent la matière organique de la litière et créent ainsi une belle couche d'humus.

Cette couche de terre est extrêmement importante car elle permet de garder un sol stable et fertile. Cela veut dire qu'elle permet aux végétaux de pousser, de se nourrir. »

Sous cette couche de humus, il y a une couche que l'on appelle : la matière minérale. Voyons comment elle se forme.

Tu vois, sous la couche de humus, on observe une couche un peu plus claire. Cette partie du sol est composée de matière minérale!

Après la création de humus, il reste encore une étape!

Comme on a pu le voir, lorsqu'ils consomment la matière organique, les animaux la rejettent. Tous ces animaux font caca, mais à quoi servent leurs excréments ?

Ils servent à nourrir les bactéries !

Les bactéries mangent le caca, c'est-à-dire les excréments produits par les animaux et les transforment en matière minérale.

Avec leurs racines, les végétaux peuvent alors se nourrir de cette matière minérale.

Conclusion (voir si on la laisse en fonction de la vidéo).

Nous venons d'apprendre beaucoup de choses passionnantes sur le sol de nos forêts!

C'est quand même étonnant... Les êtres vivants du sol sont si petits et pourtant tellement indispensables! Ils permettent aux végétaux de vivre à l'aide de leurs propres déchets!

Les feuilles que l'arbre laisse tomber lui permettent d'en faire pousser de nouvelles à chaque saison!

Annexe F.

Test de mémoire de travail verbale, MTVE.

ENTRAÎNEMENT

Essai 1 :

- 1) En automne, les arbres perdent leurs (feuilles)
 - 2) On va à la boulangerie pour acheter du (pain)
- (Rappel : feuilles, pain.)

Essai 2 :

- 1) Une pomme c'est rouge, une banane c'est (jaune)
 - 2) Le dos du chameau a deux (bosses)
- (Rappel : jaune, bosses)

Essai 3 :

- 1) Au foot, on tape avec le pied dans un (ballon)
 - 2) À la bibliothèque, on peut prendre des BD ou des (livres)
- (Rappel : ballon, livres)

TEST

Série de 2 phrases

Essai 1 :

- 1) On mange de la soupe avec une (cuillère)
- 2) Sur mon gâteau d'anniversaire, j'ai soufflé les (bougies)

Essai 2 :

- 1) Pour se faire couper les cheveux, on va chez le (coiffeur)
- 2) On se lave les mains avec de l'eau et du (savon)

Essai 3 :

- 1) On jette les ordures et les vieux papiers à la (poubelle)
- 2) Quand on voit mal, il faut porter des (lunettes)

Série de 3 phrases

Essai 1 :

- 1) Quand je vais à l'école, je mets ma trousse dans mon (cartable)
- 2) Pour planter un clou, on se sert d'un (marteau)
- 3) Le chat miaule et le chien (aboie)

Essai 2 :

- 1) L'enfant a de la fièvre, il est (malade)
- 2) On va à la pêche pour attraper des (poissons)
- 3) Le roi et la reine habitent dans un (château)

Essai 3 :

- 1) Le dimanche tous les magasins sont (fermés)
- 2) Avant de poster la lettre, j'ai collé un (timbre)
- 3) Après un bain, je m'essuie avec une..... (serviette)

Série de 4 phrases

Essai 1 :

- 1) Avant de s'endormir, on éteint la (lumière)
- 2) Quand la chambre est en désordre, il faut la (ranger)
- 3) Pour faire un trait bien droit je prends une..... (règle)
- 4) Quand on veut sentir bon, on se met du (parfum)

Essai 2 :

- 1) Pour me lever à l'heure le matin, je mets un (réveil)
- 2) Un homme c'est grand, un enfant c'est (petit)
- 3) On mange avec une fourchette et un (couteau)
- 4) À Noël, on accroche des boules dans le (sapin)

Essai 3 :

- 1) À la piscine, on apprend à (nager)
- 2) Au tennis, on envoie la balle avec une (raquette)
- 3) J'ai mis un disque pour écouter de la (musique)
- 4) Quand il y a le feu, on appelle les (pompiers)

Série de 5 phrases

Essai 1 :

- 1) Pour couvrir ma tête quand il fait froid, je mets un (bonnet)
- 2) J'avais sommeil, je suis allé me..... (coucher)
- 3) Les assiettes sont sales, je dois faire la..... (vaisselle)
- 4) 1, 2, 3, 4 sont des chiffres, A, B, C, D sont des (lettres)
- 5) Quand on s'est cassé la jambe, on marche avec des (béquilles)

Essai 2 :

- 1) Je découpe du papier avec des (ciseaux)
- 2) Dans un château hanté, il y a des (fantômes)
- 3) On va à l'école pour apprendre à lire et à..... (écrire)
- 4) Dans mon café je mets deux morceaux de..... (sucre)
- 5) Chez le boucher on achète de la (viande)

Essai 3 :

- 1) Au supermarché, on va à la caisse pour (payer)
- 2) La nuit on voit la lune, le jour on voit le (soleil)
- 3) Pour savoir l'heure, je regarde sur ma (montre)
- 4) La maîtresse écrit avec une craie sur le (tableau)
- 5) Pourriez vous me passer le sel et le (poivre)

Annexe G.

Post-test de correspondance texte-image.

Date :

Nom :

prénom :

CLASSE :

SEXE :

Masculin/ Féminin

AGE :



Bonjour !

Tu viens de visualiser une vidéo qui parle de la décomposition de la matière organique, tu as dû apprendre beaucoup de choses ! J'aimerais savoir de quoi tu te souviens. Pour cela, je vais te proposer un petit jeu 😊

Je vais te présenter des images et des phrases. Pour chaque phrase, je vais te proposer deux images. Quand tu entendais la phrase dans la vidéo, quelle image voyais-tu en même temps? Peux-tu cocher l'image qui correspond à la phrase que tu as entendue dans la vidéo ?

Condition -6

« Les Oribates sont de petits animaux qui mesurent entre 0,6 et 1,5 mm et vivent dans la litière »



« Les Oribates sont de petits animaux qui mesurent entre 0,6 et 1,5 mm et vivent dans la litière »



« Les branches et les brindilles des arbres peuvent casser et tomber ! On en retrouve alors sur le sol »



« Les branches et les brindilles des arbres peuvent casser et tomber ! On en retrouve alors sur le sol »



« Ici, on peut voir les différentes couches du sol »



« Ici, on peut voir les différentes couches du sol »



Condition -6

« Sur cette feuille, on peut voir de petits fils blancs. Ce sont des filaments de champignons. On appelle cela le mycélium »



« Sur cette feuille, on peut voir de petits fils blancs. Ce sont des filaments de champignons. On appelle cela le mycélium »



« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »



« Le cloporte mesure 0,5, à 2cm de long. On peut le rencontrer dans la litière. »



« On peut parfois voir ceci sur le sol. Ce n'est pas simplement un petit tas de terre ! »



« On peut parfois voir ceci sur le sol. Ce n'est pas simplement un petit tas de terre ! »



« Mais on peut se poser la question suivante : pourquoi l'épaisseur de la litière n'augmente pas avec le temps ? »



« Mais on peut se poser la question suivante : pourquoi l'épaisseur de la litière n'augmente pas avec le temps ? »



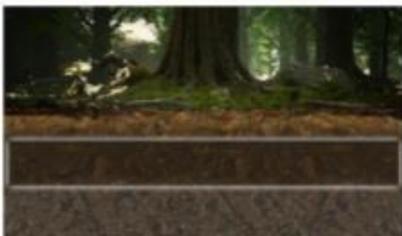
« Ils vont découper la matière organique en petits morceaux pour pouvoir la manger. »



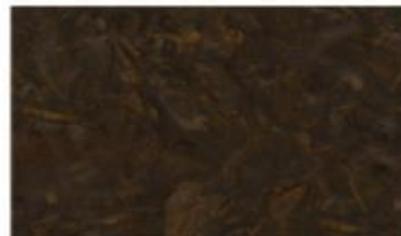
« Ils vont découper la matière organique en petits morceaux pour pouvoir la manger. »



« Cette couche de terre noire et humide est située sous la litière, c'est de l'humus. »



« Cette couche de terre noire et humide est située sous la litière, c'est de l'humus. »



Condition #6

« Tu vois, sous la couche de humus, on observe une couche un peu plus claire »



« Tu vois, sous la couche de humus, on observe une couche un peu plus claire »



« Tous ces animaux font caca, mais à quoi servent leurs excréments ? »



« Tous ces animaux font caca, mais à quoi servent leurs excréments ? »



« Ce qu'on voit ici, c'est ce que l'on appelle : la litière »



« Ce qu'on voit ici, c'est ce que l'on appelle : la litière »



« Tout d'abord, il y a le pseudo-scorpion. »



« Tout d'abord, il y a le pseudo-scorpion. »



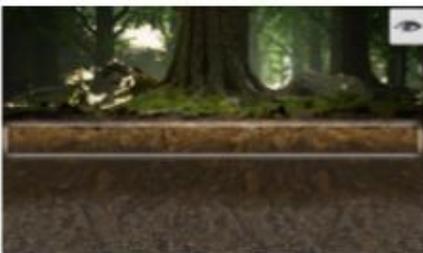
« Ils servent à nourrir les bactéries. »



« Ils servent à nourrir les bactéries. »



« La première, est une fine couche de litière. »



« La première, est une fine couche de litière. »



Condition #2

« Les animaux perdent leurs poils. On en retrouve donc par terre. »



« Les animaux perdent leurs poils. On en retrouve donc par terre. »



« Les poils et les plume sont produits par les animaux. »



« Les poils et les plume sont produits par les animaux. »



« Ils se nourrit de ce que la plupart des autres êtres vivants ne peuvent pas manger. »



« Ils se nourrit de ce que la plupart des autres êtres vivants ne peuvent pas manger. »



Condition +2

« De même, les oiseaux perdent leurs plumes, on peut donc aussi en retrouver sur le sol. »



« De même, les oiseaux perdent leurs plumes, on peut donc aussi en retrouver sur le sol. »



« D'ailleurs, nous avons besoin d'un microscope pour les observer ! »



« D'ailleurs, nous avons besoin d'un microscope pour les observer ! »



« Carnivore, il chasse les cloportes et les vers de terre. »



« Carnivore, il chasse les cloportes et les vers de terre. »



Condition -2

« Les collemboles sont de petits animaux qui mesurent moins de 3mm de long. »



« Les collemboles sont de petits animaux qui mesurent moins de 3mm de long. »



« Le travail de tous les animaux que nous venons de voir participe à la formation de ce que l'on appelle : le humus »



« Le travail de tous les animaux que nous venons de voir participe à la formation de ce que l'on appelle : le humus »



« Après la création de humus, il reste encore une étape. »



« Après la création de humus, il reste encore une étape. »



Annexe H.

Test d'orientation spatiale.

Nom :	établissement :
Prénom :	date :
Classe :	âge :

Test d'orientation spatiale

Ceci est un test destiné à mesurer votre capacité à imaginer différentes perspectives ou orientations dans l'espace. Sur chacune des pages suivantes, vous verrez une image présentant un ensemble d'objets et « un cercle fléché » avec une question sur la direction dans laquelle se trouvent certains de ces objets. Pour la question figurant sur chaque page, vous devez imaginer que vous vous tenez au niveau d'un objet (qui sera nommé au centre du cercle) et que vous faites face à un autre objet, nommé au sommet du cercle. Votre tâche est de dessiner une flèche depuis l'objet du centre du cercle et qui pointe en direction du 3ème objet. Regardez l'exemple à la page suivante. Dans cet exemple, on vous demande d'imaginer que vous vous tenez au niveau de la fleur, qui est nommée au centre du cercle, et vous regardez en direction de l'arbre, qui est nommé au sommet du cercle. Votre tâche est de dessiner une flèche pointant en direction du chat. Dans cet exemple, la flèche a été dessinée pour vous. Dans les items du test, votre tâche sera de dessiner cette flèche. Est-ce que vous voyez bien que si vous étiez à l'endroit de la fleur, faisant face à l'arbre, le chat serait dans cette direction ? Demandez maintenant à l'expérimentateur si vous avez des questions au sujet de ce que vous devez faire.

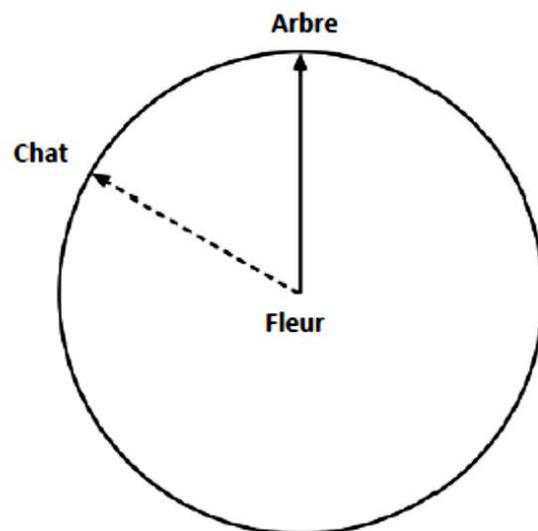
Il y a 12 items dans ce test, un item par page. Pour chaque item, la série d'objets est montrée en haut de la page, et le cercle est visible en bas de la page. Essayez de dessiner les directions exactes mais ne perdez pas trop de temps sur chacune des questions.

Vous aurez 5 minutes pour faire ce test.

Exemple

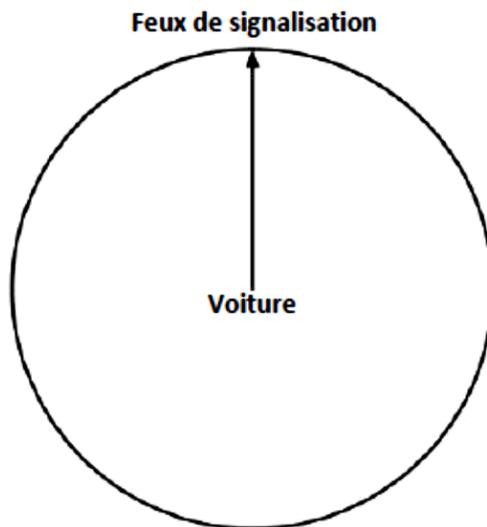


Imaginez que vous êtes positionné-e à l'emplacement de la fleur et que vous regardez l'arbre. Indiquez la direction du chat.



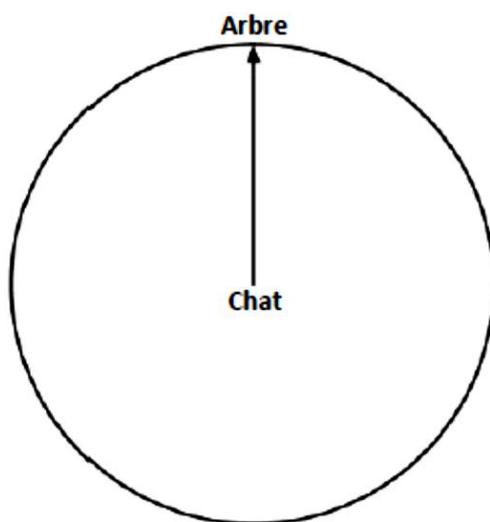


1. Imaginez que vous êtes positionné-e à l'emplacement de la **voiture** et que vous regardez les **feux de signalisation**. Indiquez la direction du **panneau stop**.



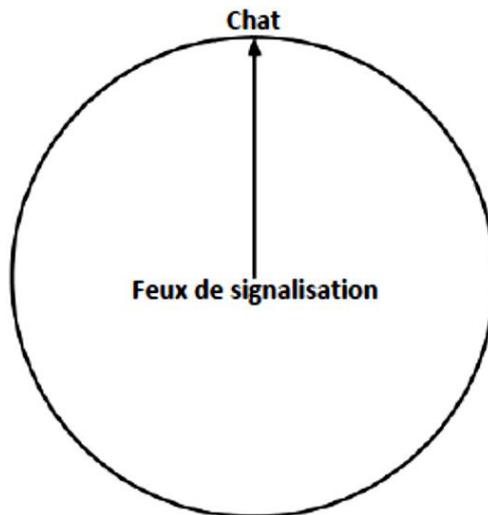


2. Imaginez que vous êtes positionné-e à l'emplacement du **chat** et que vous regardez l'**arbre**. Indiquez la direction de la **voiture**.



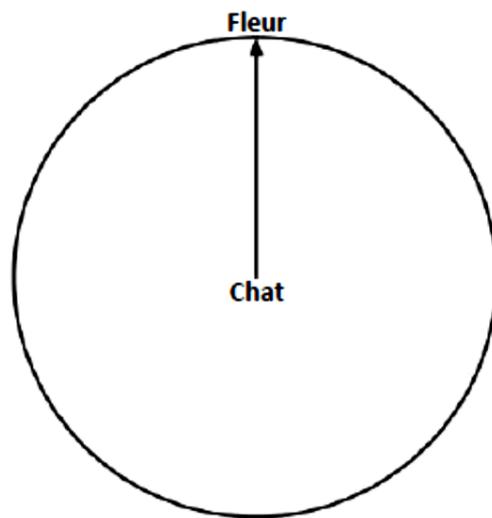


3. Imaginez que vous êtes positionné-e à l'emplacement des **feux de signalisation** et que vous regardez le **chat**. Indiquez la direction de la maison.



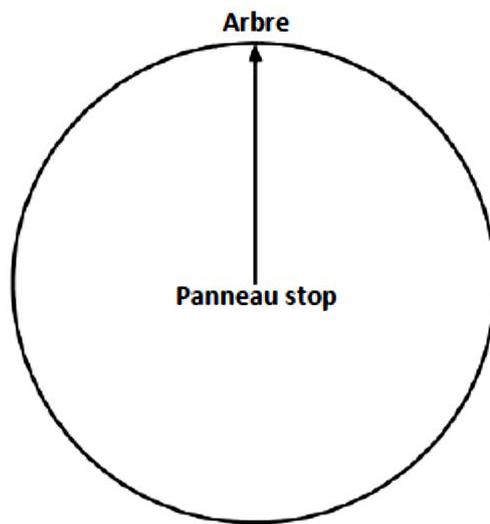


4. Imaginez que vous êtes positionné-e à l'emplacement du **chat** et que vous regardez la **fleur**. Indiquez la direction de la **voiture**.



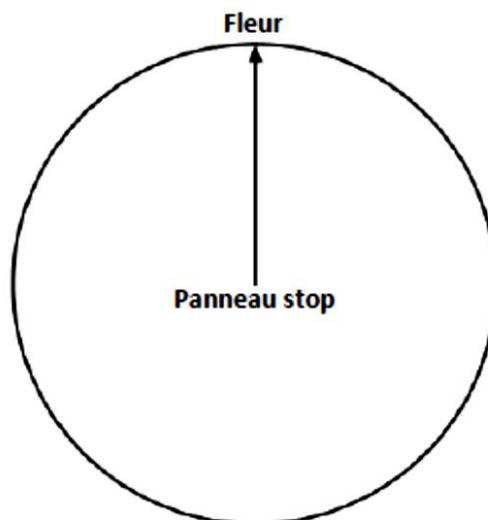


5. Imaginez que vous êtes positionné-e à l'emplacement du **panneau stop** et que vous regardez l'**arbre**. Indiquez la direction des **feux de signalisation**.



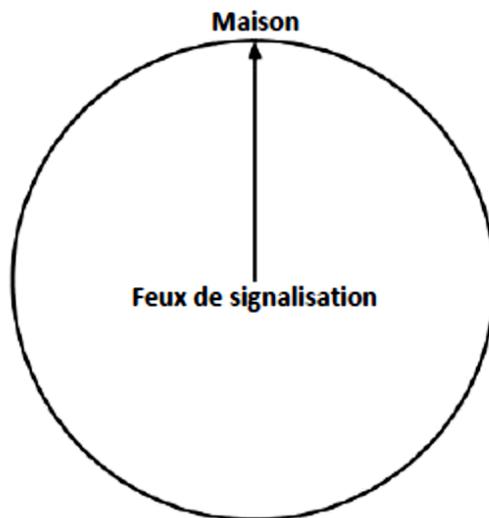


6. Imaginez que vous êtes positionné-e à l'emplacement du **panneau stop** et que vous regardez la **fleur**. Indiquez la direction de la **voiture**.



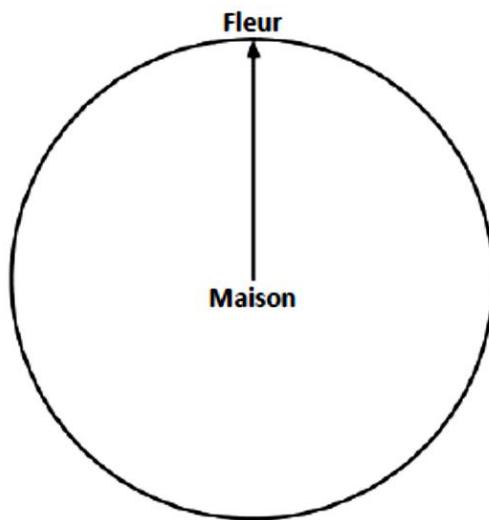


7. Imaginez que vous êtes positionné-e à l'emplacement des **feux de signalisation** et que vous regardez la **maison**. Indiquez la direction de la **fleur**.



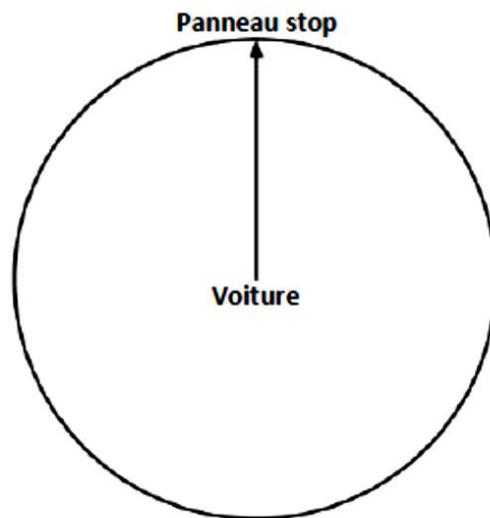


8. Imaginez que vous êtes positionné-e à l'emplacement de la **maison** et que vous regardez la **fleur**. Indiquez la direction du **panneau stop**.



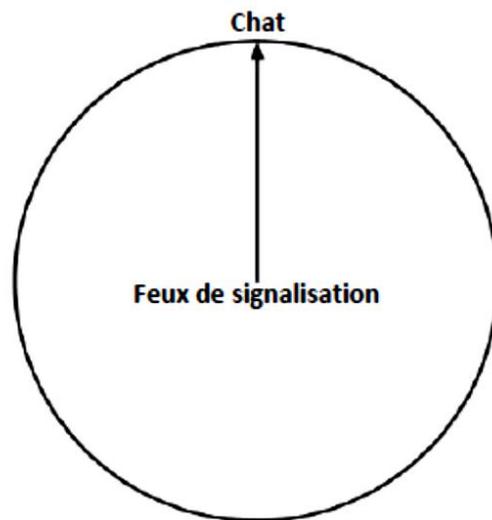


9. Imaginez que vous êtes positionné-e à l'emplacement de la **voiture** et que vous regardez le **panneau stop**. Indiquez la direction de l'**arbre**.



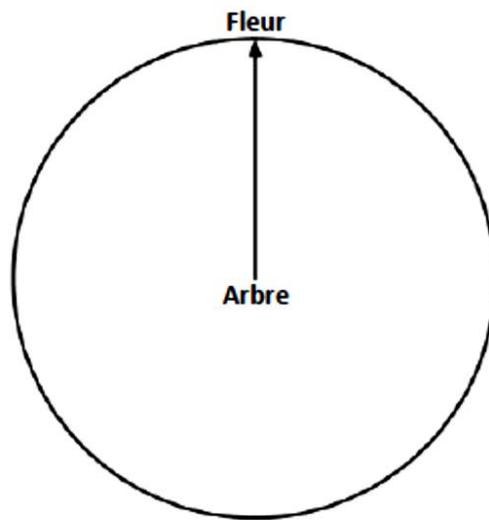


10. Imaginez que vous êtes positionné-e à l'emplacement des **feux de signalisation** et que vous regardez le **chat**. Indiquez la direction de la **voiture**.



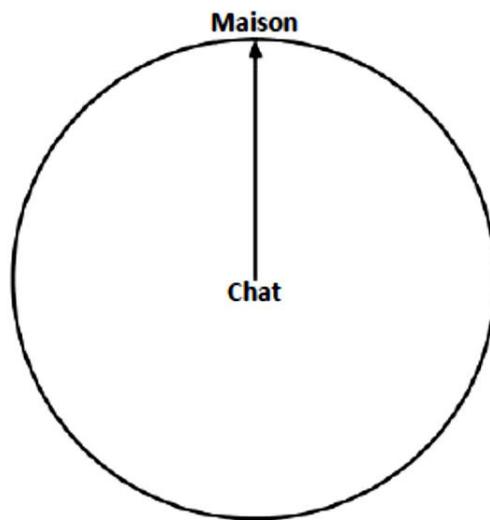


11. Imaginez que vous êtes positionné-e à l'emplacement de l'arbre et que vous regardez la fleur. Indiquez la direction de la maison.





12. Imaginez que vous êtes positionné-e à l'emplacement du chat et que vous regardez la maison. Indiquez la direction des feux de signalisation.



Annexe I.

Séries 1 & 2 du test de Daneman & Carpenter.

Entraînement

- 0.1. Les soirs d'été, nous aimions manger dans le jardin, en compagnie des abeilles.
- 0.2. Pour l'anniversaire de son mari, elle hésitait entre une ceinture et une cravate.

RAPPEL

- 0.3. Le notaire a donné rendez-vous à son client vendredi prochain en fin de matinée.
- 0.4. Ce jeune aventurier est parti en Amérique et il y a vite fait fortune.
- 0.5. Comme toujours dans ce bureau de poste, il n'y avait personne derrière le guichet.

RAPPEL

Phrases tests

Série 1

- 1.1. Elle se leva avec nonchalance et dit à son ami qu'il était un ivrogne.
- 1.2. Elle regagna le balcon de son appartement pour y admirer les toits de la cité.

RAPPEL

- 1.3. D'une main qui tremblait, il se coiffa puis, sans détourner la tête, gagna la sortie.
- 1.4. Il renversa deux cyclistes et termina sa course entre les bras d'un policier.
- 1.5. Sans rien dire, il s'asseyait près du feu et buvait un alcool de poires.

RAPPEL

- 1.6. La jeune fille est partie en abandonnant sa voiture devant le magasin.
- 1.7. Le navire se mit à prendre l'eau car sa coque avait été percée par un rocher.
- 1.8. Dès qu'elle fut arrivée sur la rive du fleuve, la troupe se disposa en cercle.

- 1.9. L'homme regarda discrètement son invité et sut alors qu'il avait atteint son but.

RAPPEL

- 1.10. La prise de vue n'était pas bonne mais cela ne diminuait en rien son éclat.
1.11. Le malade se pencha à l'oreille de son voisin pour lui confier ses angoisses.
1.12. Après avoir dévasté le village, le vieux pirate ordonna de gagner le large.
1.13. Le flacon contenait un liquide bleu clair et frais, au goût de miel et de banane.
1.14. L'incompétence des directeurs est souvent à l'origine de graves problèmes.

RAPPEL

- 1.15. A ce moment, le train entra en gare et mon ami abandonna sa lecture avec regret.
1.16. Le lendemain, il retrouva la clef de sa maison parmi les verres et les bouteilles.
1.17. Ce mauvais peintre reproduisait le même coucher de soleil sur toutes ses toiles.
1.18. Le bateau leur fit sentir quelques légères secousses lorsqu'ils montèrent sur le pont.
1.19. A peine fut-il entré chez le dentiste qu'il ressentit un pénible sentiment.
1.20. Ces sacs de semence étaient précieux car ils représentaient vraiment leur seul avenir.

RAPPEL

Série 2

- 2.1. Au cours des derniers mois, ce terrain sablonneux est devenu le dépotoir de l'île.
2.2. Ce vieil homme a consacré sa vie à l'observation du monde mystérieux des astres.

RAPPEL

- 2.3. Le lendemain, leur maison fut cernée par des individus bâtis comme des hercules.
2.4. Le marin qui manifestait tous les signes de la colère entra dans la cuisine.
2.5. Craignant encore d'être abandonné, le chien suivait son maître comme une ombre.

RAPPEL

- 2.6. Le ciel prit de délicates teintes pâles et le soleil glissa lentement vers l'horizon.
- 2.7. La vendeuse dégagea la table et emballa le vase fragile avec beaucoup de soin.
- 2.8. Elle demanda au serveur qu'il attende encore un peu avant d'apporter le dessert.
- 2.9. Les refus que les pèlerins essayèrent sur la route leur causèrent du chagrin.

RAPPEL

- 2.10. Il donnait l'impression d'avoir abandonné, comme s'il avait renoncé au succès.
- 2.11. Un étranger apparut sur le seuil et tendit à la fille un petit sac de jouets.
- 2.12. Gravement blessé, le bandit tomba sur le trottoir, les mains crispées sur le ventre.
- 2.13. Parmi ces acteurs, il y en avait un dont la prestation était d'une grande qualité.
- 2.14. Il se remit à marcher et cette fois, il était certain qu'il parviendrait au sommet.

RAPPEL

- 2.15. Elle attendit que la brume se lève et put alors admirer le lac et la vallée.
- 2.16. Il resta un moment dans un profond embarras sans pouvoir émettre le moindre son.
- 2.17. Elle rassembla ses maigres affaires et partit sans se retourner loin de ce village.
- 2.18. La chienne rampa jusqu'à la porte, puis elle se mit à gratter avec une patte.
- 2.19. La petite baraque en planches et au toit vitré nous protégeait tous de la tempête.
- 2.20. Il se fit passer pour un inspecteur en civil et entra avec beaucoup de facilité.

RAPPEL

Annexe J.

Contenu des balises informatives de l'expérimentation RVI.

Contenu des balises.

- 1- Présence d'animaux de type libellule, grenouilles, couleuvres.
- 2- Sol humide.
- 3- Essences favorisant le développement de certaines bactéries et champignons.
- 4- Erable de gros diamètre support de micro habitats.
- 5- Sol stable.
- 6- Bois mort permettant l'installation de quelques micro habitats et qui apporte des éléments fertilisants pour le sol.
- 7- Sol stable.
- 8- Présence de micro habitats sur certains arbres.
- 9- Présence de différentes essences d'arbres complémentaires.

Annexe K

Questionnaire RVI.

Questionnaire RV.

Nom Prénom :

Classe :

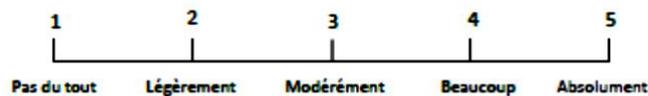
Etablissement :

Date :

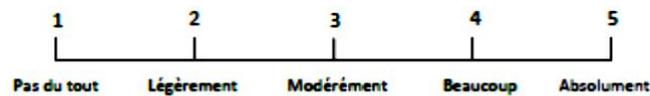
Questionnaire Réalité Virtuelle.

Pour chacune des questions, entoure la réponse qui te correspond le mieux sur cette échelle. Il n'y a pas de bonne ou de mauvaise réponse ☺

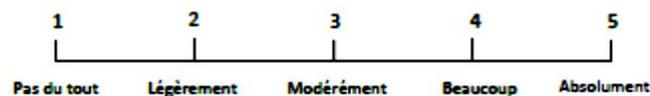
1- L'interface graphique est réaliste et facile à interpréter.



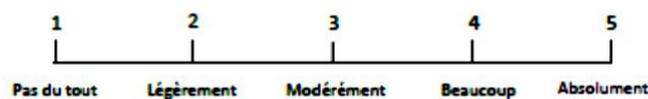
2- L'environnement virtuel pris dans sa globalité m'a permis de me sentir immergé dans cette expérience.



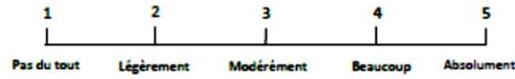
3- L'environnement virtuel paraissait réel.



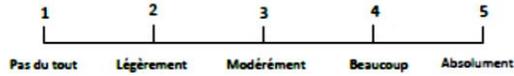
4- Il m'a été facile de reconnaître et d'identifier les éléments de l'environnement en réalité virtuelle.



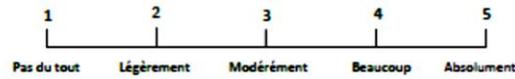
5- J'ai rapidement su me déplacer et prendre en main l'interface.



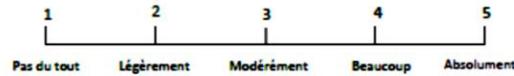
6- Les fonctionnalités (marquer un arbre, accéder à la carte etc.) m'ont été utiles.



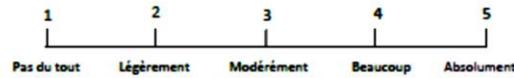
7- J'ai perdu du temps dans mon exploration à cause de la complexité de l'environnement.



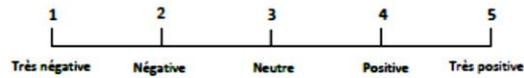
8- J'ai dû faire appel à mes connaissances pour répondre à la problématique de l'expérience.



9- Je me suis senti mal à l'aise pendant cette expérience : casque inconfortable, nausées, perte de repères...



Mon appréciation de cette expérience en réalité virtuelle a été plutôt :



Des remarques sur l'utilisation du programme et de la réalité virtuelle :
Qualités / Défauts éventuels ; Autres :

.....

.....

.....

Annexe L.

Feedback proposé aux élèves après l'exploration en RVI.



Ici nous pouvons observer plusieurs essences : le chêne, le hêtre ou l'érable.

Le sol est stable et on peut rencontrer une diversité de végétaux et arbustes comme par exemple l'aubépine ou l'euphorbe.

On distingue aussi des micro habitats.



Ici nous pouvons observer plusieurs essences : l'érable, Le frêne, Le charme, Le hêtre, Le chêne, L'épicéa.

Le sol est stable et on peut rencontrer une diversité de végétaux et arbustes comme par exemple l'aubépine ou l'euphorbe par exemple.

On distingue aussi des micro habitats.



Ici nous pouvons observer plusieurs essences : l'érable, L'aulne, le frêne et l'orme.

Le sol est humide et on peut rencontrer une diversité de végétaux et arbustes comme par exemple le saule cendre, la fougère, l'iris ou le populaire.

On distingue aussi des micro habitats et la présence d'animaux.

Annexe M.

Consignes et commandes du didacticiel RV.

Phase 1 : didacticiel.

Consigne : Tu viens d'apparaître dans le didacticiel. Tu vas devoir te familiariser avec l'environnement et les manettes.

Les actions à apprendre :

- **Se déplacer/téléporter** : viser le sol en appuyant sur le pad de la manette droite, un cylindre bleu apparaît. Le déplacer où on souhaite sur le sol et relâcher le pad pour que la téléportation ait lieu.
- **Faire apparaître la carte** : appui maintenu de la gâchette de la manette droite.
- **Obtenir une information à propos d'un élément** : viser l'élément et appuyer sur le pad de la manette gauche, un rayon bleu apparaît et laisse apparaître une bulle d'information tout en désignant l'élément visé en légère surbrillance.
- **Marquage d'un élément** : viser l'élément souhaité avec la manette gauche et appuyer sur la gâchette de cette même manette pour marquer. L'élément devient alors rouge. Ce marquage est réversible en réalisant la même action : viser l'élément marqué et appuyer sur la gâchette avec la manette gauche.

Consigne : ouvre la carte, ici tu te trouves dans la zone X, peux-tu aller dans la zone Y ? une fois qu'il est dans la zone demandée : peux-tu marquer la balise ? maintenant l'arbuste ? Et maintenant peux-tu enlever le marquage de la balise ?

- ⇒ Appropriation de la zone de déplacement : quand on arrive à la limite de la zone RV, on voit un quadrillage bleu

Annexe N.

Consignes données à l'apprenant lors de l'expérimentation RV, selon la condition expérimentale.

Consigne : maintenant, on va commencer l'expérience. Tu vas devoir explorer une parcelle de forêt dans laquelle on a délimité trois zones par des barrières. L'une d'entre elle devra accueillir une aire de pique-nique pour des touristes et des randonneurs. On souhaite donc implanter des installations de type tables, chaises, banc et poubelle. Tu vas donc devoir explorer attentivement ces trois zones et me dire, à la fin, dans laquelle il est préférable d'installer cette aire de pique-nique selon toi. Attention, il est important que tu respectes l'écosystème forestier de manière à le préserver le plus possible.

Pour le mode flashing : sois très attentif, des éléments vont clignoter, tu vas devoir les explorer un à un. Clique dessus pour afficher les informations qui lui correspondent. Prends bien le temps de lire les informations qui te sont données. Même si tu vois un autre élément se mettre à clignoter, concentre-toi sur l'élément que tu viens de sélectionner. Si tu juges que cet élément est important alors marque-le et explique-moi, en quoi il a de l'importance selon toi ? sinon, passe à l'élément suivant. Si tu le souhaites, tu peux retourner à un élément pour le marquer ou annuler un marquage réalisé auparavant. Quand tu auras terminé, je t'indiquerai dans quelle zone te rendre. As-tu des questions ? dis-moi quand tu es prêt et nous pourrons commencer.

Pour le mode sans signalisation : Je vais te laisser suffisamment de temps pour explorer cette zone. Tu peux explorer les éléments dans l'ordre que tu souhaites. Lorsque tu juges qu'un élément est important, alors marque-le et dis-moi, en quoi il est important selon toi. Tu peux bien évidemment changer d'avis et supprimer ce marquage en m'indiquant pourquoi tu changes d'opinion. Quand tu auras terminé, je t'indiquerai dans quelle zone te rendre. As-tu des questions ? dis-moi quand tu es prêt et nous pourrons commencer.