



HAL
open science

Étayage des activités de recherche d'information et d'apprentissage en environnement vidéo : apports de la segmentation et de la structuration

Salomé Cojean

► **To cite this version:**

Salomé Cojean. Étayage des activités de recherche d'information et d'apprentissage en environnement vidéo : apports de la segmentation et de la structuration. Psychologie. Université Rennes 2, 2018. Français. NNT : 2018REN20038 . tel-01902650

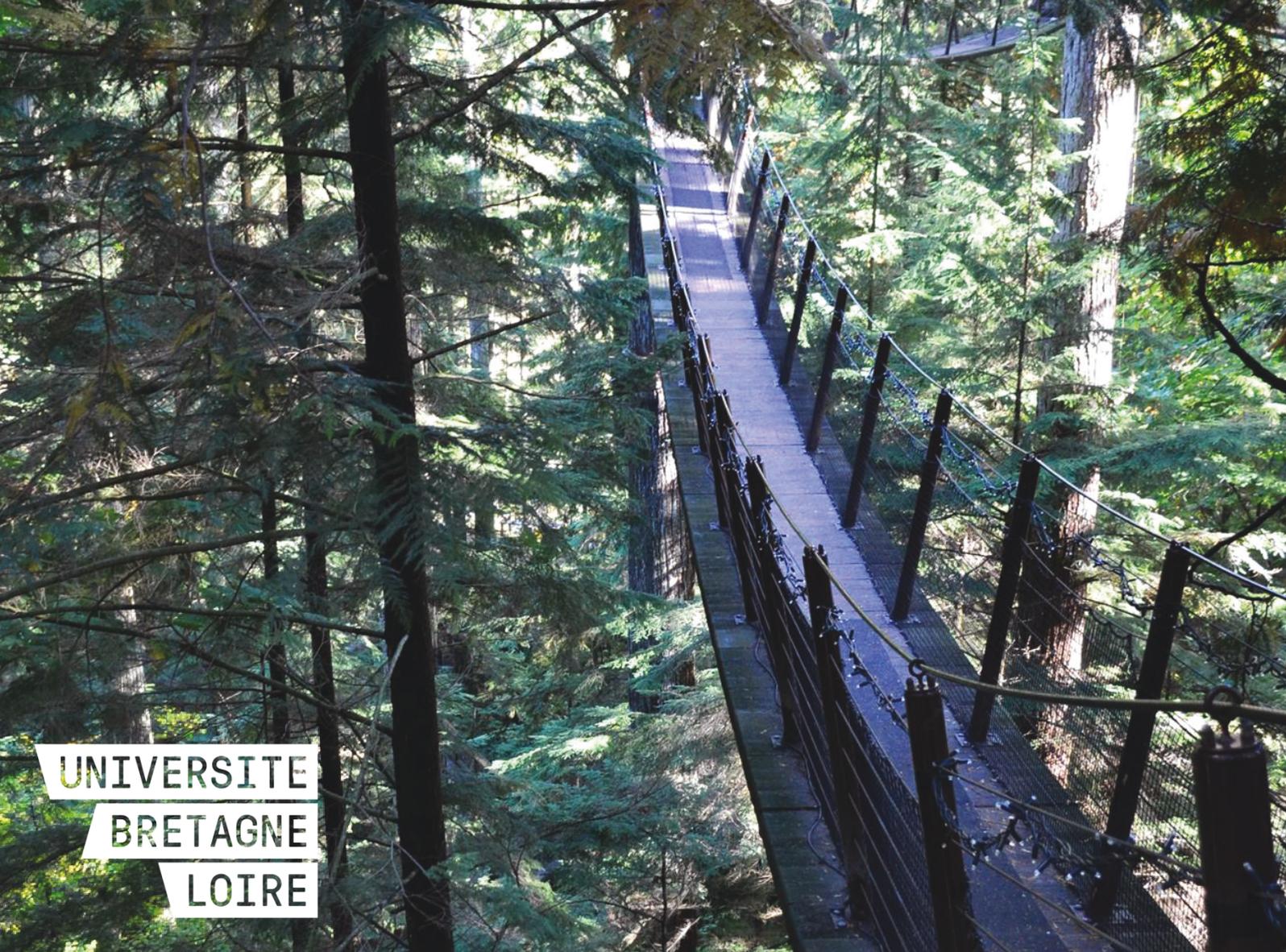
HAL Id: tel-01902650

<https://theses.hal.science/tel-01902650>

Submitted on 23 Oct 2018

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.



UNIVERSITÉ
BRETAGNE
LOIRE

THÈSE / Université Rennes 2

sous le sceau de l'Université européenne de Bretagne

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ RENNES 2

Mention : Psychologie

École doctorale : **Éducation, Langage, Interactions,
Cognition, Clinique (ED 603 ELICC)**

présentée par

Salomé COJEAN

Préparée au sein du LP3C (EA 1285)

Université Rennes 2

Laboratoire de Psychologie : Cognition,
Comportement, Communication

Étayage des activités de recherche d'information et d'apprentissage en environnement vidéo : apports de la segmentation et de la structuration

Thèse soutenue le 4 octobre 2018

devant le jury composé de :

Aline CHEVALIER

Professeure des Universités, Université Toulouse Jean Jaurès / *Examinatrice*

Jean-François ROUET

Directeur de Recherche au CNRS, Université de Poitiers / *Rapporteur*

André TRICOT

Professeur des Universités, Université Toulouse Jean Jaurès / *Rapporteur*

Éric JAMET

Professeur des Universités, Université Rennes 2 / *Directeur de thèse*



UNIVERSITÉ RENNES 2 – HAUTE BRETAGNE

Laboratoire de Psychologie : Cognition, Comportement, Communication
(LP3C)

École Doctorale Education, Langage, Interactions, Cognition, Clinique
(ED 603 ELICC)

Sous le sceau de l'Université Bretagne Loire

**Étayage des activités de recherche d'information et d'apprentissage en
environnement vidéo : apports de la segmentation et de la structuration**

Thèse de Doctorat

Discipline : Psychologie

Présentée par Salomé COJEAN

Directeur de thèse : Éric JAMET

Soutenue le 4 octobre 2018

Jury :

Aline CHEVALIER

Professeure des Universités, Université Toulouse Jean Jaurès (*Examinatrice*)

Jean-François ROUET

Directeur de Recherche au CNRS, Université de Poitiers (*Rapporteur*)

André TRICOT

Professeur des Universités, Université Toulouse Jean Jaurès (*Rapporteur*)

Eric JAMET

Professeur des Universités, Université Rennes 2 (*Directeur de thèse*)

Remerciements

Puisqu'il me semble impossible de décrire ce qu'est « l'expérience de la thèse », impossible de retranscrire toutes les émotions – décuplées – par lesquelles j'ai pu passer, impossible d'expliquer ce que je ressens à la fin de ces trois années, je tiens simplement à remercier l'ensemble des personnes impliquées de près ou de loin dans l'élaboration de cette thèse (promis, les autres phrases seront beaucoup plus courtes que celle-ci).

En premier lieu, je souhaiterais remercier Éric Jamet. Merci de m'avoir accompagnée durant ces nombreuses années (6 depuis la L3 !) sur le chemin sinueux de la recherche. Merci notamment d'avoir écouté mes quelques plaintes sur, au choix : le matériel qui ne marche pas, les participants qui ne viennent pas, les hypothèses qui ne se valident pas... Merci surtout d'avoir étayé (hihi étayage) mon travail et d'avoir contribué à ces années très enrichissantes.

Je tiens également à remercier Aline Chevalier, Jean-François Rouet et André Tricot d'avoir accepté de faire partie du jury de cette thèse. Merci d'avoir pris le temps d'examiner ce travail et de participer à ma soutenance.

Merci à l'équipe LIMAH, et tout particulièrement Pascale Sébillot et Guillaume Gravier pour leur bienveillance tout au long de ces années.

Merci bien sûr aux étudiants qui ont participé aux études menées au cours de cette thèse. Les passations ont été parfois longues mais la bonne humeur des participants a su les rendre finalement très agréables.

Merci aux collègues (ou ex-collègues) du LP3C d'avoir partagé mon quotidien ces trois dernières années. Merci général à l'ensemble des enseignants-chercheurs pour les échanges quotidiens qui ont rythmé et enrichi ce travail de thèse. Mention spéciale au « bureau des ingés » composé de Martin puis Maxime (team LIMAH forever), Mathieu, Jérémie, Julie (Djulaï comme disent les anglais tu sais). Une pensée évidemment particulière pour toute la flopée de doctorants rencontrée durant cette aventure : les « anciens » Anne-Laure (souvenirs éternels !), Anne-Laure (« du bas »),

Benjamin, Emeline, Jean, Natacha, Sylvain, les « contemporains » Anne-Sophie, Anta, Brivael, Gaëlle, Genavee, Marie-Stéphanie, Medhi, Mickael, Nicolas, Pauline, Sarah, Vanina, les « colocs de bureau » Jonathan (la référence !) et Tiphaine, et enfin Julie la « nouvelle ho regardez elle est si mignonne elle est toute petite, j'étais aussi jeune et aussi fringante en début de thèse moi aussi ? », Anthony et Laurine.

Merci tout plein à Gaëtan et Mélissa, qui sont passés au fil de cette thèse du statut de « doctorants en dev » à bien plus que cela. Merci pour ces soirées à parler de recherche mais surtout d'autres choses, merci pour ces moments passés à partager nos bonnes nouvelles et les moins bonnes, merci d'avoir été là dans ces années si précieuses.

Merci à Bleuenn, Mathilde et Cyrille pour leur amitié si importante.

Merci à ma famille pour son soutien inconditionnel. Pensée toute mignonne pour Zoé, née au milieu de cette thèse.

Et enfin, merci ne suffirait pas pour celui qui peut rentrer dans l'ensemble des catégories « collègues », « doctorants », « amis », « famille »... Nicolas, merci pour ces années avec toi, aucun mot ne pourra retranscrire avec justesse le bonheur et les instants vécus. Il me tarde de vivre la suite à tes côtés.

« Le vrai miroir de nos discours est le cours de nos vies ».

Montaigne (Livre I, chapitre 25)

Préambule

Les travaux de cette thèse ont fait l'objet de publications et communications :

Publications dans des revues à comité de lecture :

Cojean, S., & Jamet, E. (en révision). The role of scaffolding in improving information seeking in videos. *Journal of Computer Assisted Learning*.

Biard, N., Cojean, S., & Jamet, E. (2017). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior*, <https://doi.org/10.1016/j.chb.2017.12.002>

Cojean, S., & Jamet, E. (2017). Facilitating information-seeking activity in instructional videos: The combined effects of micro- and macroscaffolding. *Computers in Human Behavior*, 74, 294-302, <https://doi.org/10.1016/j.chb.2017.04.052>

Communications orales et actes de colloque :

Cojean, S. & Jamet, E. (2018). *Effets de la segmentation et de la structuration de vidéos éducatives sur l'apprentissage et la recherche d'information*. Journée des doctorants, 7 juin, LP3C, Université de Rennes 2.

Cojean, S. & Jamet, E. (2017). The impact of scaffolding on characteristics of mental models during information-seeking activity. In J. Dron & S. Mishra (Eds.), *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 292-296). Association for the Advancement of Computing in Education (AACE), 20 octobre, Vancouver, British Columbia, Canada. <https://www.learntechlib.org/p/181199/>

Cojean, S., & Jamet, E. (2016). *Facilitating the information seeking activity in instructional video: the effects of micro- and macro-scaffolding*. Meeting of EARLI SIG 6&7 "Instructional Design" and "Learning and Instruction with Computers", 24 août, Dijon, France.

Communications affichées :

Cojean, S., & Jamet, E. (2017). *Quels outils pour favoriser la recherche d'information dans les vidéos ? Effets sur les performances et les représentations mentales*. Savoir, apprentissage, éducation aujourd'hui : Regards croisés en sciences humaines et sociales - Journées transversales de l'UFR Sciences Humaines de Rennes 2, 23 mai, Rennes, France.

Poster collaboratif (Linkmedia, IRISA, CNRS ; TALN, LINA, CNRS ; CRPCC (ex LP3C), Univ. Rennes 2 ; PREFics, Univ. Rennes 2 ; IODE, Telecom Bretagne ; 2016). *Linking Media in Acceptable Hypergraphs*. Journées CominLabs à L'INRIA, 30 novembre, Rennes, France.

Table des matières

Introduction théorique	15
1. Contexte	15
1.1. Le projet LIMAH.....	15
1.2. Intérêt pour les vidéos pédagogiques	17
1.2.1. L'usage croissant des MOOC.....	18
1.2.2. Une variété de styles de présentation	21
2. L'apprentissage à partir de vidéo	24
2.1. L'apprentissage multimédia	24
2.1.1. Les modèles de Kintsch et Schnotz	25
2.1.2. Le modèle de Mayer.....	26
2.2. Le rôle du modèle mental dans l'apprentissage	28
2.3. Quelques précédentes études sur l'apprentissage en vidéo	31
3. L'activité de recherche d'information (RI) en vidéo.....	35
3.1. Définition de la RI	35
3.2. L'importance de la localisation de l'information en apprentissage	37
3.3. Les principaux modèles théoriques en RI.....	38
3.4. Le rôle d'un modèle mental pertinent en RI	44
3.4.1. Le pouvoir prédictif du modèle mental.....	44
3.4.2. Évaluation de la qualité du modèle mental.....	46
3.4.3. Distinction entre modèle mental et modèle conceptuel	49
4. Comment adapter le format vidéo pour faciliter la RI ?.....	52
4.1. Les bénéfices de l'étayage	52
4.1.1. Un équivalent au modèle conceptuel ?	52
4.1.2. Les niveaux d'étayage.....	53

4.2.	Étayage de micro-niveau : exemples de l'interactivité et de la segmentation	56
4.2.1.	Objectif : éviter le flot continu d'informations	56
4.2.2.	Les activités de micro-niveau	58
4.3.	Étayage de macro-niveau : apport de la structuration	62
4.3.1.	Objectif : favoriser la représentation de structure	62
4.3.2.	Les activités de macro-niveau	65
5.	Effets attendus de l'étayage	70
5.1.	Problématique de la thèse	70
5.2.	Effet attendu de l'étayage sur la RI	70
5.3.	Effet attendu de l'étayage sur l'apprentissage	74
	ÉTUDE 1 : Effet de l'étayage sur la performance en RI dans une vidéo	75
1.	Introduction et hypothèses	75
2.	Méthode	79
2.1.	Participants	79
2.2.	Matériel	79
2.3.	Mesures	82
2.3.1.	Intérêt préalable et compétence perçue	82
2.3.2.	Réussite à la tâche	82
2.3.3.	Temps de réponse	82
2.3.4.	Pertinence du premier clic	83
2.3.5.	Difficulté perçue et contrôle perçu	83
2.3.6.	Rappel des chapitres	83
2.4.	Procédure	84
3.	Résultats	85
3.1.	Variables contrôles (pré-questionnaire)	85

	10
3.2. Tâche de RI	86
3.2.1. Prérequis pour l'analyse statistique.....	86
3.2.2. Réussite à la tâche.....	87
3.2.3. Temps de réponse	89
3.2.4. Pertinence du premier clic.....	91
3.3. Post-questionnaire.....	93
3.3.1. Difficulté perçue.....	93
3.3.2. Contrôle perçu.....	93
3.3.3. Rappel des chapitres.....	94
4. Discussion	94
5. Conclusion.....	97
ÉTUDE 2 : Rôle de l'étayage sur les caractéristiques de la représentation mentale de la vidéo lors d'une activité de RI	99
1. Introduction et hypothèses	99
2. Méthode	102
2.1. Participants.....	102
2.2. Matériel.....	103
2.2.1. Tâche de RI.....	103
2.2.2. Tâche de localisation.....	104
2.3. Mesures.....	105
2.3.1. Intérêt préalable et compétence perçue.....	105
2.3.2. Réussite à la tâche.....	106
2.3.3. Temps de réponse	106
2.3.4. Pertinence du premier clic.....	106
2.3.5. Difficulté perçue.....	107
2.3.6. Précision de la localisation	107

	11
2.4. Procédure	107
3. Résultats	108
3.1. Prérequis pour l'analyse statistique	108
3.2. Tâche de RI	109
3.2.1. Réussite à la tâche	110
3.2.2. Temps de réponse	111
3.2.3. Pertinence du premier clic	112
3.2.4. Difficulté perçue	113
3.3. Tâche de localisation	113
4. Discussion	115
5. Conclusion	117
ÉTUDE 3 : Comment réduire l'effet d'attention divisée pendant la RI au sein d'une vidéo ?	119
1. Introduction et hypothèses	119
2. Méthode	127
2.1. Participants	127
2.2. Matériel	127
2.3. Mesures	130
2.3.1. Intérêt préalable et compétence perçue	130
2.3.2. Réussite à la tâche	131
2.3.3. Temps de réponse	131
2.3.4. Pertinence du premier clic	131
2.3.5. Difficulté perçue	131
2.4. Procédure	132
3. Résultats	133
3.1. Prérequis pour l'analyse statistique	133

	12
3.2. Pré-questionnaire	134
3.3. Tâche de RI	135
3.3.1. Réussite à la tâche	135
3.3.2. Temps de réponse	136
3.3.3. Pertinence du premier clic	138
3.3.4. Difficulté perçue.....	139
4. Discussion	140
5. Conclusion.....	143
ÉTUDE 4 : Effet de l'étayage sur la performance en apprentissage dans une vidéo	145
1. Introduction et hypothèses	145
2. Méthode	150
2.1. Participants	150
2.2. Matériel.....	150
2.3. Mesures.....	152
2.3.1. Intérêt préalable et compétence perçue	152
2.3.2. Temps d'apprentissage	153
2.3.3. Score d'apprentissage.....	153
2.3.4. Activité de révision	154
2.3.5. Difficulté perçue et contrôle perçu	154
2.3.6. Rappel des chapitres.....	154
2.4. Procédure	155
3. Résultats	156
3.1. Variables contrôles (pré-questionnaire)	156
3.2. Tâche d'apprentissage	157
3.2.1. Temps d'apprentissage	157

	13
3.2.2. Score d'apprentissage.....	157
3.2.3. Activité de révision	158
3.3. Post-questionnaire.....	159
3.3.1. Difficulté perçue.....	159
3.3.2. Contrôle perçu.....	160
3.3.3. Rappel chapitres	160
4. Discussion.....	161
5. Conclusion.....	165
ÉTUDE 5 : Effets conjoints de l'étayage et des questions de RI sur l'apprentissage en vidéo.....	166
1. Introduction et hypothèses	166
2. Méthode	170
2.1. Participants.....	170
2.2. Matériel.....	170
2.3. Mesures.....	172
2.3.1. Intérêt préalable et compétence perçue	172
2.3.1. Performance en RI	173
2.3.2. Temps d'apprentissage	173
2.3.3. Score d'apprentissage.....	174
2.3.4. Activité de révision	174
2.3.5. Difficulté perçue.....	174
2.4. Procédure	175
3. Résultats	176
3.1. Variables contrôles (pré-questionnaire).....	176
3.2. Tâche de RI	177
3.2.1. Réussite à la tâche.....	178

	14
3.2.2. Temps de réponse	178
3.2.3. Pertinence du premier clic	180
3.3. Tâche d'apprentissage	180
3.3.1. Temps d'apprentissage	180
3.3.2. Score d'apprentissage	181
3.3.3. Activité de révision	182
3.3.4. Difficulté perçue.....	183
4. Discussion	184
5. Conclusion.....	189
Discussion générale.....	190
1. L'effet de l'étayage vidéo lors d'une activité de RI	191
2. L'effet de l'étayage vidéo lors d'une activité d'apprentissage	195
3. Limites de la thèse	197
3.1. Population étudiée	197
3.2. Matériel.....	199
3.3. Mesures.....	199
4. Perspectives et travaux futurs	203
5. Conclusion.....	205
5.1. Le rôle de l'étayage en vidéo	205
5.2. Implications pédagogiques	206
Références	208
Annexes.....	237
Index des figures	246
Index des tableaux	248

Introduction théorique

1. Contexte

1.1. Le projet LIMAH

Cette thèse a été réalisée dans le cadre du projet LIMAH (*Linking Media in Acceptable Hypergraphs*)¹, financé par le gouvernement français à travers le programme d'investissement d'avenir sous le label « laboratoire d'excellence » (Labex) Cominlabs. Le projet LIMAH s'est étendu sur quatre ans, du 1er mars 2014 au 30 septembre 2018. L'ambition de ce projet était de développer un logiciel facilitant l'exploration d'archives, notamment des articles de presse en ligne, journaux télévisés, émissions de radio, etc. La voix contenue dans les fichiers audio ou vidéo sera automatiquement transcrite sous forme textuelle afin d'avoir une représentation homogène des documents. La particularité de ce logiciel réside dans le fait qu'il propose automatiquement des liens explicites entre les articles qu'il recouvre. Ces liens sont caractérisés de différentes manières, comme :

- Même thématique ;
- Lien de cause/conséquence ;
- Temporalité (avant/après) ;
- « en réaction à »
- « résumé de »
- Etc.

Dans ce logiciel, les opinions émises via les réseaux sociaux sont également analysées et présentées avec les articles auxquels elles se réfèrent.

Deux objectifs sont poursuivis à travers la création d'un tel logiciel. Il s'agit dans un premier temps de proposer un nouvel accès à la navigation dans les données d'informations. Cet outil est ainsi plutôt à destination des professionnels des métiers de l'information et de la communication (e.g., journalistes, attachés de presse). Le second objectif est de faciliter l'accès à des cours en ligne, ou plus largement des

¹ <http://limah.irisa.fr/>

documents à visée pédagogique au sein du logiciel, et la recherche d'information (RI) au sein de ces cours. Il s'agira alors de favoriser la RI et l'apprentissage dans ce nouvel environnement, en proposant une interface facilitant les interactions avec l'utilisateur (e.g., l'élève, l'étudiant, l'apprenant).

Le projet LIMAH s'articule autour de trois axes majeurs : 1) création de liens et structure, 2) analyse du langage et 3) analyses de l'acceptabilité et des usages. Ces trois axes sont développés au sein de trois lots (*work packages*). Le premier lot correspond à la collection de données. Il s'agit de mettre à disposition des données aussi bien orientées média et professionnels de l'information que des données pédagogiques. Au final, les données accessibles par le logiciel sont issues de sites de journaux, de chaînes de télévision, réseaux sociaux ou encore de la plateforme CominLabs. Le deuxième lot a pour objectif la construction d'hypergraphes. Une thèse² est notamment rattachée à cette partie de développement logiciel. Il s'agit de pouvoir au final comparer des contenus entre eux, gérer la granularité des liens proposés entre ces contenus, et analyser les données issues des réseaux sociaux. Le troisième lot concerne l'évaluation et l'usage. Deux thèses, dont celle-ci³, et un post-doctorat y sont rattachés. L'objectif dans ce lot est de prendre en compte les facteurs humains, notamment en faisant des évaluations centrées utilisateur, ou en mesurant l'acceptabilité du projet auprès du public concerné. L'évolution des pratiques du point de vue sociologique et légal est également abordée.

Cette thèse, plus précisément, avait pour objectif d'étudier les effets de la segmentation et de la structuration des vidéos sur les activités de RI et d'apprentissage. Les vidéos sont en effet une source d'information grandissante sur Internet, que ce soit dans le cadre des plateformes de partage de vidéos ou encore les cours en ligne. Elles sont donc de plus en plus susceptibles de se retrouver classées par le logiciel dans un usage futur. De plus, la création de liens entre sources

² « Hypergraphes multimédias dirigés navigables : construction et exploitation », thèse de Rémi Bois en informatique, soutenue le 21 décembre 2017 à l'Université Rennes 1.

³ « Le cadre juridique des médias liés dans l'Union Européenne », thèse de Thomas Pérennou en droit.

d'informations, comme prévue dans le projet LIMAH, implique une segmentation et une structuration automatique du contenu des vidéos traitées. L'objectif de cette thèse est ainsi d'analyser les effets de la segmentation et de la structuration dans des activités qui seront des fonctions majeures du logiciel.

1.2. Intérêt pour les vidéos pédagogiques

Longtemps, les mots ont été le support privilégié de l'enseignement, et ainsi la seule ressource proposée pour l'apprentissage (Lowe & Schnotz, 2014 ; Mayer, 2014a). L'intérêt d'enrichir les documents pédagogiques avec des illustrations, de l'interactivité, des bandes sonores, des animations, ou encore des vidéos, a été questionné au cours du temps et des avancées technologiques. L'avantage de la présence des illustrations sur les performances dans un document pédagogique a été montré dans plusieurs revues de questions (e.g., Carney & Levin, 2002 ; Levie & Lentz, 1982). Ajouter un second vecteur de transmission de l'information, en ajoutant une illustration à un texte écrit, serait donc globalement bénéfique à l'apprentissage. Cet effet est appelé l'effet multimédia (*multimedia effect* ; Eitel, Scheiter, & Schüller, 2013 ; Mayer, 2014a). Dans ce contexte, le terme multimédia renvoie à la présentation d'un message sous plusieurs formats, par exemple un texte et des images (Schnotz, 2014). D'autres recherches se sont portées sur l'apport de l'interactivité à un document pédagogique. L'interactivité peut se définir comme le contrôle donné à l'utilisateur sur les informations qui lui sont transmises, en termes de temps et de contenu (Evans & Gibbons, 2007 ; Mayer & Chandler, 2001 ; Renkl & Atkinson, 2007). Enfin, la possibilité de faire apparaître le contenu illustré ou écrit d'un document de manière séquentielle a également été étudiée : plusieurs travaux ont démontré un avantage de ce type de présentation par rapport à une présentation statique (e.g., Evans & Gibbons, 2007 ; Fleury & Jamet, 2014 ; Jamet, Gavota, & Quaireau, 2008 ; Khacharem, Zoudji, Kalyuga, & Ripoll, 2013).

Dans le cadre des théories sur l'effet multimédia, des chercheurs se sont intéressés à la présentation non pas d'un texte et une/des image(s), mais à l'association des modalités écrites et verbales. Schnotz (2014) considère en effet que le terme multimédia peut être considéré à deux niveaux : selon le format de

présentation ou les modalités sensorielles invoquées. Si l'on considère les formats de présentation, cela correspond aux différentes formes de représentations que peuvent être les textes ou images ; si l'on se place au niveau des modalités sensorielles, le terme multimédia représente les multiples organes sensoriels qui peuvent être sollicités (e.g., yeux ou oreilles). Cet effet de modalité, avec par exemple un texte oral et des graphiques visuels, serait bénéfique à l'apprentissage par rapport à une présentation de l'information uni-sensorielle (Ginns, 2005 ; Low & Sweller, 2014). Cependant, il apparaît que le même matériel ne doit pas être présenté simultanément par les deux modalités mobilisées, cela pouvant conduire à un effet de redondance qui risquerait d'être délétère à l'apprentissage (Kalyuga, Chandler, & Sweller, 2004).

Enfin, des études sur l'apport des animations par rapport aux images statiques ont démontré un bénéfice sur l'apprentissage (e.g., Höffler & Leutner, 2007 ; Mayer & Moreno, 2002). Les animations permettraient en effet de représenter l'information de manière dynamique, par exemple en montrant les changements qui opèrent au cours du temps dans le processus étudié. Néanmoins, l'un des problèmes majeurs de ce type de présentation serait le flot continu d'informations qu'il engendre, et l'impossibilité de s'adapter aux caractéristiques des apprenants (Lowe & Schnotz, 2014).

Dans le domaine éducatif, se développe de plus en plus le format vidéo (Kay, 2012). Il présente l'avantage de mettre en commun les effets de l'interactivité (i.e., grâce à la présence de bouton pause par exemple ; Schwan & Riempp, 2004), de la multi-modalité (i.e., présentation simultanée de l'information visuelle et auditive ; van der Meij & van der Meij, 2014) et des animations (contenu dynamique, présentation séquentielle). L'intérêt pour les vidéos réside donc essentiellement dans le fait que ce format de présentation pourrait cumuler les divers avantages déjà étudiés séparément. De plus, cet intérêt est également fortement lié au contexte actuel, les vidéos devenant un outil pédagogique très utilisé.

1.2.1. L'usage croissant des MOOC

Aujourd'hui, les vidéos sont omniprésentes dans les environnements d'apprentissage en ligne, elles constitueraient ainsi la méthode de présentation la plus commune (Delen, Liew, & Willson, 2014). Le « cours idéal en ligne » d'après Carr-

Chellman et ses collaborateurs (2000) devrait même nécessairement inclure une vidéo, puisque selon ces chercheurs, elle apporterait un avantage certain si elle est utilisée de façon adéquate. En parallèle, l'émergence d'Internet, et son accès de plus en plus simplifié et répandu, a permis de proposer des environnements d'informations de qualité au plus grand nombre (Sharit, Hernández, Czaja, & Pirolli, 2008). Selon le treizième rapport annuel sur l'apprentissage en ligne durant les études supérieures aux États-Unis (Allen & Seaman, 2016), plus de 28% des étudiants américains ont ainsi suivi au moins un cours en ligne. Dans ce contexte, un cours en ligne est défini ainsi quand 80% ou plus du contenu du cours est délivré en ligne.

Le terme MOOC apparaît pour la première fois en 2008 (Hew & Cheung, 2014 ; Joseph & Nath, 2013). Il signifie *Massive Open Online Course* (Hew & Cheung, 2014 ; Ilioudi, Giannakos, & Chorianopoulos, 2013 ; Joseph & Nath, 2013 ; Schacter & Szpunar, 2015 ; Yadav et al., 2015), soit les cours en ligne ouverts et massifs (i.e., disponibles *via* Internet, libres d'accès et pouvant accueillir un grand nombre de participants). Leurs contenus sont essentiellement des vidéos (Yadav et al., 2015). Les MOOC sont rapidement devenus populaires dans le monde de l'éducation (van der Zee, Admiraal, Paas, Saab, & Giesbers, 2017), au point d'être qualifiés comme l'une des plus grandes révolutions en éducation de ces 200 dernières années (Yadav et al., 2015). Koutropoulos et ses collaborateurs (2012) indiquent que généralement, plus de 500 participants s'inscrivent à chaque MOOC. Ce chiffre semble néanmoins à nuancer et reflète certainement l'activité des plateformes les plus importantes (e.g., Coursera⁴, Udacity⁵, EdX⁶, Khan Academy⁷, Fun⁸).

L'une des caractéristiques clés des MOOC est la flexibilité permise par ce format : chaque participant peut décider de son niveau de participation et gérer ainsi lui-même son apprentissage (Joseph & Nath, 2013). Les participants y ont accès à

⁴ <https://www.coursera.org/>

⁵ <https://www.udacity.com/>

⁶ <https://www.edx.org/>

⁷ <https://fr.khanacademy.org/>

⁸ <https://www.fun-mooc.fr/>

tout moment et à tout endroit : les seuls prérequis pour suivre un MOOC sont en effet l'accès Internet et de l'intérêt (Joseph & Nath, 2013 ; Koutropoulos et al., 2012). À ce propos, Hew et Cheung (2014) relèvent quatre principales raisons pour lesquelles les participants s'inscrivent à un MOOC :

- Pour améliorer ses connaissances sur un sujet / un thème donné, pour apprendre ;
- Par curiosité envers les MOOC ;
- Par défi personnel, par exemple pour savoir s'ils peuvent suivre les cours d'une grande université qui dispense des cours sur une plateforme de MOOC ;
- Pour avoir un nombre élevé de certificats (i.e., attestation que l'apprenant a bien suivi le cours en entier et a participé à ce qui était requis).

Zhang et ses collaborateurs (2006) relèvent par ailleurs une liste de bénéfices permis par l'apprentissage en ligne, qui pourrait donc s'appliquer aux MOOC. Parmi ceux qui sont cités, certains s'appliquent directement aux apprenants, par exemple :

- La flexibilité permise par ce genre de format en termes de temps (i.e., horaires) et de lieu ;
- Ce genre de format favoriserait l'apprentissage auto-dirigé et auto-rythmé, c'est-à-dire que l'apprenant peut gérer lui-même son apprentissage en fonction de ses besoins et capacités ;
- Ce format pourrait créer un environnement d'apprentissage collaboratif entre apprenants et experts physiquement éloignés ;
- Un accès illimité au matériel d'apprentissage lorsqu'il est disponible en ligne.

Cependant, malgré la liste de tous ces potentiels bénéfices à l'apprentissage en ligne et aux MOOC, il est à noter le fort taux d'abandon sur ce type de plateformes : 90% des participants qui s'inscrivent à un MOOC ne finiraient pas le cours (Hew & Cheung, 2014 ; Koutropoulos et al., 2012 ; Yadav et al., 2015). Si les participants ne s'inscrivent pas tous pour apprendre le contenu du MOOC (ils peuvent, par exemple, l'avoir fait par curiosité ou revenir sur leur décision peu de temps après), la question peut se poser de savoir comment faciliter l'apprentissage des participants face à une

vidéo. En effet, ce type de format de présentation est particulier et ne peut être comparé à de l'apprentissage sur document papier, statique ou uniquement visuel. Dans une vidéo, des images dynamiques ont lieu en même temps qu'un texte oral est prononcé. L'information dispensée est donc transitoire : elle disparaît rapidement pour être remplacée par une autre (Wong, Leahy, Marcus, & Sweller, 2012). Pour étudier l'apprentissage au sein d'une vidéo, il est nécessaire d'étudier les processus à l'œuvre et de prendre en compte cette contrainte majeure qu'est son aspect transitoire, pour adapter au mieux la présentation de ce type de format sur des plateformes telles que celles proposant des MOOC.

1.2.2. Une variété de styles de présentation

Il existe pour le moment différents styles de présentation des vidéos dans les MOOC. Ces styles varient selon les plateformes, ou même au sein d'une même plateforme sur des cours différents (Reutemann, 2016), ou encore dans une même vidéo (Santos-Espino, Afonso-Suárez, & Guerra-Artal, 2016). Le type de présentation le plus utilisé par les concepteurs de MOOC est celui où le participant voit l'image du professeur parler à côté d'une présentation de schéma (*talking head lecture*) par exemple (Ilioudi et al., 2013) : selon l'étude de Reutemann (2016), 74% des cours sur l'ensemble de quatre plateformes parmi les plus connues étaient présentés sous ce format de vidéo (voir Figure 1). Un autre format de présentation reconnaissable est le style Khan, en référence aux vidéos relativement similaires disponibles sur la plateforme Khan Academy (Ilioudi et al., 2013). Il s'agit généralement d'un dessin sur tableau, accompagné d'une explication orale (voir Figure 1). Le format de présentation avec l'image du professeur intégrée au visuel présenté peut également être cité (voir Figure 1), ou encore simplement la présentation d'un diaporama accompagné de commentaires oraux (Santos-Espino et al., 2016).

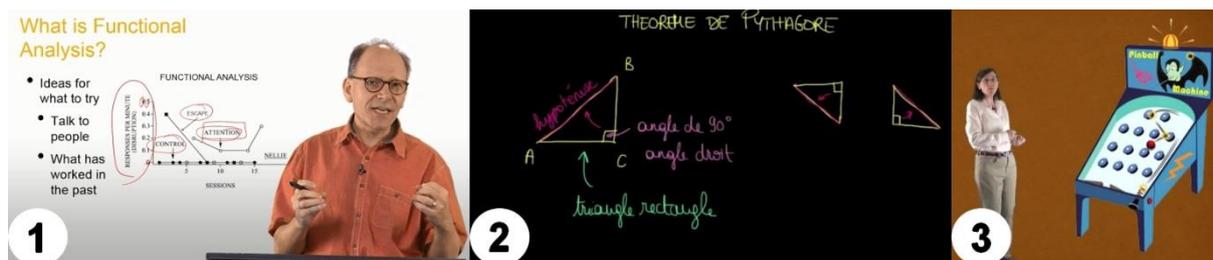


Figure 1. Exemples de style *talking-head* (1) sur la plateforme EdX⁹, de style Khan (2) sur la plateforme Khan Academy¹⁰ et d'image du professeur intégrée à l'image présentée (3) sur la plateforme Coursera¹¹

Peu de recherches ont cependant été menées sur les effets de chaque style de présentation sur l'apprentissage (Ilioudi et al., 2013). Les études réalisées jusqu'à maintenant ont surtout porté sur la comparaison de l'apprentissage en ligne avec un apprentissage en face-à-face (Schacter & Szpunar, 2015). Parmi le peu de travaux menés en ce sens, la recherche menée par Ilioudi et ses collaborateurs (2013) peut être citée, dans laquelle les auteurs comparent un apprentissage avec un livre, une vidéo de cours proposée en classe (en style *talking-head*) et une vidéo avec une voix expliquant par-dessus un dessin (style Khan). Pour cela, trois modules de mathématiques ont été répartis sur trois semaines pour chaque condition expérimentale et l'apprentissage des participants était mesuré. A la fin des trois semaines, la seule différence significative était en la faveur du style *talking-head* par rapport au livre. Chen et Wu (2015) ont quant à eux comparé l'apprentissage selon trois types de présentations de vidéo : la capture vidéo d'un cours avec voix du professeur et image en simultané, un diaporama avec la voix du professeur par-dessus, et enfin une vidéo avec les images du professeur et du diaporama intégrées. Les résultats de cette étude ont montré une supériorité du format « capture de cours »

⁹<https://courses.edx.org/courses/course-v1:ASUx+DOGx003+3T2017/courseware/e55bad90a7a34fd8aa1d8121232d96f9/8f0188c0e0ca48e19769aeb60abcb57c/?child=first>

¹⁰https://fr.khanacademy.org/math/basic-geo/basic-geometry-pythagorean-theorem/geo-pythagorean-theorem/v/the-pythagorean-theorem?ref=staff_video_picks OU <https://www.youtube.com/watch?v=j84pUBA9bjo>

¹¹<https://www.coursera.org/learn/learning-how-to-learn/lecture/75EsZ/introduction-to-the-focused-and-diffuse-modes>

sur le format « images intégrées », lui-même supérieur au format « voix sur diaporama ».

Finalement, il n'existe à l'heure actuelle pas de guide de présentation ou de standard à suivre, à destination des concepteurs de MOOC (C.-M. Chen & Wu, 2015 ; Ilioudi et al., 2013). Un article de van der Meij et van der Meij (2013) permet de recenser quelques lignes directrices de conception de vidéos, mais globalement, face au développement rapide de cette méthode d'éducation, peu de recherches ont actuellement été menées sur la manière d'améliorer l'apprentissage au sein des vidéos (Schacter & Szpunar, 2015). Rowe et Cooke (1995) soulignaient déjà l'un des enjeux centraux de cette nouvelle manière de proposer l'enseignement : lorsque l'on introduit une nouveauté technologique dans un environnement de travail, bien que le but soit d'améliorer la productivité, il peut arriver qu'elle n'apporte que de la complexité. Ainsi, l'objectif de cette thèse sera d'explorer les processus en jeu dans les activités de RI et d'apprentissage dans les vidéos, afin notamment de proposer des outils pour faciliter l'appréhension d'un tel support par de futurs participants à des MOOC. Il s'agira alors d'essayer de proposer quelques suggestions pour la conception des vidéos pédagogiques afin de promouvoir les activités pédagogiques sans amener de complexité de traitement néfaste.

2. L'apprentissage à partir de vidéo

2.1. L'apprentissage multimédia

Le terme d'apprentissage est largement répandu dans la langue française, mais néanmoins régulièrement utilisé selon des définitions différentes. L'objectif d'un apprentissage serait d'accroître les connaissances dans la mémoire à long terme (i.e., lieu de stockage de très grande capacité) de l'apprenant (Hasler, Kersten, & Sweller, 2007). Les travaux développés dans cette thèse reposeront sur la définition proposée par Mayer (2014c) : il s'agit d'apprentissage, et plus particulièrement d'apprentissage *multimédia*, lorsque les individus parviennent à construire des représentations mentales, c'est-à-dire lorsqu'ils se créent de la connaissance, à partir de mots (qu'ils soient oraux ou écrits) et d'images (e.g., illustrations, vidéos). L'apprentissage consiste en la mémorisation de l'information, mais également en la compréhension du document présenté (Jamet, 2008 ; Kintsch, 1998).

Comme précisé précédemment, l'un des objectifs de cette thèse est de proposer des formats de présentation vidéo favorisant l'apprentissage. Ainsi, il est nécessaire de pouvoir évaluer les conséquences sur les performances dans ce type d'environnement. L'apprentissage peut se mesurer par le biais de deux types d'items, chaque type correspondant à un des objectifs recherchés lorsqu'il y a apprentissage : des items de **mémorisation** et de **compréhension** (Bétrancourt, Dillenbourg, & Montarnal, 2003 ; Eitel et al., 2013). La mémorisation se définit par la capacité qu'a l'apprenant à reproduire ou reconnaître le matériel qui lui a été présenté lors de l'apprentissage (Mayer, 2014c). Elle consiste en le stockage des informations entrantes, ainsi que la possibilité qu'a l'apprenant d'avoir accès à ces informations pour les récupérer (Woodall, Davis, & Sahin, 1983). La mémorisation peut se mesurer grâce à des questions de rétention, en demandant par exemple à l'apprenant ce dont il se souvient d'un texte ou d'un cours qu'il a lu (i.e., rappel), ou encore en lui demandant de sélectionner ce qui a été présenté lors de l'apprentissage (i.e., reconnaissance). Dans les questions portant sur la mémorisation, c'est la quantité d'apprentissage qui est mesurée. La compréhension, quant à elle, implique une interaction entre l'information entrante et les connaissances antérieures (Woodall et al., 1983). L'idée est que l'apprenant parvienne à réutiliser le matériel qu'il a appris

dans de nouveaux contextes. La compréhension est mesurée par des items qu'on appelle de transfert, c'est-à-dire que les apprenants doivent résoudre des problèmes ou répondre à des questions dont les réponses n'ont pas été explicitement données dans le matériel. Évaluer l'apprentissage nécessite cependant également de comprendre au mieux les processus à l'œuvre lors de cette activité. Différents modèles théoriques ont été proposés dans la littérature.

2.1.1. Les modèles de Kintsch et Schnotz

La littérature rapporte de nombreux modèles théoriques, et il n'aurait pas été possible d'en établir une liste exhaustive. Certains modèles cependant apparaissent comme des modèles princeps largement utilisés dans les travaux actuels et semblent intéressants à présenter.

Modèle de Kintsch (1998)

Un modèle d'apprentissage a notamment été proposé par Kintsch (1998), le modèle de Construction-Intégration (*C-I model*). Selon l'auteur, l'apprentissage correspondrait au processus de transformation de l'information (écrite ou orale) en représentation mentale. Ce processus se découperait en deux phases. Premièrement, il y aurait une phase de construction, durant laquelle un modèle mental approximatif mais cohérent est construit à partir d'éléments du texte, dits locaux, mais aussi à partir des objectifs de l'apprenant et de ses connaissances préalables. La deuxième phase serait une phase d'intégration, où les constructions locales non-pertinentes sont rejetées, afin de ne garder que les représentations qui s'assemblent en un tout cohérent. La représentation du texte ainsi créée se décomposerait de plus en deux composantes : une « base de texte » (*textbase*) reprenant les éléments tirés directement du texte, et un modèle de situation (*situation model*) correspondant à un niveau de compréhension approfondi du texte. De la même manière que le décrivent Narayanan et Hegarty (1998), Kintsch (1998) postule que la compréhension dans l'apprentissage requiert de la part de l'apprenant qu'il mobilise à la fois ses connaissances préalables, l'information qui lui est présentée, ainsi que ses compétences en raisonnement.

Modèle de Schnotz (2014)

Dans la littérature, d'autres modèles de l'apprentissage ont été proposés par différents auteurs pour s'adapter à l'apprentissage dans un contexte multimédia et non plus simplement à partir d'un texte écrit. Selon le modèle de Schnotz (2014), l'apprenant accède à l'information par deux entrées possibles : une entrée visuelle ou bien une entrée auditive. Dans le modèle ITPC (*Integrative model of Text and Picture Comprehension* – modèle intégratif de la compréhension de textes et d'illustrations) de Schnotz (2014), il existe deux sous-systèmes de traitement de l'information : descriptif et représentatif. Les informations dites verbales, c'est-à-dire le texte qu'il soit écrit ou parlé, sont traitées successivement dans le système descriptif puis le système représentatif, alors que les informations imagées, c'est-à-dire les illustrations visuelles ou les sons, sont traitées par le système représentatif puis le descriptif. L'idée de deux sous-systèmes distincts mais connectés, l'un traitant les informations langagières et l'autre les informations non-verbales, a été préalablement développée par Paivio (1990). Dans le modèle développé par Schnotz (2014), l'information serait traitée à travers des canaux sensoriels de capacité limitée, et l'élaboration d'un modèle mental est centrale dans la qualité de l'apprentissage. À l'instar du modèle présenté par Kintsch (1998), le modèle mental intègre les connaissances antérieures de l'apprenant.

2.1.2. Le modèle de Mayer

Mayer (2014a) précise sa définition de l'apprentissage multimédia lorsqu'il propose la théorie cognitive de l'apprentissage multimédia (*Cognitive Theory of Multimedia Learning* – CTML). Mayer (2014c) distingue trois courants différents dans leur approche sur l'apprentissage : a) l'apprentissage implique d'augmenter ou de diminuer la force d'une connexion entre un stimulus et une réponse, le matériel multimédia est alors vu comme un système délivrant des *feedbacks* ; b) l'apprentissage consiste en la transmission d'informations entre un professeur et un apprenant passif, et le matériel multimédia délivre alors l'information nécessaire ; c) l'apprentissage est une activité de construction de sens, où l'apprenant est engagé activement dans l'activité, et où le matériel multimédia est alors une aide cognitive, servant de guide sur comment traiter l'information présentée. L'auteur défend

d'avantage la troisième position, celle d'un apprenant actif plus performant qu'un apprenant passif, autrement dit une théorie constructiviste de l'apprentissage. Le modèle de Mayer a été choisi comme référence principale pour les travaux de cette thèse notamment pour l'apport des trois principes proposés au sein du CMTL. Premièrement, l'être humain dispose de deux canaux séparés pour traiter les informations auditives et verbales d'une part, visuelles et imagées d'autre part. Deuxièmement, la quantité d'informations pouvant être traitées en même temps au sein de chaque canal est limitée. Les canaux peuvent ainsi se retrouver dans certaines circonstances en surcharge. Ces deux premiers principes sont communs avec le modèle développé par Schnotz (2014). Enfin, troisièmement, l'apprenant s'engage donc, comme précisé précédemment, dans un apprentissage actif. Il y parvient en sélectionnant les informations qu'il juge pertinentes à son apprentissage, en les organisant dans des représentations mentales cohérentes, et en intégrant ces représentations mentales avec d'autres connaissances antérieures (voir Figure 2). La construction d'un modèle mental de la situation ou du système étudié est centrale dans le modèle de Mayer, c'est également une idée récurrente dans l'essentiel des modèles d'apprentissage comme vu précédemment.

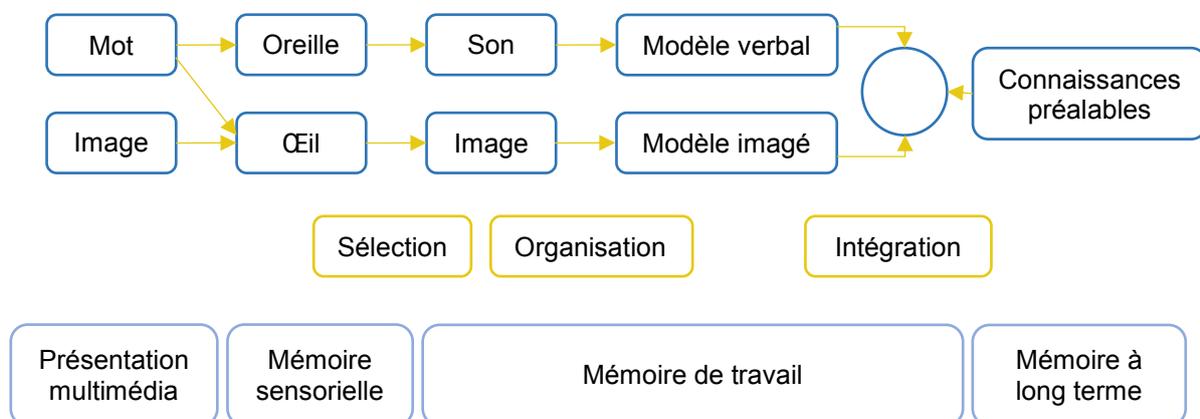


Figure 2. Modèle de la théorie cognitive de l'apprentissage multimédia de Mayer (2014a)

Dans une vidéo particulièrement, l'apprentissage implique donc la création d'une représentation mentale (i.e., d'un modèle mental ; Storey, Fracchia, & Müller, 1999) pertinente (Mayer, Mathias, & Wetzell, 2002) de la vidéo présentée. Cette notion de modèle mental est la base d'un apprentissage de qualité ; c'est pour cela qu'il convient de l'appréhender au mieux pour pouvoir favoriser l'émergence d'une représentation utile à l'apprenant.

2.2. Le rôle du modèle mental dans l'apprentissage

La notion de modèle mental est présente dans nombre de modèles portant sur l'apprentissage multimédia (e.g., Kintsch, 1998 ; Mayer, 2014a ; Narayanan & Hegarty, 1998 ; Schnotz, 2014). Selon Narayanan et Hegarty (1998), lors d'un apprentissage, l'individu se retrouve dans une situation divisée en trois composantes : le monde réel (e.g., le matériel, système ou vidéo à apprendre), la représentation externe (i.e., représentation du monde réel proposée à l'apprenant), et les représentations internes de l'apprenant (i.e., le modèle mental, pour comprendre le monde réel). Il est ainsi généralement admis que les modèles mentaux ont un rôle central dans la représentation qu'ont les individus des objets, des événements, ou plus globalement de la manière dont le monde est fait (Johnson-Laird, 1983). Cependant, il est difficile de trouver une description unique du modèle mental, tant les définitions sont multiples selon les domaines (Doyle & Ford, 1998 ; Rouse & Morris, 1986). Doyle et Ford (1998) proposent une définition qui se veut unifiée du modèle mental d'un système dynamique. Il s'agirait de la représentation interne que la personne a d'un système qui lui est externe (voir également Furlough & Gillan, 2018), représentation relativement persistante et accessible mais limitée. La structure de cette représentation serait ainsi conforme à la structure perçue du système. Les modèles mentaux visent à préserver au mieux la structure du système qu'ils représentent, mais réussissent seulement partiellement puisqu'ils se baseraient sur la perception de l'individu. De plus, les modèles mentaux ne sont pas spécifiques à une modalité sensorielle, c'est-à-dire qu'ils sont davantage abstraits que des images ou que le système représenté. Un modèle mental contient des informations différentes que, par exemple, l'image visuelle correspondante. En effet, il contient moins de détails que l'image parce que les détails jugés non pertinents sont omis dans sa construction. Cependant, le modèle mental pourrait également être considéré comme contenant davantage d'informations que l'image à laquelle il se réfère, parce que sa construction inclut de fait d'y intégrer les connaissances préalables de l'apprenant (Schnotz, 2014 ; Schnotz & Bannert, 2003).

Ainsi, un modèle mental contient les parties ou étapes clés du système présenté, ainsi que les relations entre ces différents éléments (Mayer, 2014a). L'apprenant a la capacité de tester mentalement le modèle mental qu'il possède d'un système, c'est pourquoi les modèles mentaux sont considérés comme ayant une

nature dynamique et changeante (Staggers & Norcio, 1993). Par exemple, dans le cadre de la présentation d'un système tel qu'une serrure, l'apprenant doit pouvoir être capable de se représenter un système de cause à effet dans lequel un changement à un endroit du système engendre en conséquence une modification ailleurs : s'il insère la clé dans un sens plutôt qu'un autre, que va-t-il se passer dans la serrure ? Un modèle mental comprend également trois buts (Borgman, 1986 ; Norman, 1983 ; Rouse & Morris, 1986 ; Storey et al., 1999) : 1) décrire le système qu'il représente ; 2) expliquer comment ce système fonctionne et les états observés au sein de ce système ; et 3) prédire les futurs états du système. L'apprenant doit ainsi pouvoir développer, tester et affiner ses modèles mentaux jusqu'à ce qu'ils soient suffisamment élaborés et fiables pour être efficaces dans de nouveaux contextes de résolution de problème en lien avec le système (D. Zhang et al., 2006). Selon Hegarty (2014), la compréhension qu'a un apprenant d'un système dépend de la qualité des processus mis en œuvre par celui-ci pour traiter le matériel proposé et construire sa propre représentation mentale : perception, attention, encodage, intégration, interaction, etc. Cette définition semble parfaitement adaptée dans le cadre d'un apprentissage procédural : il s'agit alors de comprendre comment fonctionne le système présenté (e.g., une serrure). Dans le cadre d'un apprentissage relatif au contenu d'un document pédagogique (e.g., un cours sur l'histoire de la serrurerie), le système étudié est l'ensemble du contenu délivré, présenté dans la vidéo par exemple. Il paraît nécessaire de préciser les trois buts du modèle mental : 1) décrire le contenu du document présenté ; 2) expliquer le contenu du document et les états observés dans le document ; 3) prédire la réponse à une question sans qu'elle ait été donnée explicitement dans le document (i.e., compréhension).

De plus, d'après Mayer (2014c), la qualité du modèle mental n'aurait une influence que sur les performances en compréhension et non en mémorisation. Pourtant, il semble que le lien entre modèle mental et mémorisation peut s'expliquer en reprenant la définition de la mémorisation : il s'agit du stockage et de la capacité de récupération de l'information (Woodall et al., 1983). Selon le modèle d'apprentissage multimédia présenté par Mayer (2014a), l'information entrante est sélectionnée, organisée et intégrée dans un modèle mental. Pouvoir récupérer l'information du document facilement en mémoire, et donc être performant à un test de mémorisation, reviendrait alors à retrouver une information précédemment organisée. De la qualité

du modèle mental et des étapes ayant précédé sa construction dépendraient donc les performances en compréhension ainsi qu'en mémorisation (voir Tableau 1 pour un résumé).

Tableau 1. Caractéristiques du modèle mental et implications dans les performances en apprentissage

Fonctions du modèle mental	Effets supposés sur l'apprentissage
Le modèle mental se crée à partir de la sélection et l'organisation de l'information entrante. Un modèle mental bien organisé faciliterait la récupération en mémoire des informations stockées.	Un modèle mental pertinent aurait un effet bénéfique sur les performances en mémorisation .
Le modèle mental de l'apprenant lui sert à expliquer le contenu du document présenté et prédire une information relative.	Un modèle mental pertinent aurait un effet bénéfique sur les performances en compréhension .

Il est généralement admis que les novices (i.e., apprenants sans connaissance ni entraînement sur le contenu étudié; Simmons & Lunetta, 1993) ont une représentation mentale des systèmes ou documents présentés moins pertinente que celle des experts dans le domaine (Staggers & Norcio, 1993). Cependant, des études suggèrent que la différence entre experts et novices ne se situe pas dans la quantité de connaissances, mais plutôt dans le fait que leurs modèles mentaux sont fondamentalement différents en fonction du niveau d'expertise (Rouse & Morris, 1986). La structure du modèle mental d'un novice serait ainsi surtout plus désorganisée que celle d'un expert dans le domaine étudié (S. Y. Chen, Fan, & Macredie, 2006). La qualité du modèle mental serait donc l'essence même d'un apprentissage de qualité. Staggers et Norcio (1993) proposent ainsi de penser la conception des documents et supports à l'apprentissage dans le but d'améliorer les capacités et possibilités des apprenants à créer des modèles mentaux appropriés. Il s'agit donc d'aider la construction de la représentation du système présenté pour qu'elle soit au plus près possible de celle attendue, et qu'elle permette un apprentissage efficace et le moins coûteux possible. Cette question, si elle a déjà été beaucoup étudiée au sein des

documents écrits puis multimédias (e.g., Höffler & Leutner, 2007 ; Mayer & Chandler, 2001), émerge ces dernières années au rythme de l'apparition du nouveau format d'apprentissage qu'offre la vidéo. Michas et Berry (2000) proposent l'idée que le caractère dynamique de la présentation de l'information au sein du format vidéo puisse permettre aux apprenants de développer un modèle mental pertinent. Aussi, ces dernières années, les travaux sur les vidéos se sont particulièrement développés.

2.3. Quelques précédentes études sur l'apprentissage en vidéo

Le nombre de recherches scientifiques portées sur l'apprentissage en vidéo augmente chaque année, surtout depuis 2007. Giannakos et ses collaborateurs (2014) recensent ainsi, au sein de huit revues scientifiques spécialisées dans l'éducation en lien avec les technologies, qu'en 2006 seulement 3.6 % des articles sont consacrés à de l'apprentissage en vidéo. En 2012, ce ne sont pas moins de 14.5 % des travaux qui portent sur ce domaine. Ces études proposant de comparer un format de présentation vidéo à un autre plus traditionnel dans un cadre d'apprentissage ont souvent eu comme cadre d'application l'apprentissage procédural. Par exemple, l'étude menée par Arguel et Jamet (2009) était composée de deux expériences, dans le but d'investiguer les bénéfices de la présentation vidéo sur l'apprentissage des gestes de premiers secours. Au sein de la première expérience, trois conditions d'apprentissage étaient comparées, changeantes selon la présentation choisie : 1) des images uniquement, 2) une vidéo seule, ou bien 3) une vidéo associée à des images (statiques). Des questions ouvertes posées aux participants devaient permettre de mesurer l'apprentissage, notamment en s'axant à la fois sur des items de compréhension et de rappel. Les résultats de cette première expérience ont montré que l'apprentissage des participants était meilleur dans la condition proposant à la fois la vidéo et les images que dans la condition montrant uniquement la vidéo. L'apprentissage dans la condition vidéo était lui-même favorisé par rapport à la condition proposant uniquement des images fixes. La deuxième expérience conduite par les auteurs a permis notamment de conclure sur un intérêt à favoriser les images dynamiques (i.e., des animations) par rapport à des images statiques au sein de la vidéo. Cette étude a notamment permis de constater que la présentation d'une vidéo seule peut être accompagnée d'aides telles que des images ou animations. Ces aides

vont pouvoir venir améliorer l'effet positif du format vidéo sur l'apprentissage. Un autre exemple d'apprentissage procédural en vidéo peut être l'étude de Schwan et Riempp (2004), où les auteurs ont comparé différents types de présentation de la vidéo elle-même. Ainsi, les participants étaient confrontés à l'une des deux vidéos suivantes : il s'agissait d'apprendre à nouer un nœud marin dans une vidéo 1) sans possibilité d'action, c'est-à-dire que l'apprenant devait regarder la vidéo en entier à chaque fois qu'il voulait la revoir, ou bien 2) avec possibilité d'interactivité, c'est-à-dire que l'apprenant pouvait faire des pauses, changer la vitesse de défilement, etc. De plus, quatre niveaux de difficulté du nœud à réaliser étaient contrôlés. Les résultats ont montré qu'il fallait davantage de temps aux participants pour apprendre à faire un nœud lorsqu'ils étaient dans la condition proposant la vidéo non-interactive. Cette différence était davantage prononcée pour les nœuds les plus difficiles à réaliser. Cette mesure du temps d'apprentissage a par la suite été décomposée en deux variables distinctes : le temps passé à regarder la vidéo ainsi que le temps dévoué à la réalisation du nœud. Si pour le temps passé sur la vidéo, le seul effet trouvé était celui de la difficulté du nœud, concernant la réalisation du nœud, une interaction a été trouvée. En effet, dans la condition proposant la vidéo non-interactive, les apprenants ont passé au moins deux fois plus de temps à réaliser le nœud que ceux dans la condition avec la vidéo interactive. Le nœud le plus complexe a quant à lui nécessité trois fois plus de temps dans la condition non-interactive que dans la condition interactive. Les auteurs de cette recherche ont également noté un fort usage de l'interactivité par les apprenants lorsque celle-ci était disponible. Cet usage semblait influencé par la difficulté du nœud : plus le nœud était complexe, plus l'interactivité était utilisée. Dans le cadre de cette étude, le format vidéo et les possibilités d'actions permises par ce format ont donc été bénéfiques à l'apprentissage d'une procédure. Les résultats d'une étude menée par Ganier et de Vries (2016) nuancent cet effet. Dans cette étude, les participants devaient apprendre une procédure médicale, la réalisation de points de suture. Trois conditions expérimentales ont été comparées, les apprenants étaient confrontés à 1) des photographies accompagnées de texte écrit, 2) une vidéo accompagnée de texte écrit, ou bien encore 3) une vidéo accompagnée de texte audio. Cinq sutures différentes constituaient l'ensemble de la tâche d'apprentissage. Lors de leur première suture, les participants des deux conditions proposant un format vidéo ont été plus rapides que ceux ayant eu des

photographies. Une interaction entre le format de présentation et le numéro de la suture à réaliser fait apparaître un effet inverse pour les procédures suivantes. Ainsi, le format vidéo serait bénéfique au début de l'apprentissage mais perdrait en efficacité par la suite. Les auteurs concluent à un traitement finalement plus profond de l'information face à des photographies plutôt que des vidéos. Ainsi, ces trois études présentées ont des résultats cohérents : le format vidéo semble favoriser l'apprentissage par rapport à un format plus classique. De plus, il semble nécessaire d'accompagner la présentation d'une vidéo pédagogique, par des animations représentant les étapes clés, ou bien en laissant des possibilités d'interaction avec la vidéo par exemple. L'apport des vidéos sur l'apprentissage est cependant contrasté, le format vidéo pouvant perdre en efficacité par rapport à des photographies au cours de l'activité.

Plus généralement, la supériorité du format vidéo sur les documents écrits est parfois discutée (Merkt, Weigand, Heier, & Schwan, 2011), et le bénéfice sur l'apprentissage n'est pas mis en évidence dans toutes les études. Ainsi, dans une étude menée par Michas et Berry (2000), là encore concernant une tâche d'apprentissage procédural, cinq formats de présentations étaient comparés : 1) texte seul, 2) dessins seuls, 3) vidéo, 4) images tirées de la vidéo uniquement, ou bien 5) texte associé aux dessins. Les résultats sur les performances à l'issue de l'apprentissage ont montré un avantage pour les conditions proposant la vidéo et celle associant texte et dessins par rapport aux autres conditions. Cependant, aucune différence significative n'a été trouvée entre le format vidéo et le format texte et dessins. Par ailleurs, Zacks et Tversky (2003) ont conduit une expérience où trois formats de présentation étaient comparés quant à leurs effets sur un apprentissage procédural : 1) texte seulement, 2) texte associé à des images, ou bien 3) texte associé à une vidéo. Les résultats n'indiquent aucun effet significatif sur la performance selon la condition expérimentale. Et si le score de mémorisation est meilleur dans la condition avec la vidéo, le temps d'apprentissage est plus long. Une seconde expérience menée à la suite de la première comparait une condition avec texte et images à une condition avec texte et vidéo. Aucun effet significatif non plus n'a été constaté sur la performance, ni sur le temps d'apprentissage, et la mémorisation était

cette fois favorisée lors de la présence d'images. Globalement, Merkt et ses collaborateurs (2011) penchent en faveur d'un effet bénéfique de la vidéo lorsqu'il s'agit d'un public apprenant constitué plutôt d'enfants, et pour un matériel peu complexe, alors que ce format sera plutôt au désavantage des publics adultes et des matériels complexes par rapport à un format papier. En résumé, certains travaux n'ont pas permis de mettre en évidence un effet positif du format de présentation vidéo. Une des explications envisagées dans la littérature peut s'interpréter sous le spectre des contraintes induites par le format vidéo lui-même. En effet, ces contraintes peuvent parfois ne pas être prises en compte lors de la conception du matériel d'apprentissage. La contrainte majeure des vidéos est probablement la **nature transitoire de l'information** proposée par ce format, contrainte non présente sur les supports papier. Elle demande que l'information soit continuellement traitée en mémoire de travail¹² par l'apprenant (Hasler et al., 2007 ; Merkt et al., 2011 ; Wong et al., 2012). L'effet bénéfique des images statiques ou dynamiques accompagnant la vidéo dans l'étude de Arguel et Jamet (2009), ainsi que celui des pauses et du temps supplémentaire accordé à l'apprenant dans l'étude de Spanjers et ses collaborateurs (2012), peuvent s'expliquer par l'hypothèse que les images et pauses viennent limiter cet aspect transitoire de l'information. Une autre contrainte à laquelle les apprenants sont souvent confrontés dans les vidéos : **accéder à une partie spécifique** de la vidéo est plus compliqué que rejoindre directement un paragraphe dans un texte (D. Zhang et al., 2006). Connaître l'organisation du contenu, et en particulier la structure du document étudié, est pourtant essentiel à son appréhension par l'individu (Sanchez, Lorch, & Lorch, 2001). En résumé, le format vidéo est prometteur et pourrait permettre de nouvelles possibilités d'apprentissage. Néanmoins, il reste nécessaire de connaître et de prendre en compte les contraintes engendrées par ce format particulier. L'objectif de cette thèse est de concevoir des environnements vidéo adaptés aux apprenants et efficaces, mais cela demande également de considérer plus largement l'activité d'apprentissage.

¹² Structure cognitive permettant le maintien temporaire et la manipulation de l'information en quantité limitée (Sweller, Van Merriënboer, & Paas, 1998).

3. L'activité de recherche d'information (RI) en vidéo

L'apprentissage est une question fondamentale lorsque l'on s'intéresse à l'éducation, et nombre de recherches (e.g., Mayer, 2014c ; Merkt et al., 2011 ; Spanjers et al., 2012) mettent en avant l'importance de cette mesure pour qualifier l'efficacité d'un document pédagogique. Cependant, avant de se focaliser sur l'activité d'apprentissage en elle-même, une autre activité mérite d'être étudiée, notamment parce qu'elle est liée à l'apprentissage : la localisation de l'information.

3.1. Définition de la RI

La littérature scientifique a peu défini, à notre connaissance, l'activité de recherche d'information (RI), bien que de nombreuses études aient été réalisées dans le domaine. Guthrie et Mosenthal (1987, p. 283) définissent ainsi l'activité de RI comme étant « la localisation de l'information dans des documents écrits se rapporte à l'activité d'individus dont le but est de repérer une information précise parmi de nombreuses autres accessibles visuellement¹³ » [notre traduction]. La recherche est intentionnelle : c'est la conséquence d'un besoin très spécifique d'informations pour atteindre un objectif précis (Carmel, Crawford, & Chen, 1992 ; Wilson, 2000). La RI est un processus de création de sens, pendant lequel l'utilisateur se forme un point de vue personnel : il assimile les informations contenues dans différentes sources, ou à différents endroits du document, avec ce qu'il connaît déjà à travers une série de choix quant à sa navigation au sein du document (Kuhlthau, 1991). L'activité de RI est à distinguer de l'activité de navigation (en anglais, *browsing*). La navigation au sein d'un document est un processus interactif, durant lequel l'utilisateur parcourt l'information et peut également élaborer une série de choix (Borgman, Hirsh, Walter, & Gallagher, 1995), notamment afin d'écarter les informations qui lui semblent non pertinentes (Rouet & Tricot, 1998). Les définitions sont relativement proches, si bien que parfois certains auteurs jugent qu'il est difficile de distinguer l'une de l'autre (e.g., Carmel et al., 1992). Ces deux concepts se distinguent parce que la navigation serait

¹³ “*Locating information in written documents refers to the performance of learners who have the goal of detecting a specific subset of information within a relatively wide array of information that is displayed for visual inspection*” (Guthrie & Mosenthal, 1987, p. 283).

peut-être plus vraisemblablement parfois une activité induite par la RI : lorsque l'utilisateur s'est vu définir ou a défini lui-même un objectif de localisation d'information, il peut naviguer au sein du document afin de parvenir à ses fins. La navigation n'est cependant pas exclusivement envisagée dans ce cadre, et peut avoir lieu en dehors de tout objectif de RI.

L'intérêt d'étudier cette activité, afin notamment de pouvoir améliorer sa mise en œuvre dans les documents pédagogiques, est lié au fait qu'elle est impliquée dans le processus d'apprentissage lui-même. L'amélioration des compétences des élèves en RI pourrait ainsi leur permettre de chercher l'information de manière plus efficace, et donc possiblement de mieux apprendre ou comprendre le contenu du document lorsque cette tâche est demandée (Rieh, Collins-Thompson, Hansen, & Lee, 2016). De plus, cette activité de localisation de l'information est courante, voire demandée quotidiennement dans l'enseignement (Guthrie & Mosenthal, 1987). En effet, l'activité de RI revêt une importance grandissante dans le monde éducatif : il est de plus en plus demandé aux élèves de chercher au sein de documents, le plus souvent électroniques, des informations en lien avec le cours (Rouet, 2003). Plus globalement, les individus, dont les enfants dans le cadre scolaire, sont de plus en plus amenés à chercher des informations sur Internet ou dans des environnements numériques (Dinet, 2014 ; Lazonder, Biemans, & Wopereis, 2000 ; Tsai, 2009).

L'implication de cette activité dans une perspective d'apprentissage permet d'envisager sous un nouvel angle les recommandations pédagogiques. Il semble donc essentiel de parvenir à comprendre les processus en jeu lors d'une tâche de RI, afin potentiellement de proposer de nouvelles méthodes pour favoriser l'apprentissage. Plusieurs auteurs ont proposé des modèles théoriques visant à expliquer les processus sous-jacents à l'activité de RI. Ces modèles diffèrent sur de nombreux aspects, mais une base commune à la plupart se dégage : l'importance du modèle mental qui, à l'instar du rôle qu'il joue dans l'apprentissage, semble également être une nécessité au bon fonctionnement d'une tâche de RI (e.g., Kitajima, Blackmon, & Polson, 2000 ; Marchionini, 1997 ; Sharit et al., 2008).

3.2. L'importance de la localisation de l'information en apprentissage

Les études menées et centrées sur l'activité de localisation pourraient éclairer et amener à considérer sous un nouvel angle celles envisagées concernant l'activité d'apprentissage. En effet, la capacité à chercher de l'information semble liée aux compétences en compréhension, si bien que ces deux activités peuvent parfois être difficiles à distinguer l'une de l'autre (Puustinen & Rouet, 2009 ; Rouet, 2003 ; Rouet & Coutelet, 2008). Par conséquent, il semble nécessaire de prendre en compte la performance en RI pour mieux comprendre l'apprentissage. De nombreuses études ont ainsi interrogé le lien entre RI et apprentissage (pour une revue de littérature, voir Rieh et al., 2016).

Cataldo et Cornoldi (1998) ont notamment voulu mesurer l'effet de la RI au sein d'un texte sur la performance à un test de compréhension. Les résultats indiquent que les participants avaient un meilleur score de compréhension lorsqu'ils avaient dû répondre à des questions (i.e., localiser des informations pour donner la réponse) au cours de leur lecture du texte plutôt qu'à la toute fin. D'après les auteurs, ce résultat suggère que les difficultés que rencontrent les apprenants dans la compréhension d'un texte seraient dues à l'incapacité ou la difficulté à chercher dans le texte pour trouver de l'information pertinente. Lorsque les participants sont aidés à le faire, le score de compréhension semble s'améliorer. Une autre étude, menée par Cataldo et Oakhill (2000), a montré que les élèves possédant un bon niveau habituel de compréhension (i.e., des bons *compreneurs*) mettaient moins de temps à rechercher des informations dans un texte que les élèves avec un niveau plus faible en compréhension. Les résultats ont également montré que le temps associé à la RI corrélait significativement et négativement ($r = - .75, p < .001$) avec les questions de compréhension posées à la fin de la lecture du texte. L'idée est ainsi avancée par les auteurs qu'une meilleure compréhension pourrait permettre une RI rapide et pertinente à l'endroit du texte où se trouve la réponse attendue. Ceci serait possible dans ce cas grâce à une meilleure correspondance entre la structure interne du texte en mémoire et la structure externe du texte sur le papier. Les bons *compreneurs* (i.e., experts en compréhension) seraient ainsi de meilleurs « chercheurs » parce qu'ils auraient une meilleure connaissance des caractéristiques structurelles du texte, ou bien parce qu'ils seraient davantage capables de distinguer le plus important du plus anecdotique. La forte corrélation

trouvée dans cette étude pourrait cependant tout aussi bien suggérer qu'une RI efficace permettrait un meilleur score de compréhension par la suite. L'explication de la performance en RI et donc en compréhension pourrait alors être une troisième variable (e.g., qualité de la représentation du document par l'apprenant).

En conclusion, compréhension et RI dans le document sont donc fortement liées. En effet, les apprenants doivent localiser des informations pertinentes avant de lire, comprendre et utiliser l'information (Merkt & Schwan, 2014 ; Rouet & Coutelet, 2008). Cette étape de localisation et sélection de l'information se retrouve par ailleurs décrite au sein du modèle de l'apprentissage de Mayer (2014a), où elle prend place avant les processus d'organisation et d'intégration des informations. Plus la complexité de la tâche augmente, plus la compréhension de l'élève semble liée à l'efficacité de l'activité de localisation (Cataldo & Oakhill, 2000) : savoir ou pouvoir correctement rechercher de l'information au sein d'un document ne serait ainsi plus une activité professionnelle ou personnelle en soi, mais pourrait aussi servir à l'apprentissage (Puustinen & Rouet, 2009). De la même manière, au sein du modèle de Mayer (2014a), cette étape de localisation et de sélection de l'information est également liée aux performances futures en mémorisation. Il semble ainsi intéressant de s'interroger plus précisément sur l'activité de RI. L'étude de Cataldo et Oakhill (2000) pourrait laisser penser que la RI se base, à l'instar de l'apprentissage, sur la qualité du modèle mental de l'apprenant confronté au document (i.e., les experts en compréhension sont meilleurs en RI parce qu'ils auraient un meilleur modèle mental du document).

3.3. Les principaux modèles théoriques en RI

Diverses études ont porté sur la comparaison de modèles définissant l'activité de RI au sein d'un document. Dinet, Chevalier et Tricot (2012) ont distingué, par exemple, les modèles plus généraux des modèles cognitivistes, c'est-à-dire davantage orientés sur les processus mentaux. D'autres modèles ont été analysés dans les études de Lazonder et Rouet (2008) ou encore Weiler (2005). Globalement, les modèles proposés dans la littérature modélisent le processus de RI en le décomposant en étapes ou phases successives, car il s'agirait d'une activité séquentielle complexe (Lazonder & Rouet, 2008). Plusieurs modèles apparaissent dans la littérature comme

étant largement utilisés, et il semble donc important de les citer, mais cette thèse se focalisera finalement sur un modèle qui semble davantage adapté au contexte précis de la RI en vidéo : le modèle de Sharit, Hernández, Czaja et Pirolli (2008).

Modèle de Guthrie et Mosenthal (1987)

L'un des premiers modèles relatifs à la localisation de l'information a été décrit par Guthrie et Mosenthal (1987). Il a été pensé dans le cadre d'une activité de RI dans des documents écrits, et les auteurs proposaient alors cinq composantes au processus. La première étape est la **formulation d'un objectif** : répondre à une question par exemple. Cette étape est cohérente avec la définition de l'activité de RI, dont l'une des caractéristiques est justement d'être conséquente à un objectif précis. L'individu effectuant la RI peut, lors de cette étape de formulation d'un but, procéder en sous-buts successifs. La deuxième étape est l'**inspection des catégories d'informations**, c'est-à-dire les chapitres du document, la table des matières, les paragraphes, etc. Ici encore, un parallèle peut être établi avec l'activité de navigation, que nous avons positionnée comme pouvant faire partie intégrante de l'activité de RI. L'individu, toujours selon ce modèle, prépare ensuite sa **séquence de navigation** dans le document : il définit par exemple des priorités quant aux catégories à inspecter. Les auteurs précisent que cette étape est souvent le reflet de l'efficacité de la RI. Il procède ensuite à l'**extraction de détails** depuis une ou plusieurs catégorie(s), après avoir distingué le plus du moins important. La dernière étape est l'**itération des précédentes étapes** jusqu'à obtention d'une solution satisfaisante pour l'individu.

Modèle de Kuhlthau (1991)

Le modèle proposé par Kuhlthau (1991) présente quant à lui six étapes qui interviendraient successivement dans le processus de RI. Il se distingue des autres modèles notamment par le fait que les domaines affectifs, cognitifs et physiques sont pris en compte par l'auteur. La première étape est appelée **initiation** : l'individu prend conscience qu'il ne possède pas l'information dont il a besoin à cet instant, par exemple parce qu'elle lui est demandée dans un exercice ou pour une nécessité personnelle. La deuxième étape est une étape de **sélection**, durant laquelle l'individu doit identifier et sélectionner le thème à investiguer pour effectuer la recherche. S'ensuivent des étapes d'**exploration** de l'information au sein du document utilisé, de

formulation d'une hypothèse durant laquelle l'individu devient davantage confiant quant à la recherche, et de **collecte**, c'est-à-dire que l'information en lien avec la question de recherche et le thème inspecté est récupérée. Enfin, la dernière étape est la **présentation** ou l'utilisation des résultats de la recherche, en d'autres termes la réponse à l'objectif initial.

Modèle de Kitajima, Blackmon et Polson (2000)

Kitajima, Blackmon et Polson (2000) ont développé le modèle CoLiDeS (*Comprehension-based Linked model of Deliberate Search*), plus particulièrement basé sur les processus cognitifs des utilisateurs lors d'une RI sur Internet. Les auteurs se sont inspirés des travaux de Kintsch (modèle de construction-intégration concernant la compréhension de texte, 1998 ; voir Introduction Théorique 2.1.1.) et ont distingué quatre processus centraux à l'œuvre. Le premier peut se traduire par le terme **analyse** (en anglais, *parsing*) et correspond à une étape de construction de représentations mentales des pages Internet explorées par l'individu. Selon les auteurs, une page Internet peut contenir entre 100 et 200 points d'intérêt (e.g., liens, paragraphes, graphiques), et ces éléments peuvent chacun attirer l'attention de l'utilisateur. Les individus construisent alors une représentation schématique de la page Internet en cours de consultation afin de gérer la complexité inhérente au format Internet. Cette représentation schématique leur permet d'éviter un sentiment de confusion lors de la navigation au sein du document. L'individu se **focalise** ensuite sur un ensemble d'éléments de la page Internet qu'il inspecte. Cette étape guidera les prochaines : la **compréhension** de ces divers éléments et la **sélection** du ou des plus pertinent(s) pour l'objectif. De plus, la sélection de l'action qui sera réalisée par la suite dépend de la compréhension qui aura été faite de l'ensemble d'éléments à l'écran sur lesquels l'utilisateur s'est focalisé. Pour planifier l'action qu'il va pouvoir mettre en œuvre, l'individu doit comprendre et comparer les conséquences probables qui auraient lieu en sélectionnant tel ou tel élément à l'écran.

Modèle de Puustinen et Rouet (2009)

Puustinen et Rouet (2009) synthétisent les études réalisées précédemment dans la littérature sur la RI. Selon les auteurs, l'activité de recherche implique en général un processus cyclique : en effet, la présence d'une étape d'itération ou d'une

boucle relançant le processus jusqu'à obtention d'un résultat satisfaisant se retrouve dans la plupart des modèles. Tout au long de ce processus, les utilisateurs commencent par construire une **représentation** de l'objectif de la recherche. Ils procèdent ensuite à la **sélection** d'une des sources ou catégories d'information présentées, puis à l'**extraction** du contenu qu'ils ont jugé pertinent et qui pourrait convenir pour répondre à l'objectif. Les individus impliqués dans la RI **intègrent** ces nouvelles informations extraites avec celles précédemment acquises puis, s'il le faut, recommencent les étapes précédentes autant de fois que nécessaire pour atteindre l'objectif de la RI.

Cette liste de modèles de RI n'a évidemment pas pour vocation d'être exhaustive, mais elle met en lumière à la fois l'homogénéité et l'hétérogénéité des propositions de modèle en RI. En effet, des étapes se recoupent parfois entre les auteurs, portant le même nom ou bien sous-tendant les mêmes processus (e.g., représentation, sélection, itération), mais d'autres points de divergences (e.g., reformulation, intégration) rendent compliquée la construction d'un modèle unifié. De nombreux autres modèles, que nous ne détaillerons pas ici, soutiennent et complètent ceux présentés (e.g., Brand-Gruwel, Wopereis, & Vermetten, 2005 ; David, Song, Hayes, & Fredin, 2007 ; Eisenberg, Johnson, & Berkowitz, 2010 ; Ingwersen, 1996 ; Marchionini, 1997 ; Rouet & Tricot, 1998 ; Sutcliffe & Ennis, 1998 ; Wopereis, Brand-Gruwel, & Vermetten, 2008). Dans la majorité des modèles présentés, la notion de format est très peu présente, voire absente. L'activité de RI pourrait donc être considérée comme similaire au sein de document au format vidéo aussi bien qu'aux livres ou pages Internet (Wilson, 2000), même si les définitions ont souvent été énoncées en référence à de la RI au sein de documents basés sur du texte.

Modèle de Sharit, Hernández, Czaja et Pirolli (2008)

Un modèle cependant nous semble à la fois le plus simple et le plus approprié pour la RI en vidéo : le modèle de Sharit et ses collaborateurs (Sharit et al., 2008). En effet, les autres modèles ont été élaborés pour la plupart en prenant en compte d'autres supports que la vidéo (e.g., Guthrie & Mosenthal, 1987 ; Kitajima et al., 2000 ; Marchionini, 1997) et, même si le modèle de Sharit est décrit spécifiquement pour de

la RI sur Internet, il nous semble possible de l'appliquer au format vidéo. L'avantage de ce modèle tient notamment du fait qu'il ne comporte que trois étapes, c'est-à-dire moins que la grande majorité des modèles présentés dans la littérature. Les étapes qu'il propose sont cependant des étapes importantes, que l'on peut retrouver chez d'autres auteurs, et qui permettent une modélisation simple et facilement adaptable à de nouvelles situations. Au sein de ce modèle, Sharit et ses collaborateurs considèrent l'activité de RI comme un processus de résolution de problème. Ils décomposent ce processus en trois sous-processus itératifs (voir Figure 3) :

- **Représentation** : lors de cette première étape, l'individu internalise l'énoncé du problème (i.e., l'objectif de la RI). Pour cela, il se crée une représentation mentale (i.e., un modèle mental) des faits qui lui sont présentés.
- **Planification** : l'individu génère alors une méthode pour arriver à une solution, c'est-à-dire qu'il imagine quelles conséquences seront liées à quelles actions, et ce qu'il faut mettre en œuvre pour parvenir à l'objectif fixé. On peut également faire un parallèle avec la définition du modèle mental, et penser que c'est à cette étape que l'individu fait tourner dans sa tête le modèle mental qu'il a précédemment construit. Il teste mentalement les diverses possibilités qui s'offrent à lui afin de déterminer laquelle mettre en action. Au sein d'une vidéo, l'idée est que l'individu va planifier à quel endroit de la barre de la navigation il faudra cliquer pour accéder à l'information recherchée. Cette étape requiert souvent de la part de l'individu qu'il divise le problème en parties ou sous-buts.
- **Exécution** : l'individu met en œuvre les actions qui ont été pensées et développées durant la phase de planification. Ce modèle de l'activité est itératif, si bien que si l'exécution ne permet pas la résolution du problème, l'individu peut modifier sa représentation mentale de la situation et planifier de nouvelles opérations à exécuter.

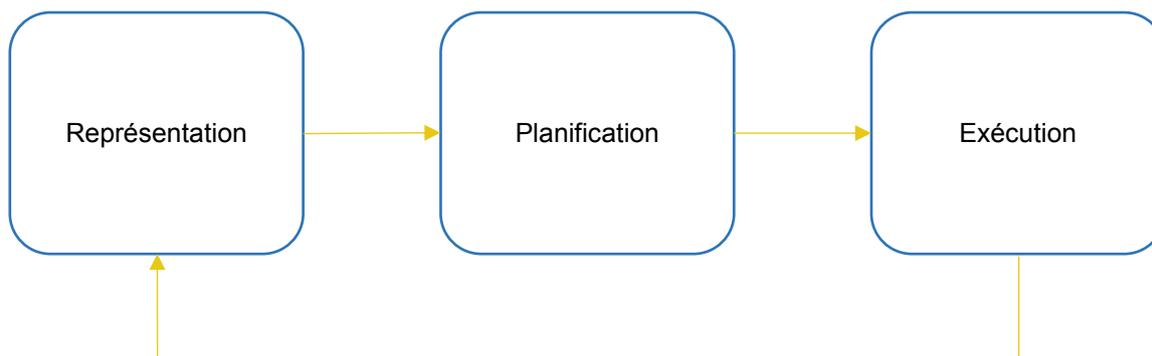


Figure 3. Schématisation du modèle de Sharit et al. (2008)

Le modèle présenté par Sharit et ses collaborateurs accorde donc une grande importance à la qualité de la représentation que l'utilisateur a de la situation : c'est pour ainsi dire la **base** de l'activité de RI. Si le modèle mental n'est pas de qualité suffisante, les étapes suivantes (i.e., planification et exécution) s'en trouveront négativement touchées. A l'inverse, un modèle mental pertinent pourra permettre la bonne réalisation des étapes suivantes, et ainsi le succès de l'activité de RI. Lazonder et Rouet (2008) s'accordent d'ailleurs pour préciser que cette étape de représentation mentale est une étape clé dans la majorité des modèles de RI qu'ils présentent dans leur article. Les différents modèles convergent ainsi tous sur les étapes clés : la première consiste en la construction d'un modèle interne (i.e., un modèle mental) de la tâche. Ensuite, les individus procèdent à un rapide balayage du document dans lequel se déroule l'activité de RI, puis une étude plus profonde du matériel a lieu. Il est important que les individus puissent estimer la pertinence de l'information qu'ils trouvent, en lien avec l'objectif de la RI. Enfin, ils transforment l'information relevée, en l'interprétant ou l'adaptant par exemple, pour l'accorder à ce qui était demandé au début de l'activité de RI.

La construction d'une représentation mentale en tant que base pour la réussite de la RI apparaît ainsi dans de nombreux modèles (e.g., Kitajima et al., 2000 ; Marchionini, 1997 ; Puustinen & Rouet, 2009 ; Sharit et al., 2008). La qualité de cette représentation est donc un élément fondamental pour l'activité de RI, de même qu'elle peut l'être pour une tâche d'apprentissage. L'importance du rôle du modèle mental s'explique au regard de sa définition même, et savoir comment favoriser sa construction par l'individu est un enjeu pédagogique essentiel.

3.4. Le rôle d'un modèle mental pertinent en RI

3.4.1. Le pouvoir prédictif du modèle mental

Selon sa définition (voir Introduction théorique 2.2), un modèle mental est une représentation interne d'un système externe. Dans le cadre d'une RI dans un support vidéo, la représentation mentale nécessaire au succès de l'activité est donc un modèle mental du système au sein duquel se trouve l'information, c'est-à-dire un modèle mental du contenu de la vidéo. Au-delà de la définition présentée précédemment, l'une des caractéristiques du modèle mental est le pouvoir prédictif qui lui est alloué. En effet, si le modèle mental représente la structure et les relations internes du système (ici, la vidéo), il aide également l'individu à prédire le comportement que peut avoir le système dans le futur (Borgman, 1986 ; Furlough & Gillan, 2018 ; Norman, 1983 ; Stagers & Norcio, 1993). Plus précisément, un modèle mental pertinent d'un système donné peut permettre à l'individu de comprendre ce système au point de pouvoir anticiper son comportement futur (Norman, 1983 ; Rowe & Cooke, 1995). Dans le cadre d'une activité au sein d'une vidéo, le modèle mental créé sera une représentation du contenu de la vidéo. Un modèle mental pertinent peut ainsi, par exemple, aider l'individu à naviguer dans la vidéo et savoir quelle information sera délivrée après un clic à tel endroit de la barre de navigation. Comme il a déjà été dit, l'individu a la capacité de le faire tourner mentalement. De cette manière, il peut imaginer ce qui se passerait s'il faisait telle ou telle action au sein du système (Slone, 2002). Il peut ainsi tester, toujours mentalement, les actions possibles, par exemple tenter d'anticiper l'affichage de la vidéo (i.e., le contenu) s'il cliquait à tel endroit plutôt qu'à tel autre dans la barre de navigation, avant de réellement interagir avec le système. L'individu teste grâce à son modèle mental les possibilités et décide finalement quelle action entreprendre sur le système. Ce processus rappelle directement l'étape de planification présentée dans le modèle de Sharit et ses collaborateurs (2008). C'est, pour ainsi dire, l'impact direct de la qualité de la représentation mentale sur l'étape de planification.

Concernant l'effet de la qualité du modèle mental sur l'étape suivante d'exécution, Marchionini (1997, p. 55), qui propose un modèle pouvant se comparer à celui de Sharit et ses collaborateurs (Sharit et al., 2008), précise : « L'exécution des

actions dans la RI est dépendante du modèle mental du système de recherche qu'a l'individu¹⁴ » [notre traduction]. Que la représentation mentale affecte la performance en RI s'explique donc par le fait que, si elle est pertinente, elle va permettre une bonne anticipation du comportement du système (i.e., planification), et donc une exécution des actions adaptée à l'objectif de la RI. Par ailleurs, il existerait un lien direct entre modèle mental et performance en RI. Les individus possédant des modèles mentaux plus développés que les autres, plus complets, seraient plus efficaces dans l'utilisation du système au sein duquel ils travaillent (e.g., Internet, vidéo). Ils feraient en effet significativement moins d'erreurs et trouveraient davantage d'informations lors d'une activité de RI que les individus avec des modèles mentaux moins complets (Staggers & Norcio, 1993 ; X. Zhang & Chignell, 2001).

Afin de pouvoir améliorer les performances des individus en RI, il semble donc important de favoriser la construction d'une représentation mentale pertinente de la vidéo (voir Tableau 2 pour un résumé). Lors des études mises en place pour ce faire, la qualité du modèle mental doit ainsi pouvoir être mesurée. Cependant, la diversité des études de la littérature scientifique nous renseigne sur un point important : il n'existe pas d'évaluation universelle (e.g., Butcher, 2006 ; Chi, Leeuw, Chiu, & LaVancher, 1994 ; Furlough & Gillan, 2018 ; Rowe & Cooke, 1995 ; Slone, 2002) du modèle mental. En fonction de la caractéristique du modèle mentale recherchée, chaque étude doit donc pouvoir définir sa propre mesure.

¹⁴ “*Execution of the physical actions to query an information source is driven by the information seeker's mental model of the search system*” (Marchionini, 1997, p. 55).

Tableau 2. Caractéristiques du modèle mental et implications dans les performances en RI

Fonctions du modèle mental	Effet supposé sur la RI
Le modèle mental permet de prédire les futurs états du document présenté.	Un modèle mental pertinent aurait un effet bénéfique sur les performances en RI.

3.4.2. Évaluation de la qualité du modèle mental

Dans la littérature, les chercheurs s'accordent généralement sur le fait qu'il est très difficile de mesurer les modèles mentaux, notamment parce qu'ils ne sont pas directement observables (Doyle & Ford, 1998 ; He, Erdelez, Wang, & Shyu, 2008 ; X. Zhang & Chignell, 2001). Il est par exemple inutile de demander à un individu de décrire son modèle mental (Norman, 1983), il est nécessaire de passer par une mesure indirecte de celui-ci. La qualité des modèles mentaux est donc généralement inférée depuis le résultat sur une tierce mesure choisie (Marchionini, 1989 ; Rowe & Cooke, 1995 ; Staggers & Norcio, 1993). La mesure choisie dépend du contexte de l'étude, il revient ainsi à l'expérimentateur de trouver la mesure la plus appropriée pour pouvoir rendre compte des différences de performance liées à la qualité du modèle mental (Rowe & Cooke, 1995).

Dans leur étude, Rowe et Cooke (1995) comparent différents types de mesure du modèle mental. Ces auteurs postulent qu'un modèle mental de qualité pour un technicien dans l'aéronautique serait associé à de meilleures performances en résolution de problèmes dans ce champ. Ils mesurent donc le score de chaque participant, ainsi que la qualité de leur modèle mental, évaluée selon les quatre techniques suivantes :

- Un entretien individuel durant lequel le technicien doit résoudre un supposé problème en identifiant les composants impliqués et en décrivant ceux du système plus globalement ;
- Une tâche d'évaluation des liens fonctionnels entre paires de composants selon des échelles ;
- De la création de diagrammes avec représentation des connexions entre composants grâce à des flèches ;

- Une résolution de problème à l'oral avec protocole de verbalisation (i.e., pensée à voix haute).

Dans cette étude, ce sont les techniques d'entretien et d'évaluation des liens entre composants qui sont apparues comme étant les plus prédictives de la performance sur une tâche réelle. La technique de la création de diagrammes n'était plus prédictive une fois les effets des deux premières techniques ôtés. L'évaluation du modèle mental par un protocole de verbalisation n'est pas ressortie dans cette étude comme étant une mesure appropriée.

Dans la littérature, diverses autres mesures des modèles mentaux ont également été proposées. Dans l'étude de Slone (2002), il est demandé aux participants de décrire le système étudié (dans l'étude, Internet). Les descriptions sont alors regroupées dans des catégories selon leur niveau de détail : vague, satisfaisante, technique, brillante, métaphorique. Des études se sont également penchées sur une technique de comparaison sur plusieurs temps de mesure, envisageant ainsi le modèle mental comme un processus mouvant (Staggers & Norcio, 1993). Par exemple, dans l'étude de Butcher (2006), les participants doivent construire un dessin sur ce qu'ils connaissent du système en question (le système cardiaque et circulatoire) avant et après une session d'apprentissage. L'amélioration de la qualité du dessin est alors vue comme une amélioration de la qualité du modèle mental. De la même manière, Chi et ses collaborateurs (1994) comparent les modèles mentaux initiaux et finaux, en les classant selon six catégories, du moins précis au plus précis. Dans plusieurs études menées par Azevedo (e.g., Azevedo, Cromley, & Seibert, 2004 ; Azevedo, Moos, Greene, Winters, & Cromley, 2008) un système de catégorisation est également utilisé. L'individu devait d'abord réaliser un pré-test, durant lequel il écrivait un essai sur ce qu'il savait du système étudié (le système circulatoire). Il naviguait ensuite dans un document afin d'apprendre son contenu, puis un post-test reprenait la consigne du pré-test. Le modèle mental était mesuré en comparant les résultats des deux tests, avant et après apprentissage. Une des trois catégories suivantes était assignée à la qualité du modèle mental : faible, intermédiaire, haute.

L'une des caractéristiques du modèle mental est qu'il permet de **prédire** le comportement futur du système avant même d'interagir avec lui ou sans avoir à le manipuler (Borgman, 1986 ; Furlough & Gillan, 2018 ; Norman, 1983 ; Stagers & Norcio, 1993). Dans cette optique, certains chercheurs ont proposé d'évaluer la qualité du modèle mental en se référant plutôt à la capacité qu'avaient les individus à prédire ce qui pourrait se passer au sein du système après telle action, ou comment parvenir à tel objectif. Ainsi, dans l'étude de Marchionini (1989), les participants devaient rechercher de l'information dans des documents écrits puis électroniques. Avant de commencer l'activité de recherche, ils devaient indiquer comment ils pourraient commencer à chercher l'information (e.g., quelle source privilégier, quels termes utiliser). Ils devaient également dire ce qu'ils pensaient trouver comme résultat à leur recherche s'ils venaient à la mettre en œuvre. De la même manière, les participants de l'étude menée par He et ses collaborateurs (2008) devaient prédire le comportement du système testé (le moteur de recherche du site d'une bibliothèque) ainsi que les résultats qui sortiraient après une RI. Puis ils recevaient une description écrite du fonctionnement du moteur de recherche ou bien s'entraînaient en naviguant librement dessus (selon le groupe expérimental). Il leur était alors re-demandé de prédire les résultats donnés par le moteur de recherche après telle ou telle requête. L'objectif était de comparer ces deux phases pour explorer les effets des conditions (i.e., description écrite ou navigation dans le système) sur le modèle mental des individus. Ces deux études se centrent sur l'activité de RI et présentent une manière similaire de mesurer le modèle mental : prédire le résultat d'une recherche au sein du système. Un modèle mental du système qui sera de qualité va permettre une RI plus performante. Selon certains auteurs (e.g., Marchionini & Shneiderman, 1988 ; Mayer, 1989 ; Stagers & Norcio, 1993), notamment au sein de la littérature dans le domaine de l'apprentissage, le modèle mental pourrait être largement déterminé par un **modèle conceptuel** que fournirait le concepteur à l'utilisateur. Il convient alors de s'intéresser à la manière de proposer un modèle conceptuel qui favoriserait un modèle mental de qualité, permettant ainsi d'obtenir un effet positif sur la RI.

3.4.3. Distinction entre modèle mental et modèle conceptuel

Il est parfois difficile de distinguer les notions de modèle mental et de modèle conceptuel (Staggers & Norcio, 1993), notamment parce que les deux concepts sont fortement liés. Ils ont pourtant des rôles bien distincts, et si le modèle mental d'un individu est inaccessible de manière directe, le concepteur d'un système de RI peut agir sur le modèle conceptuel qu'il choisira de fournir aux utilisateurs.

Un **modèle mental** représente ce que les individus ont réellement « en tête », et va guider leur utilisation des objets ou systèmes avec lesquels ils interagissent (Carmel et al., 1992 ; Norman, 1983). Il s'agit de modèles naturellement évolutifs. C'est à travers l'interaction avec un système par exemple que l'individu va construire un modèle mental de ce système. Les modèles mentaux ne sont pas forcément techniquement précis, mais doivent être fonctionnels (Norman, 1983) pour faciliter l'activité de RI par exemple pour l'individu.

Un **modèle conceptuel** est un outil, dont le but est de fournir une représentation appropriée du système avec lequel l'individu interagit. Cette représentation doit être précise, consistante et complète (Norman, 1983). Un modèle conceptuel est une représentation externe du système (Carmel et al., 1992 ; Seel, 2003), et peut apparaître sous la forme de mots ou de graphiques (Mayer, 1989). Pendant l'apprentissage, il va pouvoir servir d'organisateur de connaissances (Staggers & Norcio, 1993).

Seel (2003) distingue deux manières d'influencer la construction d'un modèle mental de la part des individus confrontés à un document d'apprentissage :

- Laisser les apprenants le construire eux-mêmes dans une optique de découverte du document ou du système étudié ;
- Fournir aux apprenants un modèle conceptuel pour qu'ils élaborent leur modèle mental en fonction de cet outil.

Un modèle conceptuel va en effet **aider l'apprenant à construire un modèle mental du système étudié** : il met en avant les actions majeures du système, par exemple les relations causales qu'entretiennent différents composants (Mayer, 1989). La question qui se pose est de savoir quelle option favoriser pour permettre à l'utilisateur de se construire un modèle mental pertinent : va-t-il construire

spontanément un modèle mental, ou est-il nécessaire de lui fournir un modèle conceptuel sur lequel l'élaboration de son modèle mental va pouvoir se baser (Borgman, 1986) ?

Dans la littérature, les auteurs s'accordent généralement pour dire qu'il est plus simple de donner aux utilisateurs un modèle conceptuel plutôt que les laisser en inférer un (Staggers & Norcio, 1993). Il est en effet plus facile pour un apprenant débutant d'assimiler une explication du système telle que fournie par le modèle conceptuel plutôt que de devoir en induire une individuellement (Seel, 2003). Dans le cas spécifique de la RI, un système de recherche peut se définir comme étant constitué de deux composantes : la base de données contenant l'information délivrée par le système (i.e., dans le cas d'une vidéo, il s'agirait du contenu de la vidéo) et l'interface du système (i.e., le format de présentation de l'environnement vidéo). L'interface permet à l'individu d'avoir accès à l'information du système et peut s'envisager comme étant le modèle conceptuel du système (Marchionini, 1989). Les modifications de l'interface peuvent par exemple concerner les possibilités d'action sur la vidéo (e.g., bouton pause), l'ajout d'une table des matières, ou encore la présentation de la barre de navigation.

Il est du devoir des concepteurs de développer des systèmes de recherche et du matériel pédagogique qui vont pouvoir aider les utilisateurs ou apprenants à construire des modèles mentaux utilisables (Norman, 1983). Proposer un modèle conceptuel pensé en amont va permettre de soutenir l'élaboration des modèles mentaux des utilisateurs. L'objectif est que le modèle mental ainsi construit soit adapté au système étudié et plus approprié qu'un modèle mental construit par l'utilisateur lui-même sans aide extérieure (Norman, 1983). L'intérêt est notamment de favoriser une meilleure performance en RI ou un apprentissage plus efficient, surtout lorsqu'il s'agit de tâches complexes (e.g., Borgman, 1986). Lorsque le système de recherche consiste en un environnement vidéo, proposer un modèle conceptuel revient donc à modifier l'interface avec laquelle les utilisateurs vont pouvoir interagir. Pourtant, comme évoqué précédemment (voir Introduction théorique 1.2.2), il n'existe aujourd'hui aucun guide sur lequel les concepteurs de vidéos pourraient se baser afin de présenter ces documents de manière à favoriser les performances des utilisateurs.

L'objectif de cette thèse est d'évaluer l'influence de différents paramètres dans la présentation des vidéos (i.e., le modèle conceptuel fourni) sur des tâches de RI ou d'apprentissage. L'influence peut se mesurer en termes de performances, mais également de qualité du modèle mental, et l'objectif est ainsi d'accompagner au mieux un individu au sein d'un environnement vidéo selon la tâche qu'il a à accomplir et selon les ressources mises à sa disposition.

4. Comment adapter le format vidéo pour faciliter la RI ?

La conception de vidéos pédagogiques doit s'attacher à respecter les besoins et ressources cognitives des utilisateurs ou apprenants. Une conception éloignée de ces considérations ne saurait être optimisée pour des tâches d'apprentissage ou de RI. L'une des manières de proposer un format de présentation approprié serait d'ajouter de l'étayage à la vidéo pédagogique. L'étayage peut prendre diverses formes mais se distingue particulièrement, dans le cas des vidéos, en deux niveaux complémentaires (i.e., micro- et macro-niveaux). En effet, les environnements vidéo présentent deux types de difficultés spécifiques au format des environnements vidéo (i.e., la fugacité de l'information et la manque de représentation de structure). Ces difficultés seront développées au sein de cette thèse, et l'étayage peut s'envisager de façon à pallier ces deux aspects négatifs. L'objectif de l'étayage étant d'améliorer l'exploration du document (Kim & Hannafin, 2011), un étayage pertinent devrait ainsi pouvoir contribuer à une RI efficiente, mais également à un meilleur apprentissage.

4.1. Les bénéfices de l'étayage

4.1.1. Un équivalent au modèle conceptuel ?

La définition originelle accordée à l'étayage (i.e., *scaffolding*) décrit l'intervention d'un adulte ou d'un pair expert qui vient aider l'apprenant à atteindre un objectif dans une tâche de résolution de problème, en agissant tel un tuteur ou un guide (Azevedo et al., 2008 ; Reiser, 2002, 2004 ; Sharma & Hannafin, 2007 ; Wood, Bruner, & Ross, 1976). La personne venant étayer l'apprenant agit alors comme un soutien, mais ne fait pas la tâche à la place de l'individu. Au sein de la littérature, il semble que la notion d'étayage se soit adaptée aux nouvelles formes de conception d'environnements d'apprentissage. Ce n'est plus le professeur ou les pairs qui viennent aider l'apprenant, mais la technologie (Delen et al., 2014 ; Quintana, Krajcik, & Soloway, 2002 ; Reiser, 2004). Par exemple, l'étayage prendra la forme d'outils qui vont soutenir l'apprenant dans la régulation de sa compréhension d'un hypermédia (Azevedo et al., 2004), ou bien en changeant la tâche de manière à ce que l'individu puisse réussir ce qui aurait été hors de sa portée sinon (Reiser, 2002, 2004). Il s'agit toujours de fournir à

l'apprenant un soutien pour l'aider à appréhender le système ou la tâche, mais les possibilités technologiques ont permis de repenser l'étayage de manière à ce que le document pédagogique lui-même puisse en proposer.

L'objectif final de l'étayage est ainsi d'améliorer l'apprentissage du matériel proposé par l'individu (Azevedo et al., 2004), mais aussi, selon Reiser (2002, 2004), de faciliter la planification. Cette étape étant au centre du modèle de RI proposé par Sharit et ses collaborateurs (2008), un étayage pertinent peut donc agir à la manière d'un modèle conceptuel fourni par le concepteur du document. Il pourra alors servir de base pour la représentation mentale qu'a l'individu du système présenté, et devrait être bénéfique à l'apprentissage et à la planification dans le cadre d'une activité de RI. Selon leur définition (voir Introduction théorique 3.4.3), les modèles conceptuels :

- sont fournis par le concepteur du document pédagogique (Norman, 1983) ;
- vont permettre aux utilisateurs de développer des modèles mentaux utilisables (Borgman, 1986 ; Norman, 1983) ;
- se retrouvent au niveau de l'interface, notamment pour permettre un accès facilité à l'information (Marchionini, 1989).

Les éléments participant à la définition des modèles conceptuels sont ainsi similaires à ceux relatant les particularités de l'étayage. En effet, à la manière d'un modèle conceptuel, les auteurs s'accordent également à dire qu'un étayage pertinent permettra à l'individu de concevoir un modèle mental approprié, notamment en identifiant les éléments importants du système et les relations entre eux (Azevedo et al., 2004 ; Kim & Hannafin, 2011).

4.1.2. Les niveaux d'étayage

Les bénéfices de l'étayage étant intéressants, la question est alors de savoir comment l'étayage peut être pensé et conçu pour améliorer l'interaction entre l'individu et le matériel présenté (Reiser, 2004). Les concepteurs de documents pédagogiques, et plus précisément d'environnements vidéo pédagogiques, ont besoin de guides pour concevoir des étayages adaptés aux tâches qui seront demandés aux utilisateurs (Quintana et al., 2002). La littérature semble présenter une grande variété de techniques d'étayage, qu'il soit fait par des humains ou des supports numériques

(Azevedo & Hadwin, 2005), mais peu d'études se focaliseraient sur l'effet précis des étayages afin de savoir quelles techniques sont efficaces (Azevedo et al., 2004).

La littérature s'entend généralement sur plusieurs types d'étayage qui peuvent être mis en œuvre dans un document pédagogique : procédural, conceptuel, métacognitif, stratégique ou déclaratif (Azevedo & Hadwin, 2005 ; Cagiltay, 2006 ; Kim & Hannafin, 2011). Chaque type d'étayage soutiendrait ainsi le développement de connaissances particulières.

Kim et Hannafin (2011) apportent un éclairage sur ces différents types d'étayage. Ainsi, l'étayage conceptuel aide l'individu à identifier le problème auquel il fait face, l'apprentissage qui lui est demandé, et les ressources cognitives qui lui seront nécessaires pour effectuer la tâche demandée. L'étayage métacognitif vise à développer les connaissances de l'individu sur son propre apprentissage, en d'autres termes l'aider à « savoir ce qu'il sait ». L'étayage stratégique, quant à lui, permettrait à l'individu de considérer différentes approches de résolution de problèmes. Selon Azevedo et Hadwin (2005), l'étayage déclaratif remplace l'étayage stratégique, et consiste en une aide pour l'acquisition de connaissances en elles-mêmes. Enfin, l'objectif de l'étayage procédural est de guider l'utilisateur pour qu'il sache au mieux comment utiliser l'environnement pédagogique à sa disposition, et de lui éviter d'investir des ressources cognitives dans sa navigation au sein de l'environnement (Azevedo & Hadwin, 2005 ; Kim & Hannafin, 2011). Selon Sharma et Hannafin (2007), ce type d'étayage peut ainsi être considéré comme lié à la conception d'interface (*interface design*), en comparaison avec les trois autres types d'étayage qui vont davantage consister en du design cognitif (*cognitive design*). Pour fournir un étayage efficace, il est nécessaire de saisir les problèmes auxquels les utilisateurs peuvent faire face, lors d'apprentissage ou de RI par exemple. L'étayage viendra en effet contraindre et organiser la navigation de l'individu dans l'environnement pour l'aider à manipuler l'information le plus efficacement possible (Reiser, 2004 ; Sharma & Hannafin, 2007). L'étayage peut prendre des formes différentes selon l'objectif souhaité : présentation de schémas à compléter (e.g., Mulder, Bollen, de Jong, & Lazonder, 2016), questions posées avant le visionnage de la vidéo à apprendre (e.g., Carpenter & Toftness, 2017), décomposition d'une tâche complexe en indiquant les actions successives à réaliser (Reiser, 2004), ou encore des sous-objectifs durant

l'apprentissage, par exemple nommer, dessiner ou décrire certains éléments précis du document (e.g., Azevedo et al., 2004).

Le format vidéo est un format aux caractéristiques spécifiques à prendre en compte pour faciliter la navigation de l'utilisateur à travers le matériel. Ainsi, dans le cas des documents vidéo, deux niveaux d'étayage sont rendus possibles – voire nécessaires – car deux contraintes majeures inhérentes à ce format coexistent. En effet, dans le cas des vidéos, l'information délivrée est dite transitoire (Merkt et al., 2011), c'est-à-dire qu'il s'agit d'un flot d'informations coûteux à traiter pour l'utilisateur car le rythme imposé par la présentation de la vidéo ne correspond pas toujours à ses capacités cognitives. De plus, les vidéos n'ont pas de représentation de structure visible, c'est-à-dire que l'organisation de l'information n'est pas transparente pour l'utilisateur comme elle peut l'être dans un document écrit sur papier (van der Meij & van der Meij, 2014).

Cependant, l'avantage majeur des environnements interactifs, comme les vidéos, est qu'ils peuvent laisser la possibilité aux utilisateurs de gérer et manipuler la présentation de l'information qu'ils délivrent (Scheiter & Gerjets, 2007). Ce format s'oppose aux environnements dits contrôlés par le système (Scheiter, 2014). L'interactivité consiste en la possibilité de laisser l'utilisateur décider quelle information apparaît (i.e., « quoi ? ») et de quelle manière elle apparaît (i.e., « comment ? »). En d'autres termes, dans un environnement interactif, l'individu peut contrôler le **rythme** et le **contenu** de la présentation de l'information (Evans & Gibbons, 2007 ; Lawless & Brown, 1997 ; Scheiter, 2014 ; Schwan & Riempp, 2004). Permettre à l'individu de contrôler le rythme, décider de quand apparaît l'information, va permettre de réduire l'effet négatif de la fugacité de l'information. Ces activités se regroupent sous le nom d'activités de micro-niveau (Delen et al., 2014 ; Merkt et al., 2011) : l'étayage qui peut être mis en place à ce niveau s'appelle donc étayage de micro-niveau, ou micro-étayage.

Pouvoir connaître et décider du contenu de l'information délivrée va pouvoir être rendu possible par la présentation visuelle de la structure de la vidéo. Parce qu'il favoriserait des activités de macro-niveau, cet étayage est un étayage de macro-niveau, ou encore appelé macro-étayage (Merkt et al., 2011). L'usage de l'interactivité,

caractéristique propre au format vidéo, peut donc être réfléchi en amont pour diminuer les difficultés également propres à ce format. Une interactivité étayée va ainsi permettre à l'utilisateur d'adapter la présentation du document à ses capacités cognitives (Lawless & Brown, 1997 ; Schwan & Riempp, 2004).

4.2. Étayage de micro-niveau : exemples de l'interactivité et de la segmentation

Le premier niveau d'étayage qui peut être mis en place dans une vidéo est donc un étayage permettant de contrôler le rythme de présentation de la vidéo. Cet étayage est appelé micro-étayage parce qu'il correspond à la possibilité pour l'utilisateur d'effectuer des activités de micro-niveau au sein de la vidéo : mettre sur pause, retour en arrière, avance rapide, mais également se déplacer dans une barre de navigation (*timeline*) segmentée.

4.2.1. Objectif : éviter le flot continu d'informations

Les vidéos, de la même manière que les animations, présentent l'information de manière transitoire et non statique : l'information affichée n'est pas persistante (Höffler & Leutner, 2007 ; Merkt et al., 2011). Cette caractéristique permet de représenter un contenu dynamique, présentant par exemple des changements de forme, de couleur ou de position au cours du temps. Lowe et Schnotz (2014, p. 515) définissent les animations – qu'ils distinguent des vidéos uniquement par leur représentation graphique et non tirées du réel – comme une « présentation imagée dont la structure ou d'autres propriétés change(nt) au cours du temps, et permet ainsi la perception d'un changement continu ¹⁵ » [notre traduction]. Cependant, le traitement en mémoire de travail de cette information délivrée en continu est cognitivement coûteux pour l'utilisateur de la vidéo (Merkt et al., 2011 ; Schwan & Riempp, 2004). Contrairement à l'information disponible à partir de documents écrits statiques, la nature transitoire de l'information issue des vidéos amène des difficultés supplémentaires qui peuvent parfois excéder les capacités cognitives des utilisateurs :

¹⁵ “A constructed pictorial display that changes its structure or other properties over time and so triggers the perception of a continuous change” (Lowe & Schnotz, 2014, p. 515).

il peut y avoir trop d'informations, délivrées trop rapidement pour un traitement optimal (Bétrancourt et al., 2003 ; Doolittle, Bryant, & Chittum, 2015 ; Hasler et al., 2007 ; Lowe, 1999 ; Merkt et al., 2011 ; Spanjers et al., 2012). Dans le cas d'une lecture sur document écrit, où l'information est en permanence disponible, les individus peuvent relire ou sauter certains passages, et surtout ajuster leur rythme de lecture à leurs besoins et capacités cognitifs (Merkt et al., 2011 ; Schwan & Riempp, 2004 ; Singh, Marcus, & Ayres, 2012). En revanche, les vidéos peuvent être parfois présentées sans aucune possibilité de contrôle de la part de l'utilisateur, qui subit ainsi le flot d'informations.

De nombreuses études ont été focalisées sur l'usage des animations dans les documents pédagogiques, avant l'intérêt dans la littérature pour les environnements vidéo. Une méta-analyse, conduite par Höffler et Leutner (2007) sur 76 comparaisons entre animations et images statiques, a permis de conclure à un effet globalement positif des animations sur l'apprentissage. Néanmoins, certaines études ne réussissent pas à reproduire les effets bénéfiques de l'animation : dans les quatre études menées successivement par Mayer et ses collaborateurs (2005), les résultats en termes d'apprentissage étaient meilleurs lors d'un apprentissage sur un support papier (donc statique) par rapport à la présentation d'animations. La difficulté majeure amenée par les animations dans ce cas peut être le flot continu d'informations délivré par le matériel pédagogique, non adapté aux caractéristiques des individus en situation d'apprentissage (Lowe & Schnotz, 2014). En effet, lorsqu'il regarde une animation, dont l'information est transitoire, l'individu doit intégrer les nouvelles informations qui apparaissent à l'écran avec ses connaissances en long terme (voir Introduction théorique 2.1.2) mais également avec les informations présentées juste avant dans l'animation, parce qu'elles ne sont déjà plus disponibles (Hasler et al., 2007). Ces différentes manipulations de l'information ont lieu en mémoire de travail, cependant les travaux précédents ont clairement identifié une limitation dans la capacité de traitement de l'information dans un même temps au sein de cette mémoire (e.g., Mayer, 2014a ; Schnotz, 2014 ; Sweller, Van Merriënboer, & Paas, 1998). La solution qui apparaît est de permettre à l'utilisateur d'un document pédagogique de contrôler le rythme de présentation de l'information. Ainsi, l'idée est qu'il puisse adapter l'apparition de nouvelles informations à ses capacités cognitives.

La possibilité de contrôler le rythme de délivrance de l'information a notamment été testée sous la forme de présentation séquentielle. L'objectif de ce format de présentation est d'afficher les différentes parties d'une présentation dynamique de contenu de manière successive plutôt qu'en simultané (Khacharem et al., 2013). Pour y parvenir, l'individu peut par exemple avoir à disposition un bouton sur lequel il doit cliquer pour faire apparaître la suite de l'animation (e.g., Mayer & Chandler, 2001). La littérature permet de conclure à un effet positif de la présentation séquentielle par rapport à une présentation statique des informations illustrées sur les performances en termes d'apprentissage ou de localisation de l'information (e.g., Bétrancourt et al., 2003 ; Fleury & Jamet, 2014 ; Jamet et al., 2008 ; Khacharem et al., 2013 ; Mayer & Chandler, 2001). Dans une vidéo, tester l'effet des possibilités de contrôle va notamment permettre de donner des pistes aux concepteurs d'environnements vidéo pédagogiques afin qu'ils créent des matériels dont la complexité est maintenue à un niveau raisonnable pour l'utilisateur (Merkt et al., 2011). Contrôler le rythme dans une vidéo permettrait donc aux utilisateurs de l'adapter à leurs capacités cognitives, et ce contrôle serait rendu possible par l'ajout dans l'environnement vidéo d'options telles que des boutons pause, reprendre, avancer ou retour en arrière (Hasler et al., 2007 ; Wouters, Tabbers, & Paas, 2007). Il est également possible de rythmer la vidéo avec des indices visuels et temporels : les segments (Wouters et al., 2007). Les segments ont pour objectif de diviser la barre de navigation (donc la vidéo) en différentes parties correspondant à des contenus différents. Ces possibilités de contrôle du rythme visent à faciliter pour l'utilisateur la navigation dans la vidéo à un niveau appelé **micro-structure**. Dans un texte, les activités de micro-structure ont lieu au niveau des phrases, voire mot à mot, et correspondent à une compréhension locale du document (Kintsch, 1998 ; Kintsch & Yarbrough, 1982). Dans une vidéo, faciliter le traitement de l'information à un niveau local est rendu possible par des activités de **micro-niveau**.

4.2.2. Les activités de micro-niveau

Les activités de micro-niveau regroupent un ensemble d'actions sur la vidéo telles que faire pause, ralentir le rythme, revenir en arrière, etc. (Delen et al., 2014). Plus globalement, elles permettent à l'individu qui les réalise de naviguer à un niveau

très précis de l'information (Merkt et al., 2011). L'ajout d'un bouton pause a été étudié de nombreuses fois dans la littérature et fait directement suite aux travaux sur l'ajout d'une fonction permettant une présentation séquentielle d'un contenu statique. Si l'étude de Schwan et Riempp (2004) montre une supériorité du format interactif (i.e., avec possibilité de faire pause, de contrôler la vitesse, etc.) par rapport à une vidéo non-interactive, l'effet positif apparent des pauses sur l'apprentissage en vidéo est en réalité plus contrasté. Il apparaît en effet, selon Wouters et ses collaborateurs (2007), que les individus, lorsqu'ils sont novices sur le thème abordé dans le matériel d'apprentissage, ne sauraient pas quand ni où réaliser des pauses dans la vidéo. D'autres études ont ainsi été conduites dans l'objectif de pallier cette difficulté, en segmentant cette fois la vidéo. L'idée est de permettre aux utilisateurs de voir la vidéo en segments discrets plutôt qu'en une présentation continue, et les segments sont représentés par une pause automatique entre les parties distinctes de la vidéo (e.g., Doolittle et al., 2015 ; Hasler et al., 2007). Selon le principe de segmentation énoncé par Mayer et Pilegard (2014), les individus devraient apprendre plus efficacement lorsque le support multimédia est présenté en segments dont le rythme est contrôlé par l'apprenant plutôt qu'en une présentation continue.

Pauses et segmentation sont néanmoins souvent peu distinguées, la segmentation étant régulièrement associée à des pauses imposées à la fin d'une partie de la vidéo (e.g., Biard, Cojean, & Jamet, 2017 ; Hasler et al., 2007 ; Singh et al., 2012). Par exemple, dans l'étude menée par Hasler et ses collaborateurs (2007), quatre conditions expérimentales ont été comparées, correspondant à quatre présentations différentes d'une animation :

- Une condition sans animation, avec uniquement de l'information audio délivrée sans contrôle possible de la part du participant ;
- Une version où l'animation se déroule de manière continue, sans contrôle possible non plus ;
- Une version de l'animation segmentée en onze parties, faisant pause automatiquement à la fin de chaque partie et nécessitant que le participant clique pour reprendre l'animation ;

- Une condition dans laquelle l'animation était proposée avec des boutons pause et lecture que le participant pouvait utiliser à son gré.

Les résultats sur les performances à un test d'apprentissage ont montré un effet du contrôle de la part de l'apprenant : ils se sont avérés meilleurs dans les deux conditions leur laissant contrôler le rythme que dans les deux conditions avec une présentation continue du matériel. Pourtant, les participants dans la condition proposant les boutons pause et lecture ont peu utilisé ces moyens d'action sur la vidéo. De la même manière, Biard, Cojean et Jamet (2017) ont comparé les effets d'une vidéo non interactive (i.e., le visionnage se faisait alors forcément en entier sans contrôle possible), d'une vidéo interactive (i.e., possibilité pour l'utilisateur de faire des pauses quand il le souhaitait) et d'une vidéo segmentée (i.e., pauses imposées lors de segments prédéfinis au sein de la vidéo). La tâche consistait en un apprentissage procédural. De même que pour les études précédentes, les participants ont peu utilisé l'option pause quand ils en avaient la possibilité. Finalement, les résultats de cette étude ont montré une supériorité du format segmenté sur les deux autres formats.

En revanche, d'autres études ont analysé plus finement la distinction des effets des pauses et de la segmentation. Spanjers et ses collaborateurs (2012) ont conduit une étude dans laquelle ils ont comparé quatre conditions expérimentales mettant chacune en jeu une animation : avec ou sans pause(s), et avec ou sans indice(s) de temporalité. La première condition consistait en une condition contrôle, dans laquelle l'animation n'était pas segmentée. Dans la deuxième condition, les participants visionnaient une animation segmentée par des pauses de deux secondes. Dans la troisième condition, l'animation proposée était segmentée par des écrans noirs durant chacun une demi-seconde à la fin d'un segment. La quatrième condition présentait l'animation segmentée par des pauses et des écrans noirs, durant deux secondes à chaque fois. Aucune autre possibilité de contrôle sur l'animation n'était donnée au participant. L'objectif de cette étude était de distinguer ce qui avait un rôle dans l'effet bénéfique de la segmentation : est-ce grâce aux pauses imposées dans les études mesurant cet effet, ou grâce aux indices temporels induits qui divisent la vidéo en unités de sens ? Les résultats ont indiqué que les deux explications étaient complémentaires et jouaient un rôle dans les bénéfices observés de la segmentation dans une vidéo. Cependant, elles sont également difficiles à distinguer, dans le sens

où les pauses peuvent également fournir des indices de temporalité par elles-mêmes. Il en revient que les pauses permettraient donc de proposer à l'apprenant du temps supplémentaire pour la mise en œuvre des processus cognitifs inhérents à l'apprentissage. Elles permettraient en outre le maintien en mémoire de travail de l'information de chaque segment quand il se termine, sans avoir à recevoir de nouvelles informations avant d'avoir fini son traitement. Un effort mental moindre a enfin été observé en présence des indices de temporalité. La segmentation rendrait ainsi les événements plus saillants au sein de la vidéo, et elle permettrait à l'apprenant de réduire la recherche de ces marqueurs temporels et leur charge cognitive. Suite à ces résultats, deux études menées par Merkt et ses collaborateurs (2018) ont permis de comparer une explication de réduction de l'effet de la fugacité de l'information à une explication de structuration du contenu par les pauses. Quatre conditions expérimentales ont été construites : 1) une vidéo continue sans possibilité de faire pause, 2) une vidéo avec pauses imposées à des moments prédéfinis et cohérents avec la structure du document, 3) une vidéo avec des marqueurs à des moments prédéfinis (i.e., aux mêmes endroits que les pauses dans la condition 2) mais sans pause, et enfin 4) une vidéo avec pauses à des moments ne correspondant pas à la structure du document. Les résultats n'ont pas permis de conclure à un avantage de l'une ou l'autre des explications.

De manière plus générale, la segmentation aiderait l'utilisateur à naviguer à des endroits spécifiques, ou des moments dans la vidéo, qui lui semblent pertinents (Wouters et al., 2007). D'autres études concluent sur un effet bénéfique de la segmentation sur le rappel lorsque l'apprenant réalise lui-même la segmentation d'une tâche procédurale (Flores, Bailey, Eisenberg, & Zacks, 2017), ou encore que plus la vidéo est segmentée – en nombre de parties – plus l'apprentissage sera bon (Doolittle et al., 2015). La segmentation est souvent associée à des pauses imposées, qui sont utilisées pour indiquer à quel endroit se situent les différents segments dans le document (e.g., Doolittle et al., 2015 ; Mayer & Pilegard, 2014 ; Spanjers et al., 2012). La littérature conclut pourtant à des rôles distincts dans le contrôle du rythme par l'utilisateur (e.g., Spanjers et al., 2012). De plus, il semble rare aujourd'hui de trouver des vidéos sans bouton permettant de faire pause puis de reprendre la lecture. La

segmentation n'est quant à elle pas très répandue. Pourtant, l'hypothèse peut être faite que, si le simple fait de signifier par des indices visuels où se trouvent les segments peut suffire à aider la localisation de l'information, ajouter des pauses imposées à chaque segment pourrait s'avérer inutile dans le cadre d'une activité de RI. En effet, les utilisateurs auraient davantage besoin d'un affichage permanent de la segmentation plutôt que des pauses dans une vidéo dont le but n'est pas de la regarder en entier. Une segmentation affichée sur la barre de navigation de la vidéo, par exemple, devrait alors être bénéfique à la navigation au sein de la micro-structure. Il s'agirait alors d'étayage de micro-niveau, ou **micro-étayage**. La possibilité laissée à l'individu de faire des pauses devrait quant à elle être instaurée de base dans une vidéo, ses effets ayant été éprouvés dans de nombreuses études (même si elle semble paradoxalement peu utilisée).

4.3. Étayage de macro-niveau : apport de la structuration

Le second niveau d'étayage, macro-étayage ou étayage de macro-niveau, dans une vidéo a pour but de laisser l'apprenant maîtriser le contenu de l'information délivrée. Pour cela, la structure sémantique de la vidéo peut être présentée à l'individu afin qu'il n'ait pas à la retrouver de lui-même. Ce niveau d'étayage a pour objectif de faciliter la navigation dans la vidéo à un niveau global. Grâce à cet étayage, l'utilisateur de la vidéo est aidé dans la réalisation d'activités de macro-niveau. Le macro-étayage correspond généralement à la présentation d'une table des matières relative au contenu présenté dans le document.

4.3.1. Objectif : favoriser la représentation de structure

Comprendre la structure d'un texte, c'est-à-dire repérer les chapitres ou les thèmes abordés au cours du document par exemple, est une composante essentielle dans le traitement de texte mais coûteuse pour le lecteur (Sanchez et al., 2001). Afin de limiter cette difficulté rencontrée à la lecture d'un texte ou pendant le visionnage d'une vidéo, Sanchez et ses collaborateurs (2001) ont testé l'effet de la signalisation de l'organisation du document. Il s'est avéré que faire apparaître distinctement la structure d'un texte favoriserait le rappel par les lecteurs des thèmes abordés (i.e., les chapitres du texte), ainsi que la manière dont ils étaient organisés entre eux. Ces

auteurs ont par ailleurs mis à l'épreuve deux hypothèses quant à l'avantage procuré par les signaux organisationnels sur la mémorisation :

- Hypothèse de sélection : lorsque les signaux organisationnels sont absents, le lecteur utilise une stratégie de lecture linéaire, c'est-à-dire qu'il encode les éléments du texte comme une liste d'idées sans structure. Avec les signaux organisationnels, les lecteurs favorisent une stratégie de structure, l'encodage du texte est ainsi organisé hiérarchiquement, par sujets et sous-sujets, etc.
- Hypothèse d'implémentation : les lecteurs matures (i.e., adultes) utilisent une stratégie de structure dans tous les cas, que les signaux soient présents ou absents du texte. Ils parviennent d'eux-mêmes à être conscients de l'importance de la structure, représentant les différents sujets traités. La présence de signaux organisationnels facilite simplement la mise en place de cette stratégie de structure en rendant l'organisation du texte plus saillante.

Les auteurs concluent, avec l'aide des résultats issus d'une étude précédente et d'une nouvelle étude, en faveur de l'**hypothèse de sélection**. Les lecteurs d'un texte ne seraient ainsi pas conscients de l'importance de la structure du document si elle ne leur est pas fournie.

La mémorisation des éléments d'un texte est donc favorisée par la présence d'indices de structure. C'est également le cas de la RI : en effet, localiser une information dans un document requiert de la part des individus qu'ils soient conscients de l'organisation de ce document, et du contenu qu'il propose, en s'aidant alors des signaux organisationnels à leur disposition (Guthrie & Mosenthal, 1987 ; Lorch, 1989 ; Puustinen & Rouet, 2009). Les titres des différentes parties constituant le document pédagogique sont particulièrement utiles pour aider à l'interprétation du contenu du document (Lorch, 1989). Ils faciliteraient l'accès direct à une partie spécifique du document selon le choix de l'individu. Il peut ainsi décider quelle information voir, et la navigation au sein du document est rendue plus aisée. Cette activité semble pertinente dans le cadre d'une tâche de RI.

Cependant, il apparaît que la majorité des études testant les effets des signaux organisationnels présentées dans la littérature a été réalisée sur des documents écrits

(e.g., Brooks, Dansereau, Spurlin, & Holley, 1983 ; Lorch, 1989 ; Lorch & Lorch, 1996 ; Rouet, Vidal-Abarca, Erbou, & Millogo, 2001a ; Sanchez et al., 2001). La particularité des documents écrits est que le lecteur peut avoir une vue globale assez facilement de l'ensemble du contenu, en survolant les pages par exemple (van der Meij & van der Meij, 2014). Par ailleurs, des chercheurs se sont intéressés aux hypertextes : il s'agit de multiples documents connectés entre eux par des liens que peuvent utiliser les lecteurs pour passer d'un texte à un autre (Amadiou et al., 2015 ; DeStefano & LeFevre, 2007). La structure non linéaire spécifique aux hypermédias implique des choix à réaliser de la part de l'utilisateur, et un coût cognitif supplémentaire par rapport à un document classique (DeStefano & LeFevre, 2007 ; Vörös, Rouet, & Pléh, 2011). Connaître la structure de l'hypertexte est donc nécessaire : la présentation d'une carte conceptuelle (i.e., vue globale des concepts principaux de l'hypertexte et des relations entre ces concepts ; Amadiou & Salmerón, 2014 ; Potelle & Rouet, 2003) permettrait de fournir la macro-structure de l'hypertexte. Cela favoriserait également une représentation mentale pertinente de la structure et des contenus de l'hypertexte, et réduirait le coût cognitif associé au traitement d'un tel document (Amadiou & Salmerón, 2014).

Dans le cas des vidéos, si le contenu est présenté linéairement, il n'est pas toujours possible d'accéder directement à certaines parties, en effectuant un survol comme dans les documents écrits. Il est donc délicat de naviguer dans une vidéo pauvre en interactivité, à part en visionnant la vidéo de manière linéaire en attendant de trouver l'information recherchée (D. Zhang et al., 2006). Il semble donc important d'avoir d'autres possibilités que la navigation linéaire dans une vidéo dans le cas d'une activité de RI. Ces possibilités de contrôle du contenu agissent au niveau de la **macro-structure**. La macro-structure représente la structure globale du texte, elle organise notamment les éléments de la micro-structure de manière hiérarchique (Kintsch, 1998 ; Kintsch & Yarbrough, 1982). D'après Amadiou et Salmerón (2014), dans le cadre de travaux sur les cartes conceptuelles, fournir la macro-structure du document étudié permettrait également de renseigner sur l'organisation sémantique de ce document. Dans une vidéo, faciliter le traitement de l'information à un niveau global et donc l'identification sémantique est rendu possible par des activités de **macro-niveau**.

4.3.2. Les activités de macro-niveau

Les activités de macro-niveau ont lieu à un niveau plus global que les activités de micro-niveau. Elles ne sont pas relatives à la navigation au sein d'une phrase, mais se réfèrent à la navigation à travers le document entier (Merkt et al., 2011). Les activités de macro-niveau sont rendues possibles grâce notamment aux tables des matières ou index. L'idée est de permettre de localiser, relier et comparer différentes parties du document étudié (Chun & Plass, 1996 ; Merkt et al., 2011). Le terme de signaux organisationnels est parfois utilisé et regroupe plusieurs possibilités d'activités de macro-niveau. Il s'agit plus précisément d'ajouter un mot ou une partie de phrase (i.e., un titre) annonçant un contenu particulier avant qu'il ne soit présenté et détaillé. Les signaux présentent ainsi la structure du document pour aider l'individu qui navigue dans le document à identifier l'organisation hiérarchique du contenu (Loman & Mayer, 1983 ; Lorch, 1989 ; Spyridakis & Standal, 1987). Ils ont ainsi un effet positif sur la mémorisation du texte et la compréhension du document, en permettant notamment à l'individu de saisir l'organisation de l'information (Lorch, 1989). Des effets bénéfiques dans les tâches de RI peuvent également se retrouver dans la littérature.

Lorch et ses collaborateurs (2011), dans une série d'expérimentations, ont ainsi testé l'effet de la présence de titres dans un document écrit, afin de distinguer quel niveau d'information organisationnelle ils pouvaient apporter. Les auteurs ont ainsi voulu savoir si les titres permettaient une organisation hiérarchique (i.e., informations sur les liens entre parties) ou bien une organisation séquentielle (i.e., informations sur l'ordre des parties dans le document) du contenu. Quatre expériences ont été menées :

- Dans la première expérience, deux conditions ont été comparées. La première condition présentait un document dont les titres changeaient de typographie (e.g., surlignage, majuscules, mise en gras) en fonction du niveau hiérarchique de la partie du document qui y était rattachée (i.e., organisation hiérarchique). Dans la seconde condition, les titres étaient tous présentés de la même manière, sans distinction selon le niveau hiérarchique (i.e., organisation séquentielle). Les résultats concernant le nombre moyen de titres rappelés et bien positionnés ont montré une supériorité du format avec les titres hiérarchisés.

- La deuxième expérience reprenait la mise en œuvre de la première, mais la hiérarchie du document était représentée cette fois par des numéros associés aux titres, toujours selon le niveau (e.g., 1.1.1, 1.1.2). Dans la condition non hiérarchisée, les titres étaient associés à de simples chiffres qui se suivaient, de 1 à 24. Les résultats ont mené aux mêmes conclusions que la première étude. Les auteurs ont alors postulé que ce ne sont pas les titres qui sont utiles, mais l'information hiérarchique communiquée par les titres.
- L'expérience 3 consistait en une tâche de RI, dans un document dont les titres étaient composés soit uniquement d'une indication chiffrée (e.g., partie 1, partie 2), soit d'une couleur (e.g., partie bleue, partie verte). Il apparaît que les participants sont davantage efficaces en RI (i.e., plus rapides pour localiser l'information demandée) lorsque les parties sont numérotées.
- La quatrième et dernière expérience consistait également en une tâche de RI, cette fois dans un document présentant les titres associés avec soit une organisation hiérarchique (e.g., 1.A.1, II.C.2), soit une organisation séquentielle (e.g., 1, 2). Les individus réalisant la RI ont été plus rapides dans la condition avec les titres présentés de manière séquentielle.

En conclusion de cette série d'études, les auteurs n'ont pu conclure qu'à des effets distincts des deux types de présentation des titres. Les titres communiquant l'organisation hiérarchique du document se sont avérés plus judicieux lors d'une tâche de rappel du plan, alors que les titres communiquant l'organisation séquentielle du document ont davantage aidé lors des tâches de RI.

Une étude menée par Dreher et Guthrie (1990) a montré que les participants (dans cette étude, des enfants) les plus efficaces en RI, c'est-à-dire ceux qui ont passé le moins de temps total à réaliser la tâche, ont passé davantage de temps sur l'index et la table des matières. Ces deux aides favorisent la sélection des catégories et ont lieu au début de la tâche de RI. Les individus efficaces en RI ont ainsi dévoué moins de temps que les autres à extraire les informations après. L'activité de planification, lorsqu'elle a été favorisée et bien réalisée, a donc mené à une meilleure RI. Ces résultats sont consistants avec le modèle précédemment présenté (voir Introduction théorique 3.3) de Sharit et ses collaborateurs (2008), qui soulignait l'importance de la planification dans la RI, et l'intérêt de faciliter cette étape grâce à une bonne

représentation mentale du document. Les signaux organisationnels sont des outils favorisant les différentes étapes dans la réussite d'une tâche de RI. L'utilisation de stratégies dites *top-down*, c'est-à-dire le fait de partir d'une table des matières, d'un index, etc., est un prédicteur significatif de la vitesse de recherche chez les enfants (Rouet & Coutelet, 2008). De la même manière, les participants retrouvent plus facilement les informations qui sont indexées dans le document que celles qui ne le sont pas (Yussen, Stright, & Payne, 1993).

Dans une recherche réalisée sur un document textuel (Krug, George, Hannon, & Glover, 1989), l'effet de la présence ou de l'absence des titres et d'un plan présenté en première page a été étudié. La condition expérimentale qui présentait aux individus le texte accompagné de titres et du plan s'est révélée supérieure en termes de scores sur un test de mémorisation que les conditions avec un seul type de signal, ces conditions étant elles-mêmes supérieures à la condition sans titre ni plan. Frank et ses collaborateurs (1989) ont réalisé une étude similaire à celles-ci, mais au sein d'un document vidéo. Durant les quinze minutes que durait la vidéo, les titres des parties présentées apparaissaient ou non selon la condition, de même qu'un plan sur papier de la vidéo était disponible ou non pour l'apprenant. Deux tests étaient proposés après la phase d'apprentissage : un rappel immédiat sans relecture possible des notes, et un rappel différé avec possibilité de relecture des notes prises durant l'apprentissage. Les résultats ont montré une interaction significative entre les variables « titres » et « plan » : les participants étaient meilleurs lorsque les deux étaient présents ensemble, ou lorsqu'aucun des deux n'était présent. Des hypothèses ont alors été formulées par les auteurs pour expliquer l'effet négatif qui avait lieu lorsqu'un seul de ces types de signaux était présent dans la vidéo. Lorsque le plan est proposé sans les titres, les participants le regarderaient tout au long de l'apprentissage pour connaître ce qui sera présenté par la suite dans la vidéo. De ce fait, ils prêteraient moins d'attention à la vidéo en elle-même. Lorsque seuls les titres au cours de la vidéo sont disponibles, les participants manqueraient de la vision globale permise par le plan. La mise en place

de ces deux types de signaux en même temps permettrait d'annuler les effets négatifs propres à chacun¹⁶.

D'autres études ont plus particulièrement permis de se concentrer sur les stratégies déployées avec ou sans ce type d'aide. Ainsi, une analyse des mouvements oculaires durant la lecture d'un texte, réalisée par Hyönä (1995), a démontré que les individus lisent plus lentement les phrases annonçant un nouveau sujet que les autres phrases au sein du document. Cet effet a lieu lors de la première lecture du texte, et ne se reproduit pas lors des nouvelles lectures. L'hypothèse alors avancée est qu'en première lecture, les individus doivent construire une représentation mentale du document. Ils utilisent la structure du texte pour améliorer leur représentation et s'en servent par la suite comme d'un guide pour les lectures suivantes. Dans le même sens, Rouet et ses collaborateurs (2001a) ont testé l'influence d'une vue globale structurée du document sur les stratégies de RI. Deux conditions étaient comparées : un texte était présenté, sa première page comprenait soit le titre et une vue globale du document (i.e., une table des matières), soit uniquement le titre. Les résultats ont indiqué que la présentation d'un signal organisationnel a eu pour effet de réduire le nombre de paragraphes étudiés pendant la RI. L'hypothèse soutenue par les auteurs voulant que la vue globale du document aide la localisation de l'information en favorisant et structurant la représentation mentale des individus est consistante avec les résultats trouvés dans cette étude et les études précédentes, ainsi qu'avec le modèle proposé par Sharit et ses collaborateurs (2008). Par ailleurs, les effets positifs sur l'apprentissage des titres dans un texte ont également été discutés en termes d'effet sur la construction du modèle mental par Mayer et ses collaborateurs (1984).

En résumé, la présence de signaux organisationnels semble donc bénéfique pour les utilisateurs lors d'activités d'apprentissage ou de RI. L'ajout d'une table des matières (i.e., affichage des titres des parties du document) est le type de signal le

¹⁶ Dans cette étude (Frank, Garlinger, & Kiewra, 1989), les groupes expérimentaux se composent globalement de peu de sujets (parfois 10 ou 11 participants). Les statistiques et les conclusions apportées par les auteurs sont donc à prendre avec précaution.

plus couramment utilisé dans la littérature. De ce fait, puisqu'il semble soutenir la navigation au sein de la macro-structure du document, l'ajout d'une table des matières peut constituer un étayage de macro-niveau, autrement dit un **macro-étayage**.

La segmentation d'une vidéo (i.e., micro-étayage) permettrait également à l'utilisateur de contrôler le rythme de la vidéo et de naviguer en son sein selon ses besoins et capacités cognitives. La structuration de la vidéo (i.e., macro-étayage) permettrait quant à elle de maîtriser le contenu de la vidéo et d'identifier plus aisément une information cible. Les bénéfices en termes de modèle mental sont à évaluer selon ces deux types d'étayage, chacun favorisant à sa manière l'appréhension de l'environnement vidéo par l'utilisateur. En effet, une représentation mentale pertinente de la vidéo est la base d'une RI et d'un apprentissage efficaces.

5. Effets attendus de l'étayage

5.1. Problématique de la thèse

Dans un contexte pédagogique, les activités les plus demandées aux étudiants aujourd'hui sont des activités de RI et d'apprentissage. Elles ont lieu de plus en plus dans des environnements multimédias, et notamment au sein de vidéos. La RI et l'apprentissage sont fortement liés : la localisation de l'information intervient dans le processus d'apprentissage. De plus, une variable commune apparaît : le besoin d'avoir une représentation mentale du document pertinente. Dans le cas de la RI, cette représentation, de la vidéo par exemple, va constituer la base du processus de localisation. Dans le cas de l'apprentissage, c'est de la qualité du modèle mental que vont dépendre les performances finales.

Afin de faciliter la RI et l'apprentissage, l'apport d'étayage doit servir de modèle conceptuel fourni à l'individu : en RI, il va servir de représentation externe de la vidéo ; en apprentissage, il va faciliter la construction d'un modèle mental adapté. Des contraintes doivent cependant être prises en compte. Les deux niveaux possibles d'étayage interagissent fortement, c'est-à-dire que les informations qu'ils transmettent se répondent entre elles, et le coût cognitif associé à l'intégration des informations issues de ces deux niveaux doit être réduit au possible. Enfin, dans le cas d'une tâche d'apprentissage, l'individu doit pouvoir être engagé dans l'activité pour utiliser au mieux les ressources à sa disposition, telles qu'un étayage.

Cette thèse s'attachera à évaluer et comprendre le rôle de l'étayage lors d'activité de RI ou d'apprentissage au sein d'un environnement vidéo. Deux niveaux d'étayage sont possibles, et il s'agira également de distinguer les spécificités de chacun. Les expérimentations mises en place serviront à tester l'effet de l'étayage en RI ou en apprentissage, ainsi que différentes méthodes pour faciliter le traitement de l'information et améliorer les performances des individus.

5.2. Effet attendu de l'étayage sur la RI

Le succès du processus de RI dépend de l'efficacité de la localisation de l'information (Lawless & Brown, 1997). L'étayage, comme vu précédemment, vient

ainsi aider l'utilisateur dans la localisation de l'information, en agissant notamment sur le premier niveau du processus de RI selon Sharit et ses collaborateurs (2008) : la représentation du document. Les effets de l'étayage dans des vidéos ont déjà été testés, surtout en ce qui concerne l'effet des pauses (e.g., Biard et al., 2017 ; Schwan & Riempp, 2004 ; Spanjers et al., 2012 ; Wouters et al., 2007). Les pauses sont en effet considérées comme de l'étayage de micro-niveau (en ce qu'elles permettent de contrôler le flot d'information délivrée). Cependant, l'effet des pauses ayant fait l'objet de nombreux articles dans la littérature, et la présence de boutons pause s'étant largement développée dans la présentation des vidéos, il semble désormais intéressant de se pencher sur un autre type de micro-étayage : la segmentation de la vidéo. Ainsi, deux niveaux d'étayage peuvent potentiellement avoir un effet sur la RI dans les vidéos : la segmentation (i.e., micro-étayage) et la structuration (i.e., macro-étayage). Ces deux niveaux, s'ils sont utilisés conjointement au sein du même environnement vidéo, permettraient de lier les informations qu'ils transmettent.

Une étude menée par Merkt et Schwan (2014) a examiné les effets de différentes possibilités d'étayage sur des tâches de RI et d'apprentissage. Quatre conditions expérimentales étaient comparées :

- Une condition présentant un texte illustré de 46 pages, reprenant la transcription de la vidéo présente dans les trois autres conditions ;
- Une condition présentant une vidéo non-interactive, avec seulement un bouton de démarrage de la vidéo ;
- Une condition avec vidéo « ordinaire » (*common video*), c'est-à-dire que l'utilisateur pouvait se déplacer dans la vidéo en se servant de la barre de navigation, et faire pause à tout moment ;
- Une condition avec vidéo « améliorée » (*enhanced video*) dans laquelle l'utilisateur pouvait se déplacer dans la vidéo avec la barre de navigation, faire pause, et/ou utiliser une table des matières et un index sur le contenu de la vidéo, accessibles par des boutons situés à côté de la vidéo.

En d'autres termes, la condition « vidéo ordinaire » servait à tester l'effet des activités de micro-niveau, permises par le déplacement dans la barre de navigation et la possibilité de faire pause (micro-étayage). La condition « vidéo améliorée » peut quant à elle être associée à un étayage de micro-niveau (i.e., le même que dans la condition

« vidéo ordinaire ») ainsi qu'un étayage de macro-niveau, grâce à la présence d'une table des matières et d'un index. Concernant l'hypothèse sur la RI, les auteurs s'attendaient à ce que la vidéo améliorée permette des performances meilleures que la vidéo ordinaire, elle-même supérieure à la vidéo non-interactive (i.e., sans étayage). Onze questions étaient posées aux participants, nécessitant chacune une RI dans la vidéo. Les résultats ont montré que les participants ayant cherché les informations dans le texte écrit ou dans la condition vidéo améliorée ont trouvé davantage de réponses que les participants avec la vidéo ordinaire, eux-mêmes meilleurs que ceux avec la vidéo non-interactive. Seulement, en se penchant plus précisément sur les stratégies employées lors de l'activité de RI, Merkt et Schwan (2014) n'ont pu dégager que l'usage de l'index comme prédicteur de la performance en RI. Ni la table des matières, ni la navigation dans la vidéo, ni l'utilisation du bouton pause n'ont eu d'effet. Deux limites cependant se posent quant à cette étude. Premièrement, les auteurs n'ont comparé une condition vidéo « contrôle » (i.e., sans étayage, vidéo non-interactive) qu'avec des vidéos proposant du micro-étayage (i.e., condition vidéo ordinaire) ou bien à la fois du micro- et du macro-étayage (i.e., condition vidéo améliorée). Une condition proposant uniquement de l'étayage de macro-niveau (e.g., table des matières, index) aurait pu permettre d'évaluer les effets spécifiques de ce type d'étayage. De plus, nous supposons que la présentation de la table des matières dans cette étude a souffert d'un biais potentiel lié à l'éloignement avec le document en question, la vidéo. En effet, les utilisateurs devaient cliquer sur un bouton à côté de la vidéo pour avoir accès à la table des matières, l'image de la vidéo n'étant alors plus disponible. Une telle présentation pourrait alors engendrer un effet d'attention divisée (lorsque l'individu doit prêter attention à l'information issue de sources différentes et les intégrer mentalement)¹⁷ qui serait néfaste pour l'utilisateur (Ayres & Sweller, 2014). Selon le principe de contiguïté spatiale, les performances seraient améliorées si les différentes sources d'informations ayant un lien entre elles (ici, la vidéo et la table des matières, ou plus généralement la vidéo et son étayage) étaient affichées côte à côte sur l'écran (Johnson & Mayer, 2012 ; Mayer, 2005 ; Mayer & Fiorella, 2014 ; Moreno & Mayer,

¹⁷ *“Instructional split attention occurs when learners are required to split their attention between and mentally integrate several sources of physically or temporally disparate information, where each source of information is essential for understanding the material”* (Ayres & Sweller, 2014, p. 206).

1999). Une solution aurait donc été de placer la table des matières sur le même écran que la vidéo, afin de réduire l'effort induit par les allers-retours successifs entre les deux sources d'information (Mayer & Fiorella, 2014). Ce principe de contiguïté spatiale a été éprouvé dans nombre d'études (e.g., Cierniak, Scheiter, & Gerjets, 2009 ; Erhel & Jamet, 2006 ; Johnson & Mayer, 2012) axées sur l'apprentissage, et nous supposons alors qu'il peut s'appliquer également lors d'une activité de RI. Les différentes sources d'information (ici, les différents niveaux d'étayage et la vidéo correspondante) doivent pouvoir se répondre les unes aux autres au sein du même écran dans l'environnement vidéo.

Finalement, l'étayage est attendu comme ayant un effet positif sur les performances en RI. Deux niveaux d'étayage peuvent être envisagés dans un environnement vidéo : un micro-étayage (i.e., segmentation), et un macro-étayage (i.e., structuration). Les effets de ces deux niveaux devraient être meilleurs encore lorsqu'ils sont présentés en même temps à l'utilisateur, sur le même écran. Ces hypothèses feront l'objet de l'**étude 1** de la présente thèse.

Un effet positif de l'étayage est également attendu sur le modèle mental de l'utilisateur. En effet, si l'étayage aide à la performance en RI, c'est qu'il doit améliorer la représentation du document (cf. modèle développé par Sharit et al., 2008). Dans le cas des animations, et de leur effet sur les performances après une tâche d'apprentissage, Höffler et Leutner (2007) ont proposé l'idée qu'elles fournissaient un modèle externe pour une représentation interne à l'apprenant. Dans le cas d'une activité de RI, si l'étayage est envisagé comme un modèle conceptuel fourni à l'utilisateur (voir Introduction théorique 4.1.1), alors il aiderait à la construction d'un modèle mental pertinent (Borgman, 1986 ; Norman, 1983). Il s'agira alors de savoir si l'étayage aide la représentation interne qu'a l'utilisateur du document (i.e., son modèle mental de la vidéo), ou bien s'il sert de représentation externe sans servir à la construction d'un modèle mental. Cette hypothèse peut être envisagée dans l'idée que la construction d'un modèle mental de la vidéo n'est pas nécessairement utile lors d'une consigne de RI, contrairement à lors d'un objectif d'apprentissage. Cette question du rôle de l'étayage sur le modèle mental de l'utilisateur sera étudiée lors de l'**étude 2** de cette thèse.

Enfin, l'effet d'attention divisée qui pourrait apparaître entre l'étayage et la vidéo, peut également s'envisager à un second niveau. En effet, les deux niveaux d'étayage (i.e., segmentation et structuration), lorsqu'ils sont présents conjointement dans l'environnement vidéo, sont supposés interagir fortement l'un avec l'autre. L'objectif de l'**étude 3** sera donc de comparer différents formats de présentation des deux niveaux d'étayage, afin d'éviter un traitement supplémentaire coûteux et néfaste pour l'utilisateur dans le cadre d'une tâche de RI.

5.3. Effet attendu de l'étayage sur l'apprentissage

Comme développé précédemment, des effets positifs de l'étayage sont attendus lors d'une activité de RI. Or, la pertinence de la localisation de l'information serait la base pour un apprentissage efficace (voir Introduction théorique 3.2). Des effets positifs de l'étayage peuvent donc être attendus également lors d'une activité d'apprentissage. De plus, pour un apprentissage pertinent, l'apprenant doit être engagé dans un processus actif de traitement de l'information (voir Introduction théorique 2.1.2), c'est-à-dire qu'il devra pouvoir internaliser l'étayage présenté en tant que modèle conceptuel pour la construction d'un modèle mental pertinent. L'hypothèse que dans le cadre d'une activité centrée sur l'apprentissage l'apprenant va utiliser l'étayage pour être plus actif, et donc que l'étayage améliore les performances d'apprentissage, sera testée et discutée lors de l'**étude 4** de cette thèse.

Cette hypothèse ne tient que si l'on suppose que la seule présence d'une consigne d'apprentissage va suffire à engager l'apprenant dans un processus actif. Dans le cas où ce ne serait pas le cas, ou du moins où ce ne serait pas suffisant, la question se pose de savoir comment rendre l'apprenant plus actif et l'inciter à traiter l'étayage mis à sa disposition. Des études ont été réalisées sur l'effet des tests (i.e., *testing effect*), qu'ils aient lieu avant ou pendant la phase d'apprentissage du document (Fernandez & Jamet, 2017 ; Mayer, Dow, & Mayer, 2003, étude 3 ; Reynolds, Standiford, & Anderson, 1979). Selon la définition du *testing effect*, un effet positif des questions posées sur le matériel d'apprentissage, par rapport par exemple à une simple révision du contenu, est attendu sur l'apprentissage (Carpenter & Toftness, 2017). L'effet de l'introduction d'une tâche de RI avant une phase d'apprentissage en vidéo sera testé au cours de l'**étude 5** de cette thèse.

ÉTUDE 1 : Effet de l'étayage sur la performance en RI dans une vidéo¹⁸

1. Introduction et hypothèses

L'hypothèse générale de cette première étude est que l'étayage va servir de modèle conceptuel dans le cas d'une RI au sein d'une vidéo. Ainsi l'étayage soutiendrait la représentation mentale qu'a l'utilisateur du document vidéo. La représentation mentale correspond au premier sous-processus de la localisation développé par Sharit et ses collaborateurs (2008). Concernant les deux autres étapes (i.e., planification et exécution), elles devraient également être facilitées de deux manières par l'apport d'étayage : grâce à l'utilisation d'un modèle mental amélioré (conséquence directe de la représentation mentale facilitée) et les outils fournis (i.e., marqueurs, chapitres). Ces outils, permettant des actions concrètes dans l'environnement vidéo par de simples clics, peuvent s'entendre sous le terme **interactivité** (Scheiter, 2014 ; Scheiter & Gerjets, 2007). L'ajout d'interactivité dans un document multimédia a déjà fait l'objet d'études et Zhang et ses collaborateurs (2006) ont notamment démontré son apport dans l'allègement des difficultés des utilisateurs dans une tâche de navigation au sein d'une vidéo. Si aucun modèle conceptuel n'est fourni, par absence d'étayage, les utilisateurs devraient avoir besoin de davantage de temps pour se construire un modèle mental par eux-mêmes et pour améliorer leurs performances. Un seul niveau d'étayage sur les deux possibles devrait fournir un modèle conceptuel incomplet, et donc nécessiter que l'utilisateur se construise par lui-même un modèle mental pour le compléter. Dans les cas où le modèle mental doit être construit, une fois celui-ci créé, les performances de l'utilisateur devraient alors devenir aussi bonnes que celles rendues possibles grâce à la provision d'un modèle conceptuel par étayage.

¹⁸ Cette étude a fait l'objet d'une publication dans une revue internationale à comité de lecture : Cojean, S., & Jamet, E. (2017). Facilitating information-seeking activity in instructional videos: The combined effects of micro- and macroscaffolding. *Computers in Human Behavior*, 74, 294-302, <https://doi.org/10.1016/j.chb.2017.04.052>

L'objectif de cette première étude est donc d'investiguer les effets potentiels des deux niveaux d'étayage (i.e., micro- et macro-niveaux) sur la RI dans un environnement vidéo, ainsi que l'effet de leur interaction quand ils sont utilisés conjointement et qu'ils se font référence. Dans le cadre de cette étude, le micro-étayage correspond à une segmentation qui prend la forme de marqueurs le long de la barre de navigation de la vidéo (voir Figure 4). La navigation ainsi permise devrait pouvoir faciliter les activités de micro-niveau, c'est-à-dire le **contrôle** du flot d'informations délivrées et le traitement de l'information à un niveau local dans la vidéo. Le macro-étayage, quant à lui, est testé sous la forme d'une table des matières sur le côté de la vidéo (voir Figure 4). L'objectif est alors d'inciter les activités de macro-niveau, c'est-à-dire la représentation structurelle du document et son organisation **sémantique**.

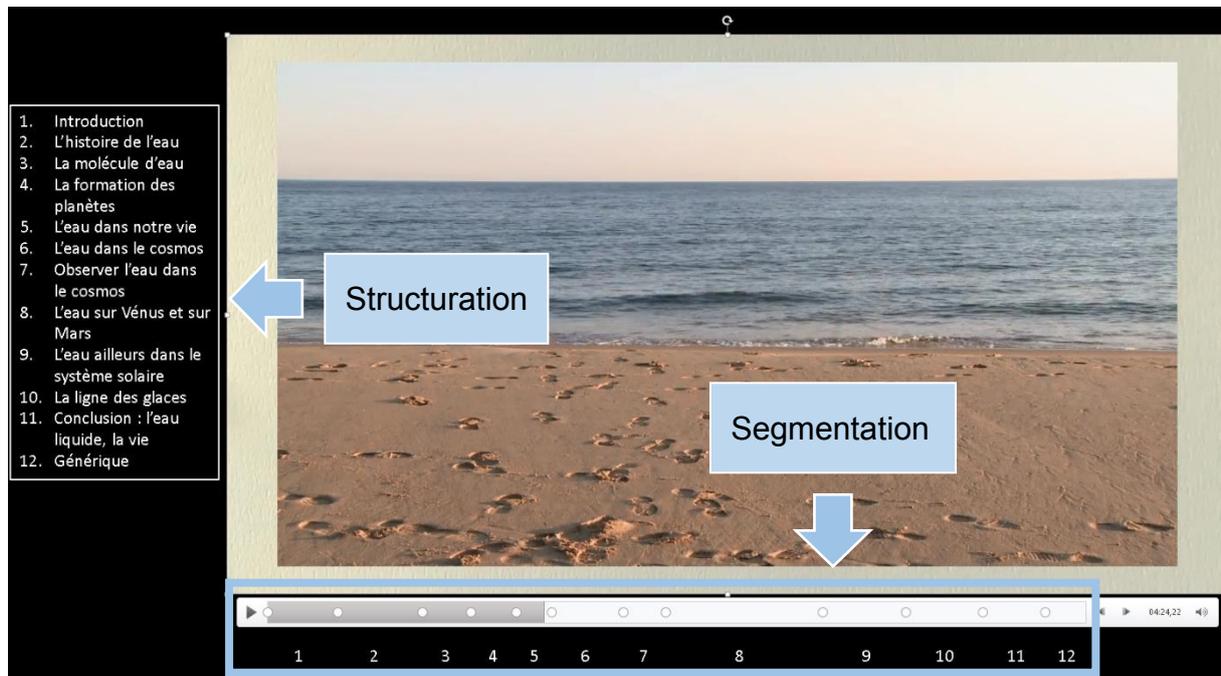


Figure 4. Segmentation et structuration au sein d'un environnement vidéo (étude 1)

La composante dynamique de l'évolution des performances au cours de la RI fera l'objet d'un regard particulier. Pour cela, il a été décidé d'analyser les réponses des participants sur neuf questions de RI successives, dont toutes les réponses se trouvaient au sein de la vidéo proposée. Six hypothèses ont donc été formulées :

Hypothèse 1. Réussite à la tâche :

Selon certains auteurs (e.g., Cataldo & Oakhill, 2000 ; Janosky, Smith, & Hildreth, 1986), la réussite à la tâche (i.e., localisation de la réponse attendue) serait un indicateur approprié pour distinguer une bonne RI. Les hypothèses concernant cette variable sont donc les suivantes :

- a) Les participants bénéficiant de l'étayage de micro- et/ou macro-niveau(x) devraient avoir de meilleures performances à la tâche de RI que les participants ne bénéficiant pas d'étayage. Cet effet devrait être maximal lorsque les deux niveaux d'étayage sont présents conjointement.
- b) Pour les participants bénéficiant d'un seul niveau d'étayage (micro- ou macro-niveau), ou pour les participants ne bénéficiant d'aucun étayage, leur performance à chaque question devrait s'améliorer au cours du temps. Cela refléterait la construction d'un modèle mental pertinent par l'utilisateur. Les participants bénéficiant conjointement des deux niveaux d'étayage devraient avoir une performance optimale, donc stable, au cours du temps.

Hypothèse 2. Temps de réponse :

Le temps passé à chercher la réponse attendue peut être considéré comme une mesure de l'efficacité de la RI (e.g., Cataldo & Oakhill, 2000 ; Rouet & Coutelet, 2008 ; Stagers & Norcio, 1993), les hypothèses concernant cette variable sont donc les suivantes :

- a) Les participants bénéficiant de l'étayage de micro- et/ou macro-niveau(x) devraient avoir un temps de réponse plus court (i.e., passer moins de temps à la RI) que les participants ne bénéficiant pas d'étayage. Cet effet devrait être maximal lorsque les deux niveaux d'étayage sont présents conjointement.
- b) Pour les participants bénéficiant d'un seul niveau d'étayage (micro- ou macro-niveau), ou pour les participants ne bénéficiant d'aucun étayage, leur temps de réponse à chaque question devrait diminuer au cours du temps. Cela devrait refléter la construction d'un modèle mental pertinent par l'utilisateur. Les participants bénéficiant des deux niveaux d'étayage conjointement devraient avoir une performance optimale, donc stable, au cours du temps.

Hypothèse 3. Pertinence du premier clic :

Une représentation mentale pertinente d'un système (i.e., environnement vidéo) doit permettre à l'utilisateur d'être plus précis quant à ses interactions avec le système (Staggers & Norcio, 1993). Les hypothèses relatives à la pertinence de l'utilisation de la vidéo en RI sont les suivantes :

- a) Les participants bénéficiant de l'étayage de micro- et/ou macro-niveau(x) devraient faire, pour chaque question, des premiers clics plus pertinents (i.e., plus proches du segment de vidéo contenant l'information recherchée) que les participants ne bénéficiant pas d'étayage. Cet effet devrait être maximal lorsque les deux niveaux d'étayage sont présents conjointement.
- b) Pour les participants bénéficiant d'un seul niveau d'étayage (micro- ou macro-niveau), ou pour les participants ne bénéficiant d'aucun étayage, la pertinence de leur premier clic à chaque question devrait s'améliorer au cours du temps. Cela devrait refléter la construction d'un modèle mental pertinent par l'utilisateur. Les participants bénéficiant des deux niveaux d'étayage conjointement devraient avoir une performance optimale, donc stable, au cours du temps

Hypothèse 4. Difficulté perçue :

Les participants bénéficiant de l'étayage de micro- et/ou macro-niveau(x) devraient percevoir la tâche comme étant moins difficile que les participants ne bénéficiant pas d'étayage. Cet effet devrait être maximum lorsque les deux niveaux d'étayage sont présents conjointement.

Hypothèse 5. Contrôle perçu :

Les participants bénéficiant de l'étayage de micro-niveau (i.e., marqueurs sur la barre de navigation) devraient ressentir davantage de contrôle sur la tâche que les participants ne bénéficiant pas de micro-étayage.

Hypothèse 6. Rappel des chapitres :

Les participants bénéficiant de l'étayage de macro-niveau (i.e., table des matières) devraient rappeler davantage de chapitres de la vidéo que les participants ne bénéficiant pas de macro-étayage.

2. Méthode

2.1. Participants

80 participants (59 femmes, 21 hommes) de l'Université Rennes 2 ont pris part à l'étude de manière volontaire. En moyenne, ils avaient 21.33 ans (E.-T. = 3.05). Les participants ont été recrutés au sein de l'université, notamment par des affiches postées dans les lieux de vie ou de passage (halls, couloirs, cafétérias...). Pour les remercier de leur participation, ils ont tous été gratifiés d'un ticket de cinéma. L'étude s'est déroulée conformément aux principes de la déclaration d'Helsinki.

2.2. Matériel

La vidéo qui a été utilisée provient du site Canal U (<http://www.canal-u.tv/>). Le titre de la vidéo était *Du Temps, de l'espace et de l'eau* (Doressoundiram, 2012), et sa durée était d'environ treize minutes. La vidéo abordait notamment le sujet de la présence d'eau sur Terre ou bien dans le reste de l'Univers, en présentant les manières de la détecter et son utilité. Comme le proposait déjà le site Internet, la vidéo a été segmentée en douze chapitres. Pour présenter la vidéo avec des marqueurs sur la barre de navigation et/ou des chapitres sur le côté, un environnement vidéo spécifique présenté aux participants sur un écran d'ordinateur a été conçu grâce au logiciel Powerpoint (voir Figure 5).

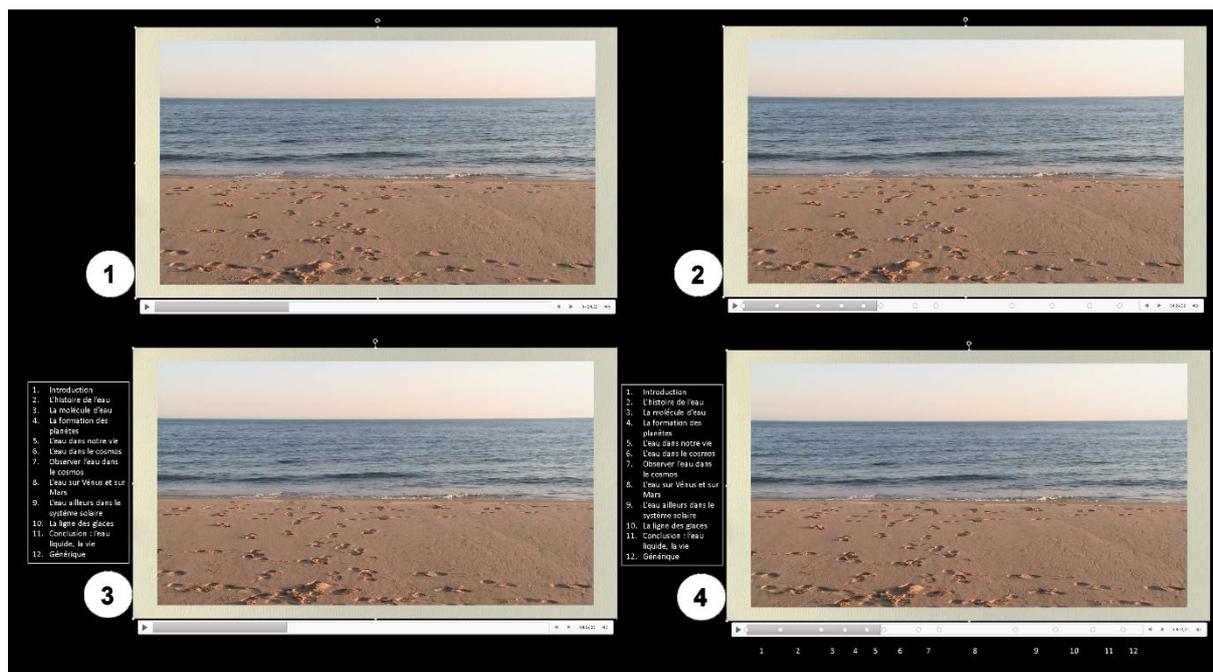


Figure 5. Environnements vidéo dans les conditions sans étayage (1), micro-étayage (2), macro-étayage (3) et double niveau d'étayage (4) de l'étude 1

Une barre de navigation sous la vidéo permettait aux participants de naviguer dans l'ensemble de la vidéo comme ils le souhaitent. Un plan factoriel en 2 x 2 a été utilisé, ce qui donnait quatre conditions expérimentales : la vidéo avait ou non des marqueurs sur la barre de navigation (douze sections correspondant aux douze chapitres, micro-étayage), de même qu'elle avait ou non une table des matières présentée sur le côté (macro-étayage). Un logiciel¹⁹ permettait d'enregistrer l'écran diffusant la vidéo et donc les informations de navigation des participants (e.g., clics, temps de consultation). À côté de l'écran d'ordinateur diffusant la vidéo, chaque participant avait une tablette tactile, sur laquelle les neuf questions RI étaient présentées, les unes à la suite des autres (e.g., « Sur quel satellite de Jupiter existerait-il un océan d'eau liquide ? » ; voir Annexe 1). À l'instar de plusieurs études portant sur la RI (e.g., He et al., 2008 ; Merkt & Schwan, 2014), un temps maximal a été défini pour répondre à chaque question. Chaque question était limitée à cinq minutes. Un chronomètre au-dessus de l'énoncé de la question indiquait le temps restant au participant pour répondre, et le décompte se lançait dès que le texte apparaissait. Un bouton sous la question permettait au participant de passer à la

¹⁹ Flashback by Blueberry Software, <https://www.flashbackrecorder.com/fbhome/>

question suivante s'il y avait répondu en moins de cinq minutes. Dans le cas où il n'y aurait pas répondu en cinq minutes, la question laissait automatiquement place à la suivante sans possibilité de retour en arrière. Les neuf questions étaient présentées dans un ordre contre-balancé, par groupes de trois questions, donnant ainsi trois ordres différents de présentations (1-2-3, 2-3-1 et 3-1-2). Le contre-balancement des questions a été mis en place pour limiter les effets d'ordre (Benton & Daly, 1991) et pour pouvoir analyser un effet du temps sur certaines des variables mesurées. Les participants répondaient aux questions sur neuf feuilles différentes placées sur le bureau, entre l'écran d'ordinateur et la tablette tactile. Après les questions de RI, s'ensuivait un post-questionnaire sur l'écran d'ordinateur. Des questions sur la difficulté perçue et le contrôle perçu étaient présentées dans un ordre aléatoire afin de limiter les effets d'ordre (Benton & Daly, 1991).

Les participants ont été aléatoirement répartis dans les quatre groupes correspondant aux quatre conditions expérimentales. Dans la **condition sans étayage** ($n = 19$), les participants pouvaient utiliser la barre de navigation pour se déplacer au sein de la vidéo, et pouvaient l'arrêter ou la redémarrer à n'importe quel moment. Le temps restant de la vidéo ainsi que la durée totale de celle-ci étaient indiqués. Ni la table des matières, ni les marqueurs sur la barre de navigation n'étaient disponibles (voir Figure 5). Dans la **condition micro-étayage**, les participants ($n = 19$) avaient le même matériel qu'en condition sans étayage, mais bénéficiaient en plus de marqueurs sur la barre de navigation. Ces marqueurs segmentaient la vidéo en douze sections correspondant aux douze chapitres développés, bien que les participants n'aient pas accès aux titres de ces chapitres (voir Figure 5). Dans la **condition macro-étayage**, les participants ($n = 20$) étaient exposés au même matériel qu'en condition sans étayage, à la différence qu'ils avaient une table des matières sur le côté de la vidéo. Les douze chapitres de la vidéo y étaient représentés, mais aucune référence n'y était faite dans la barre de navigation. Il n'était pas non plus possible de cliquer sur le titre d'un chapitre pour avoir accès à son début au sein de la vidéo (voir Figure 5). Enfin, dans la **condition à double niveau d'étayage**, les participants ($n = 22$) étaient confrontés à une vidéo présentant à la fois la table des matières et les marqueurs sur la barre de navigation correspondant aux chapitres. Afin de s'assurer que les deux niveaux d'étayage faisaient référence l'un à l'autre, des chiffres ont été assignés à

chaque chapitre et reportés sous chaque segment correspondant dans la barre de navigation (voir Figure 5).

2.3. Mesures

2.3.1. Intérêt préalable et compétence perçue

Avant la tâche de RI, deux questions étaient posées à chaque participant. Il s'agissait de vérifier que leur intérêt préalable quant au sujet développé dans la vidéo (i.e., l'eau dans l'Univers) et leur compétence perçue sur ce sujet étaient équitablement distribués parmi les quatre conditions expérimentales. Les participants devaient indiquer sur une échelle en 11 points leur degré d'intérêt (« Sur une échelle de 0 à 10, à quel point ce sujet vous intéresse-t-il ? ») et à quel niveau ils situeraient leur compétence sur ce sujet (« Sur une échelle de 0 à 10, à quel point vous estimez-vous compétent sur ce sujet ? »).

2.3.2. Réussite à la tâche

La réussite à chacune des neuf questions était évaluée de la manière suivante : une réponse était considérée comme correcte quand le participant avait noté la réponse attendue à la question (i.e., l'information recherchée) et le moment correspondant dans la vidéo où il l'avait trouvée. Une réponse était considérée comme incorrecte lorsque la réponse notée par le participant était incorrecte, lorsque le temps indiqué dans la vidéo était incorrect, ou lorsqu'il manquait au moins une de ces deux informations. Par définition, lorsque le participant ne trouvait pas la réponse en cinq minutes, la question passait à la suivante, et la réponse était considérée comme incorrecte.

2.3.3. Temps de réponse

Grâce au logiciel enregistreur d'écran, la durée de chaque RI (pour chaque question) par les participants a été analysée. Les temps de réponse ont été calculés à partir du moment où le participant commençait la recherche (i.e., premier clic dans la vidéo) jusqu'au moment où il trouvait l'information recherchée (i.e., bouton pause activé). Pour rappel, le temps de réponse était limité à cinq minutes maximum.

2.3.4. Pertinence du premier clic

L'utilisation de l'enregistreur d'écran permettait également d'avoir accès à la localisation du premier clic des participants lors de chacune de leurs RI. Pour chaque question, c'est-à-dire à chaque fois que le participant se mettait à chercher une information, il s'agissait de noter à quel point le segment contenant le premier clic du participant était loin du segment contenant la réponse attendue. Par exemple, si la réponse à la question se trouvait dans le chapitre 8 (donc segment 8), et que le premier clic du participant se faisait dans le segment 2, nous pouvions attribuer un score de 6 (8-2). Un score faible (i.e., taux d'erreur faible) signifie donc que le participant était proche de la réponse attendue, il est donc synonyme de grande pertinence. À l'inverse, un fort score (11 maximum, pour 12 segments) représente une faible pertinence du premier clic.

2.3.5. Difficulté perçue et contrôle perçu

Pour évaluer la difficulté et le contrôle perçus, des items ont été adaptés de précédentes études sur ces concepts (Kraft, Rise, Sutton, & Røysamb, 2005 ; Trafimow, Sheeran, Conner, & Finlay, 2002). Trois items de difficulté perçue ont finalement été utilisés (e.g., « Pour moi, effectuer cette recherche d'information était difficile » ; voir Annexe 5), ainsi que trois items de contrôle perçu (e.g., « J'ai eu un contrôle total sur la recherche d'information à effectuer » ; voir Annexe 6). Les participants devaient indiquer sur une échelle de Likert en 7 points (de 1 à 7) le degré avec lequel ils étaient d'accord avec l'affirmation proposée. Des alphas de Cronbach ont été réalisés, avec un résultat de $\alpha = .81$ pour les items de difficulté perçue et $\alpha = .75$ pour les items de contrôle perçu.

2.3.6. Rappel des chapitres

Après la tâche de RI, il était demandé aux participants de nommer les différentes parties présentées dans la vidéo, sans leur préciser qu'il y en avait douze. Pour les participants dans les groupes macro-étayage et double niveau d'étayage, il s'agissait d'une tâche de mémorisation, puisque la table des matières avec le nom des chapitres était présente lors de la RI. Pour les participants des groupes sans étayage

et micro-étayage, qui n'avaient donc jamais vu la table des matières, il s'agissait alors davantage d'une tâche d'inférences.

2.4. Procédure

Une fois les participants accueillis, il leur était proposé de s'installer à un bureau. Ils devaient alors compléter deux questions (pré-questionnaire) sur leur intérêt préalable à propos de « L'Eau dans l'Univers », ainsi que leur compétence perçue sur le sujet. L'expérimentateur donnait et expliquait ensuite la consigne de l'étude (répondre à des questions en cherchant la réponse dans la vidéo ; voir Annexe 7) et décrivait le matériel utilisé (table des matières et/ou marqueurs sur la barre de navigation selon les groupes expérimentaux). Tous les participants ont été tenus informés de l'ensemble des mesures qui seraient effectuées, et il leur a été précisé qu'ils pouvaient se retirer de l'étude s'ils le souhaitaient. L'expérimentateur démarrait ensuite l'enregistrement de l'écran, et les participants mettaient un casque audio pour écouter la vidéo. Ils pouvaient commencer la tâche quand ils le souhaitaient, en cliquant sur la tablette pour que la première question apparaisse. Ils avaient ensuite cinq minutes pour répondre à chaque question, en naviguant au sein de la vidéo. Une fois qu'ils avaient trouvé la réponse à une question et l'endroit dans la vidéo où elle apparaissait, ils l'écrivaient sur la feuille de papier correspondant. Ils pouvaient alors passer à la question suivante. S'ils ne trouvaient pas la réponse dans les cinq minutes, la question laissait automatiquement place à la suivante. Les participants n'inscrivaient donc rien sur la feuille de papier correspondant à la question non-répondue, et continuaient la tâche avec la question suivante. Après les neuf questions, l'expérimentateur arrêtait la vidéo et laissait le participant répondre au post-questionnaire (difficulté perçue, contrôle perçu, rappel des chapitres, sexe et âge du participant) sur l'écran d'ordinateur. La procédure est résumée en Figure 6. Après avoir répondu au post-questionnaire, l'expérimentateur gratifiait chaque participant d'un ticket de cinéma pour les remercier de leur participation à l'étude.

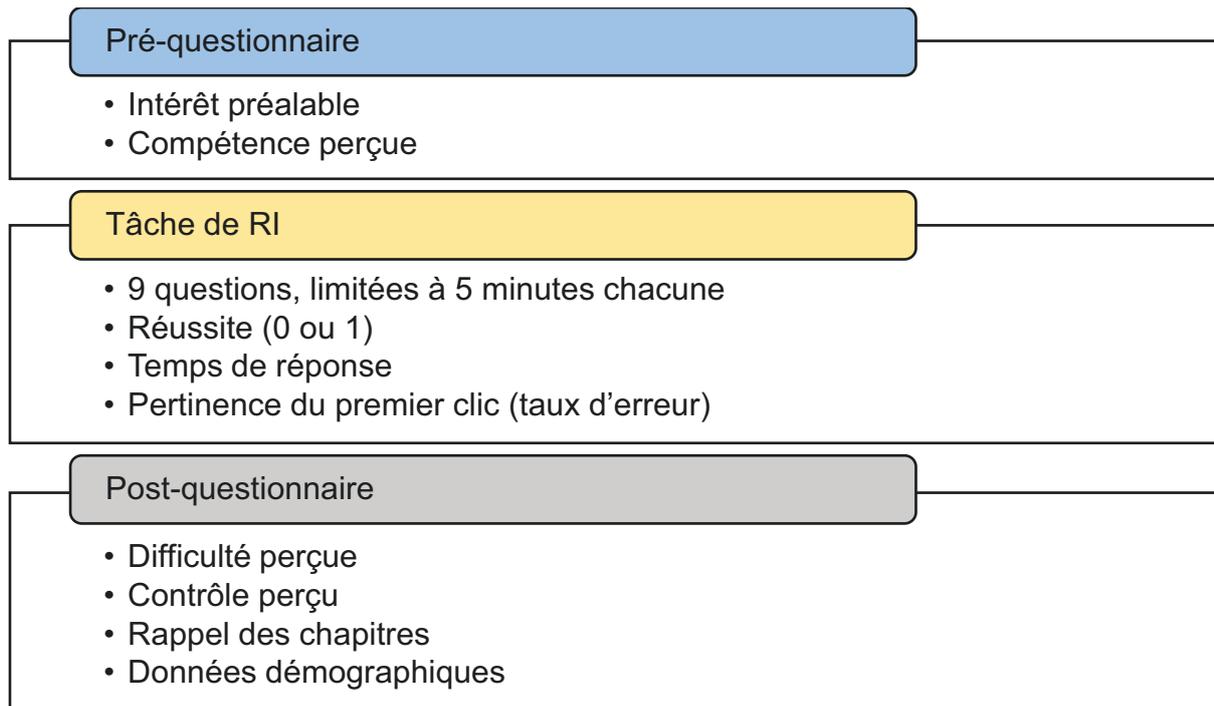


Figure 6. Procédure de l'étude 1

3. Résultats

Le logiciel R a été utilisé pour l'ensemble des analyses statistiques (R Core Team, 2018).

3.1. Variables contrôles (pré-questionnaire)

L'objectif du pré-questionnaire était de s'assurer que les résultats des participants des quatre conditions expérimentales ne différaient pas quant à leur intérêt préalable et leur compétence perçue sur le sujet présenté dans la vidéo. Des analyses de variance (ANOVA) ont été conduites et n'ont révélé aucune différence significative entre les conditions, que ce soit sur l'intérêt préalable $F(3, 76) = 1.13, p = .341, \eta^2_p = .04$, ou sur la compétence perçue, $F(3, 76) = 0.29, p = .833, \eta^2_p = .01$ (pour les données descriptives, voir Tableau 3)²⁰.

²⁰ Pourcentage de variance expliquée. Il semble communément admis qu'un effet (mesuré à l'aide d'un éta-carré partiel η^2_p) est considéré comme faible à .01, moyen à .06 et fort à partir de .14 (Cohen, 1988).

Tableau 3. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 1)

Condition	Intérêt préalable		Compétence perçue	
	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	6.26	1.73	3.53	1.22
Micro-étayage	6.32	1.57	3.05	2.12
Macro-étayage	5.80	2.17	3.35	2.06
Double niveau d'étayage	6.82	1.65	3.55	1.95

3.2. Tâche de RI

3.2.1. Prérequis pour l'analyse statistique

Chaque participant s'est vu poser neuf questions, ce qui signifie que l'indépendance des données n'est pas respectée dans ce cas (Field, Miles, & Field, 2012). Pour contrer cette limite, des modèles linéaires mixtes ont été utilisés (Gueorguieva & Krystal, 2004). Il s'agit d'une méthode statistique qui permet de prendre en compte la non-indépendance des données et d'analyser des données répétées. Pour tester l'effet d'une variable avec la méthode des modèles mixtes, deux modèles imbriqués sont comparés entre eux à chaque fois. Plus spécifiquement, pour chaque effet testé, deux modèles sont comparés : un sans la variable (i.e., modèle de base, modèle nul), et un avec la variable testée (modèle plus complexe avec davantage de degrés de liberté ; e.g., Baayen, Davidson, & Bates, 2008). Pour s'assurer de la contribution de chaque variable testée, la différence de déviance (test de Chi-2) entre ces deux modèles imbriqués est calculée. Le seuil de significativité choisi pour les valeurs p a été fixé à $\alpha = .05$. Dans cette étude, pour chaque variable dépendante, les effets principaux de l'étayage micro-niveau et de l'étayage macro-niveau, ainsi que leurs effets additif et d'interaction ont été testés. Les effets additif et d'interaction avec le rang de la question (i.e., effet du temps) ont également été évalués. En tout, six modèles différents ont été comparés entre eux :

- un modèle nul n'intégrant pas l'effet de la condition expérimentale (M0) ;
- un modèle incluant l'effet du micro-étayage (M1a) ;

- un modèle incluant l'effet du macro-étayage (M1b) ;
- un modèle prenant en compte l'effet additif des deux niveaux d'étayage (M2) ;
- un modèle incluant l'effet d'interaction des deux niveaux d'étayage (M3) ;
- un modèle incluant l'effet additif du rang de la question (M4) ;
- un modèle prenant en compte l'effet d'interaction du rang de la question (M5).

Par exemple, pour tester l'effet du micro-étayage sur la réussite à la tâche, deux modèles ont été comparés : un qui n'incluait pas de prédicteur, et un incluant la variable micro-étayage. Si ce dernier réduisait significativement la déviance ($p < .05$), il était alors considéré comme étant meilleur que le premier (i.e., cette variable indépendante a une influence significative sur cette variable dépendante).

Les modèles sont comparés entre eux uniquement lorsque les degrés de liberté sont différents, en d'autres termes, lorsque des variables sont ajoutées, ou que les liens entre elles diffèrent. Les modèles M1a et M1b ne peuvent pour cette raison être comparés entre eux. Pour chaque variable dépendante, les modèles ont été testés dans cet ordre : M0 avec M1a, M0 avec M1b, le plus satisfaisant (entre M0, M1a et M1b) avec M2, le plus satisfaisant (M0, M1a, M1b, M2) avec M3, le plus satisfaisant (M0, M1a, M1b, M2 ou M3) avec M4, et enfin le plus satisfaisant à ce stade (entre M0, M1a, M1b, M2, M3 ou M4) avec M5. Dans chaque modèle, un effet aléatoire sur les questions a été ajouté afin d'absorber les éventuelles différences de difficultés entre les questions. De la même manière, un effet aléatoire sur les participants afin de prendre en compte la non-indépendance des données a été intégré (Baayen et al., 2008). La librairie lme4 (Bates, Mächler, Bolker, & Walker, 2015) a été utilisée pour estimer les modèles mixtes au sein de R.

3.2.2. Réussite à la tâche

Concernant les scores de réussite, pour chaque question il a été codé 0 (réponse incorrecte) ou 1 (réponse correcte). Ces données étant de nature binomiale, des régressions logistiques ont été employées pour les traiter (Field et al., 2012). Les données des temps de réponse et de la pertinence du premier clic ont quant à elles été traitées avec des régressions linéaires. Les résultats indiquent un effet significatif ($\chi^2(1, N = 720) = 9.36, p = .002$) du macro-étayage (M0 comparé à M1b), sur la

réussite à la tâche, mais pas d'effet significatif ($\chi^2(1, N = 720) = 3.48, p = .062$) du micro-étayage (M0 et M1a). Aucune contribution significative ($\chi^2(1, N = 720) = 3.69, p = .055$) du micro-étayage sur le macro-étayage n'a été trouvée (comparaison de M1b et M2), ni aucun effet significatif d'interaction entre ces deux niveaux (M1b et M3), $\chi^2(2, N = 720) = 5.77, p = .056$. Le modèle M1b (i.e., un effet significatif du macro-étayage) est donc considéré comme étant le meilleur modèle. Les données descriptives indiquent que la réussite à la tâche est plus élevée lorsque le macro-étayage est présent par rapport aux deux autres conditions expérimentales (voir Figure 7).

De plus, le changement au cours du temps de la réussite à la tâche a été analysé, en intégrant le rang de la question comme prédicteur. Plus précisément, il s'agissait d'étudier la réussite à la tâche selon la condition et la position de la question, de 1 (première question posée) à 9 (dernière question posée). Le modèle intégrant un effet du macro-étayage (M1b, sélectionné comme le meilleur modèle précédemment) a été comparé aux modèles additif (i.e., effet additif du macro-étayage et du rang de la question, M4) et d'interaction (i.e., effet d'interaction du macro-étayage et du rang de la question, M5). Les résultats ont montré un effet additif du macro-étayage et du rang de la question (M1b comparé à M4), $\chi^2(1, N = 720) = 4.80, p = .028$, ainsi qu'un effet d'interaction entre le macro-étayage et le rang de la question sur la réussite à la tâche (M4 et M5), $\chi^2(1, N = 720) = 6.74, p = .010$. Le modèle M5, correspondant ici à une interaction entre le macro-étayage et le rang de la question, apparaît finalement comme le modèle le plus approprié. Les statistiques descriptives indiquent que la différence entre les conditions avec macro-étayage et les deux autres conditions semblent s'amoinrir au cours du temps (selon le rang de la question), puisque leur score de réussite semble s'améliorer au cours du temps (voir Figure 7).

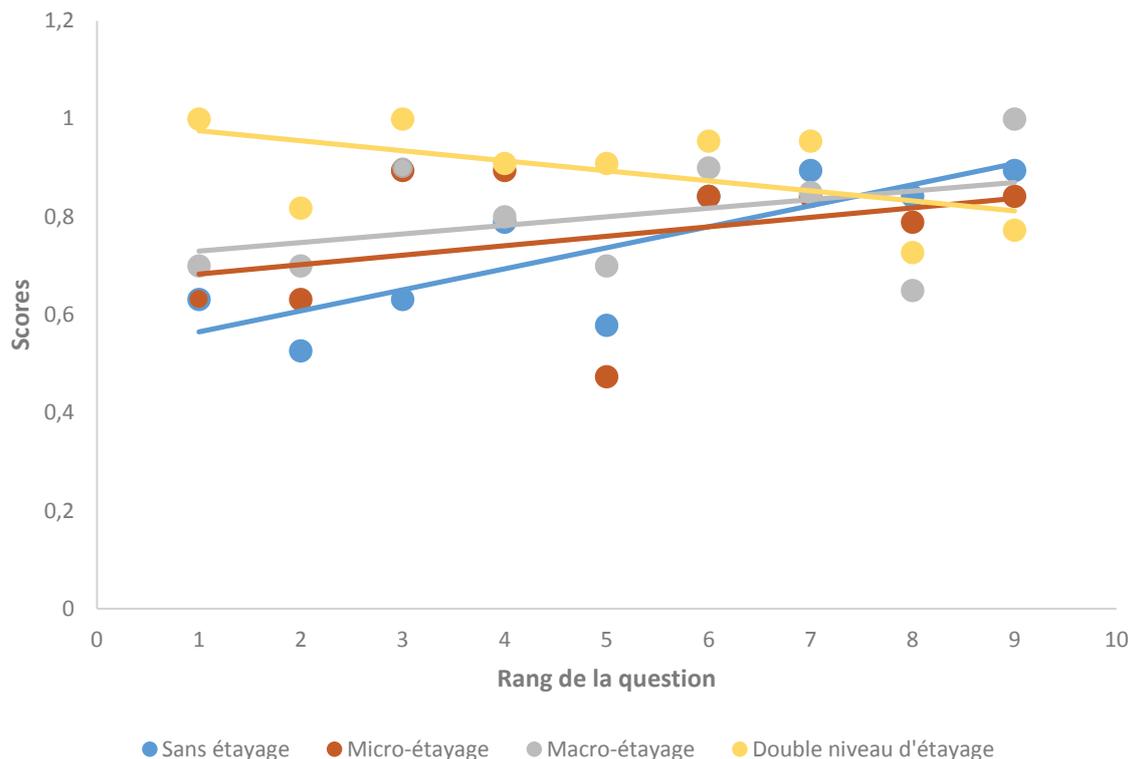


Figure 7. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 1

3.2.3. Temps de réponse

Afin d'analyser les données relatives aux temps de réponse, il a été décidé de distinguer les réponses et les échecs. Un échec est considéré comme tel lorsque le participant n'a pu trouver la réponse à une question dans les cinq minutes allouées. Un test du Chi-2 a été réalisé pour déterminer si le nombre d'échecs différait selon les conditions (voir Tableau 4 pour des données descriptives). Les résultats indiquent une différence significative, $\chi^2(3, N = 720) = 17.18, p < .001$. Les données descriptives indiquent que les participants ont eu moins d'échecs quand ils se trouvaient dans la condition double niveau d'étayage. Pour analyser les temps de réponse, les temps des questions se terminant en échec ont été retirés des données et il a été choisi de se focaliser sur les temps de réponse des recherches réussies.

Tableau 4. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 1)

Condition	Nombre d'échecs	% (condition)	% (total)
Sans étayage	22	12.87	3.06
Micro-étayage	20	11.70	2.78
Macro-étayage	15	8.33	2.08
Double niveau d'étayage	4	2.02	0.56

Les temps de réponse à chaque question ont été analysés selon les conditions expérimentales. Les résultats ont montré que le micro-étayage (M0 comparé à M1a), $\chi^2(1, N = 659) = 7.37, p = .007$, et le macro-étayage (M0 et M1b), $\chi^2(1, N = 659) = 12.57, p < .001$, ont chacun un effet sur les temps de réponse. Les résultats ont montré également un effet additif des étayages micro- et macro-niveaux (M1b et M2²¹), $\chi^2(1, N = 659) = 8.00, p = .005$, ainsi qu'un effet d'interaction entre ces deux niveaux d'étayage (M2 et M3), $\chi^2(1, N = 659) = 14.54, p < .001$. Le modèle M3 apparaît donc comme étant le meilleur modèle. Les statistiques descriptives indiquent que les temps de réponse sont plus courts dans la condition double niveau d'étayage que dans les trois autres (voir Figure 8).

D'autres analyses ont été menées pour déterminer si le temps alloué à chaque RI changeait tout au long de la tâche. Les résultats ont montré un effet additif du rang de la question (M3 comparé à M4), $\chi^2(1, N = 659) = 34.21, p < .001$, ainsi qu'un effet d'interaction entre les deux niveaux d'étayage et le rang de la question (M4 et M5), $\chi^2(3, N = 659) = 11.08, p = .011$. Le modèle M5, correspondant à une double interaction entre niveaux d'étayage et rang de la question, est finalement le plus approprié. Les statistiques descriptives indiquent que la différence entre la condition double niveau d'étayage et les trois autres conditions semble s'effacer au cours du temps (selon le rang de la question), puisque leur temps de réponse semble diminuer au cours du temps (voir Figure 8).

²¹ La comparaison des modèles M1a et M2 est également significative, $\chi^2(1, N = 659) = 13.21, p < .001$.

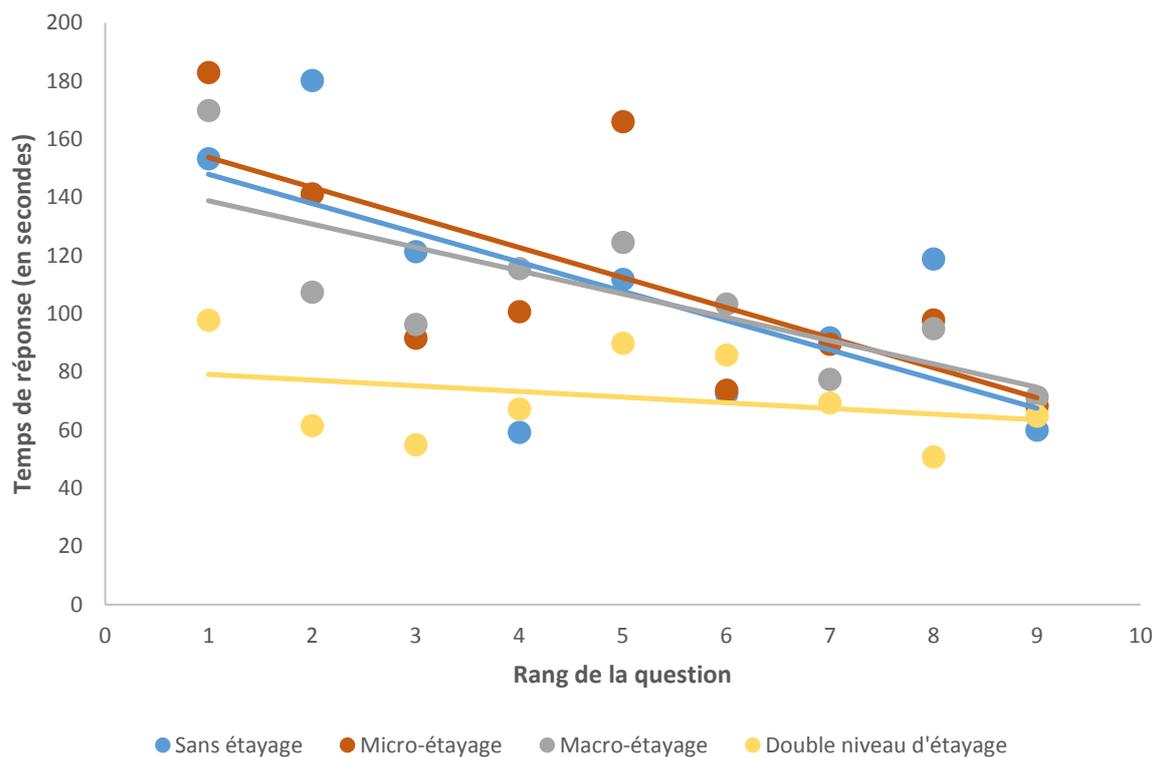


Figure 8. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 1

3.2.4. Pertinence du premier clic

Les analyses quant à la pertinence du premier clic se sont portées sur les données issues du taux d'erreur des participants lors du début de leur recherche. En effet, c'est la distance entre le segment cible (contenant la réponse attendue) et le segment cliqué par le participant qui a été mesurée. Les résultats ont montré que le micro-étayage (comparaison de M0 et M1a), $\chi^2(1, N = 720) = 5.92, p = .015$, et le macro-étayage (M0 et M1b), $\chi^2(1, N = 720) = 15.59, p < .001$, ont chacun un effet significatif sur le taux d'erreur dans la pertinence du premier clic. Les résultats ont également indiqué un effet additif des deux niveaux d'étayage (M1b et M2²²), $\chi^2(1, N = 720) = 6.55, p = .011$, ainsi qu'un effet d'interaction entre ces deux niveaux (M2 et M3), $\chi^2(1, N = 720) = 6.51, p = .011$. Le modèle M3 paraît être le plus satisfaisant. Les données descriptives indiquent que le taux d'erreur est plus bas dans

²² La comparaison des modèles M1a et M2 est également significative, $\chi^2(1, N = 720) = 16.21, p < .001$.

la condition double niveau d'étayage que dans les trois autres conditions, cela correspond donc à une pertinence plus élevée du premier clic (voir Figure 9).

Le changement au cours du temps (i.e., le changement au cours de la tâche) du taux d'erreur dans la pertinence du premier clic a également été analysé. Les résultats ont révélé un effet additif du rang de la question (M3 comparé à M4), $\chi^2(1, N = 720) = 43.64, p < .001$, ainsi qu'une interaction entre les deux niveaux d'étayage et le rang de la question (M4 et M5), $\chi^2(3, N = 720) = 14.78, p = .002$. Le modèle M5, incluant une double interaction entre les niveaux d'étayage et le rang de la question, semble donc être le modèle le plus approprié. Les statistiques descriptives indiquent que la différence entre la condition double niveau d'étayage et les trois autres conditions semble s'effacer au cours du temps (selon le rang de la question), puisque leur taux d'erreur semble diminuer au cours du temps, et de ce fait leur pertinence semble augmenter (voir Figure 9).

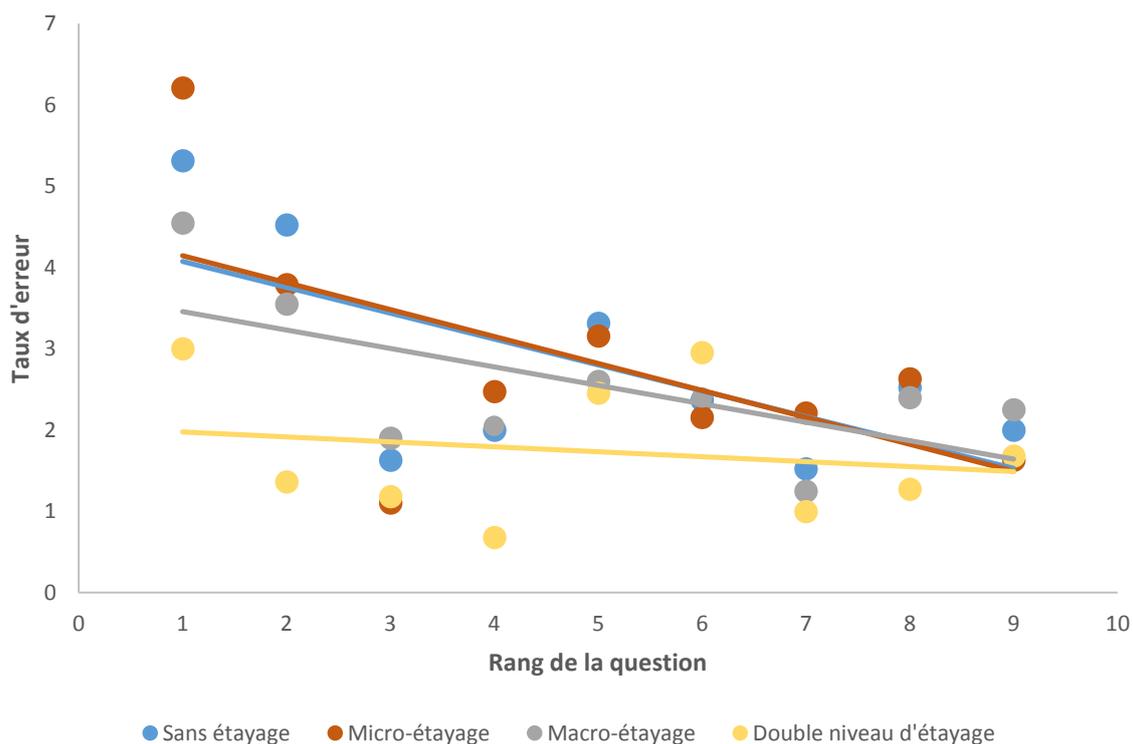


Figure 9. Taux d'erreur moyens des participants en RI sur le premier clic selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 1

3.3. Post-questionnaire

3.3.1. Difficulté perçue

Concernant la difficulté perçue, une ANOVA en 2 x 2 a révélé des effets principaux du micro-étayage, $F(1, 76) = 11.53$, $p = .001$, $\eta^2_p = .13$, et du macro-étayage, $F(1, 76) = 7.08$, $p = .010$, $\eta^2_p = .09$, ainsi qu'une interaction entre ces deux niveaux, $F(1, 76) = 5.04$, $p = .028$, $\eta^2_p = .06$. Les données (voir Tableau 5 pour les données descriptives) montrent que les participants qui ont eu une perception de la difficulté plus basse que les autres sont uniquement ceux dans la condition double niveau d'étayage.

Tableau 5. Données descriptives pour les variables de difficulté perçue, contrôle perçu et nombre de chapitres rappelés en fonction de la condition expérimentale (étude 1)

Condition	Difficulté perçue		Contrôle perçu		Nombre de chapitres rappelés		Distance de Levenshtein	
	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	8.74	2.60	13.21	3.78	4.68	1.95	7.84	1.50
Micro-étayage	8.05	3.61	16.16	2.57	5.58	2.24	7.47	1.87
Macro-étayage	8.55	3.46	14.80	3.29	6.50	2.33	6.45	2.19
Double niveau d'étayage	4.91	1.88	17.00	3.10	7.41	3.22	5.82	2.40

3.3.2. Contrôle perçu

En ce qui concerne le contrôle perçu, une ANOVA en 2 x 2 a révélé un effet principal du micro-étayage, $F(1, 76) = 12.66$, $p < .001$, $\eta^2_p = .14$, mais pas d'effet principal du macro-étayage, $F(1, 76) = 3.12$, $p = .082$, $\eta^2_p = .04$ ni d'interaction entre les deux niveaux, $F(1, 76) = 0.27$, $p = .605$, $\eta^2_p = .004$. Les données (voir Tableau 5 pour les données descriptives) montrent que les participants affirment avoir perçu davantage de contrôle sur la tâche uniquement quand ils ont bénéficié de micro-étayage.

3.3.3. Rappel des chapitres

Concernant le nombre de chapitres rappelés (qu'ils soient corrects ou non), une ANOVA en 2 x 2 a révélé un effet principal du macro-étayage, $F(1, 76) = 10.77$, $p = .002$, $\eta^2_p = .12$, mais pas d'effet principal du micro-étayage, $F(1, 76) = 2.58$, $p = .112$, $\eta^2_p = .03$ ni d'interaction entre les deux niveaux, $F(1, 76) = 0.00$, $p = .990$, $\eta^2_p = .00$. Les données (voir Tableau 5 pour les données descriptives) indiquent que les participants rappellent un plus grand nombre de chapitres uniquement quand ils ont bénéficié du macro-étayage.

Cependant, la comparaison du nombre de chapitres rappelés a semblé insuffisante. En effet, cette mesure ne prenait pas en compte la pertinence des titres rappelés, c'est-à-dire qu'un participant rappelant beaucoup de chapitres ne correspondant pas à ceux de la vidéo avait tout de même un score élevé. L'évaluation de la précision des titres, autrement dit si l'idée énoncée correspondait à celle traitée dans la vidéo, a donc semblé nécessaire. Une donnée complémentaire, appelée distance de Levenshtein (Yujian & Bo, 2007)²³, a été mesurée sur les idées énoncées dans les chapitres rappelés, et non plus sur leur simple nombre. Dans ce cas, c'est le nombre de concepts (i.e., idées) rappelés (et présents dans la vidéo) qui a été compté, et leur ordre a été pris en compte. Les résultats (ANOVA en 2 x 2) sur cette variable ont montré un effet principal du macro-étayage, $F(1, 76) = 11.38$, $p = .001$, $\eta^2_p = .13$, mais pas d'effet principal du micro-étayage, $F(1, 76) = 1.24$, $p = .270$, $\eta^2_p = .02$, ni d'interaction entre les deux niveaux, $F(1, 76) = 0.08$, $p = .774$, $\eta^2_p = .00$. Les données (voir Tableau 5 pour les données descriptives) indiquent que les participants rappellent un plus grand nombre de chapitres (i.e., ont un plus faible score d'erreur) lorsqu'ils ont bénéficié du macro-étayage.

4. Discussion

L'hypothèse principale proposée dans cette étude était la suivante : l'étayage des activités de micro- et/ou macro-niveau(x) dans un environnement vidéo aiderait les utilisateurs à élaborer une représentation mentale du document, facilitant ainsi la

²³ La distance de Levenshtein correspond au coût minimal requis pour transformer une séquence donnée en une autre, grâce à des actions telles que l'insertion, la substitution ou la suppression (Yujian & Bo, 2007).

RI. Chaque niveau d'étayage aurait notamment des effets spécifiques sur l'activité de RI. Ainsi, il était supposé que l'étayage aiderait l'efficacité et la précision de la RI, notamment en optimisant le contrôle perçu par l'utilisateur et/ou en rendant la structure du document plus saillante.

Concernant les scores de réussite à la tâche, les résultats ont montré un effet significatif du macro-étayage, mais pas d'effet du micro-étayage. Les participants bénéficiant d'un étayage de macro-niveau (i.e., table des matières à côté de la vidéo) ont eu davantage de réponses correctes que les autres participants. Aucun avantage de l'étayage de micro-niveau n'a cependant été noté, ce qui ne valide que partiellement l'**hypothèse 1**. Les résultats trouvés sur les temps de réponse selon la condition ont révélé que les deux niveaux d'étayage, combinés, ont aidé à réduire le temps alloué aux RI. Globalement, les participants bénéficiant à la fois de micro- et de macro-étayage ont passé moins de temps sur la tâche de RI que les participants des trois autres conditions expérimentales. Dans ces conditions (i.e., micro- ou macro-étayage ou sans étayage), le temps alloué à chaque recherche semble diminuer au cours de la tâche. Ces résultats sont consistants avec l'**hypothèse 2**. Le même pattern de résultats a pu être observé pour le taux d'erreur du premier clic (i.e., inverse de la pertinence) sur chaque question. Les participants ont fait moins d'erreurs (i.e., ont été plus pertinents, plus près du segment cible) quand ils ont bénéficié du double niveau d'étayage par rapport aux participants dans les trois autres conditions. Ceux-ci semblent avoir réduit leur nombre d'erreurs au fur et à mesure de l'avancée de la tâche. Ces résultats valident l'**hypothèse 3**. Les résultats observés sur ces trois variables (i.e., réussite à la tâche, temps de réponse et pertinence du premier clic) mettent en lumière les effets bénéfiques de l'étayage des documents vidéo dans une activité de RI. L'étayage de macro-niveau aiderait ainsi à améliorer la précision de la réponse, et quand les deux niveaux d'étayage (i.e., micro- et macro-niveaux) sont combinés, ils aideraient à améliorer l'efficacité et la pertinence de la RI. Ces résultats s'ajoutent à de précédents (e.g., Merkt & Schwan, 2014) en confirmant la supériorité d'un étayage de double niveau (i.e., *enhanced video*) par rapport à d'autres formats vidéo. Cette étude apporte néanmoins un nouveau regard sur l'évolution des comportements des utilisateurs au cours de la tâche. Il était supposé que l'étayage aurait des effets positifs parce qu'il fournirait un modèle conceptuel prêt à l'emploi de la vidéo. La présence d'une table des matières et de marqueurs sur la barre de

navigation rendraient ainsi plus aisés la localisation et l'accès à l'information recherchée. Dans le cas des scores de réussite à la tâche, seul l'étayage de macro-niveau semble porter de l'information sémantique. Ceci pourrait être la raison pour laquelle le macro-étayage est le seul à avoir un effet sur ce score de réponse à une question.

Par ailleurs, quand un niveau d'étayage manque, ou lorsqu'aucun étayage n'est présent, les utilisateurs doivent développer leur propre modèle mental, c'est-à-dire une représentation mentale du(des) niveau(x) d'étayage manquant(s) dans la vidéo. Cette construction pourrait être la raison des différences de performance entre les conditions expérimentales, et expliquer pourquoi ces différences s'atténuent au cours de la tâche de RI au point de disparaître à la fin. Une fois la construction de la représentation mentale élaborée, lorsque l'étayage n'était pas proposé, les individus peuvent alors l'utiliser comme leur propre étayage mental. Deux conclusions apparaissent alors : 1) lorsqu'un étayage est proposé, il servirait de modèle conceptuel, donc externe, à l'utilisateur, et serait immédiatement utilisé pour la RI, améliorant ainsi les performances ; 2) en l'absence d'étayage, les utilisateurs construiraient leur propre représentation mentale de l'environnement vidéo, au point qu'ils finiraient par avoir des performances semblables à ce qu'ils auraient eu dès le début s'ils avaient bénéficié d'étayage. Ce processus de construction semble coûteux cognitivement, comme le démontrent les résultats sur la difficulté perçue, confirmant par ailleurs l'**hypothèse 4**.

De plus, apporter un étayage de micro-niveau dans un environnement vidéo améliore significativement la sensation de contrôle sur la situation, en accord avec l'**hypothèse 5** et les recherches précédentes dans la littérature (e.g., Delen et al., 2014 ; Merkt et al., 2011). Enfin, le rappel des chapitres (nombre et concepts) est amélioré avec la présence d'étayage de macro-niveau, confirmant l'**hypothèse 6** et allant dans le sens d'études précédentes sur les indices hiérarchiques et leur effet sur le rappel du texte (e.g., Lorch et al., 2011 ; Sanchez et al., 2001). Ces résultats mettent en lumière les caractéristiques spécifiques des différents niveaux d'étayage. Alors que le micro-étayage est utilisé pour naviguer au sein de la vidéo, le macro-étayage, quant à lui, renseigne sur le contenu sémantique de l'information. L'activité de RI peut ainsi être considérée comme une activité complexe et à deux niveaux (i.e., navigation et identification sémantique). Ces deux niveaux sont complémentaires, et leur

combinaison semble avoir un effet significatif sur les performances en RI. La construction par l'utilisation d'un modèle mental pertinent peut surmonter les difficultés rencontrées en l'absence d'étayage, mais elle est bien plus coûteuse en termes de temps et de demandes cognitives. En résumé, l'étayage aide les utilisateurs à être performants dès la première tentative, une performance qu'ils n'auraient atteinte qu'après plusieurs questions sans étayage.

Cette étude comporte cependant quelques limites. Premièrement, la tâche de RI était constituée de neuf questions à suivre, ce qui peut sembler finalement peu naturel. Dans une situation classique de RI, les utilisateurs de la vidéo peuvent avoir à rechercher une information précise, voire plusieurs, mais l'enchaînement des neuf questions sans avoir visionné la vidéo en entier avant de commencer est une situation qui peut manquer vraisemblablement de caractère écologique. Deuxièmement, alors que notre principale hypothèse concerne les modèles mentaux des participants et leur construction au cours du temps, aucune variable dépendante directement liée à ce concept n'a été utilisée. Plusieurs auteurs (e.g., Azevedo et al., 2004 ; He et al., 2008 ; Marchionini, 1989) ont discuté de la manière d'avoir accès aux modèles mentaux durant une tâche sans s'appuyer uniquement sur les performances. Cette réflexion fera par ailleurs l'objet de l'étude 2.

5. Conclusion

Les résultats de cette étude indiquent que les micro- et macro-niveaux d'étayage jouent un rôle important dans l'activité de RI au sein d'une vidéo. Au sein de l'état de l'art précédemment présenté, il a été relevé que la localisation de l'information était une étape commune aux modèles théoriques dans le domaine. Les résultats de cette étude ont permis de définir deux composantes de la localisation d'information : l'identification de l'information recherchée et la navigation dans la vidéo pour la trouver. Ces deux composantes seraient soutenues respectivement par des étayages de macro- et de micro-niveaux. Concrètement, les résultats de cette étude nous permettent de faire quelques recommandations sur la conception d'environnements vidéo quand ils sont utilisés pour des activités de RI. La présence des deux niveaux d'étayage (i.e., segmentation de la barre de navigation et

structuration en chapitres) permet à l'utilisateur d'effectuer sa tâche d'une manière peu coûteuse, rapide et efficace.

Il est également supposé que l'étayage fournirait à l'utilisateur une représentation externe de la vidéo, et que sans étayage l'utilisateur doit construire un modèle mental par lui-même. En effet, la construction d'un modèle mental peut se réaliser sans modèle conceptuel (Borgman, 1986), et serait peut-être justement le moyen de pallier l'absence d'un tel soutien. Pour tester cette hypothèse, la prochaine étude se focalisera ainsi sur la manière de mesurer cette construction de modèle mental, sans se fier uniquement aux scores de performance et sans altérer la tâche de RI. Finalement, si l'activité de RI a pour conséquence la construction d'un modèle mental opérationnel, nous pouvons également supposer que cela pourrait améliorer l'apprentissage (Johnson-Laird, 1983 ; Norman, 1983).

ÉTUDE 2 : Rôle de l'étayage sur les caractéristiques de la représentation mentale de la vidéo lors d'une activité de RI²⁴

1. Introduction et hypothèses

La première étude menée a démontré un effet majeur de la présence d'un étayage de double niveau (i.e., segmentation et structuration) sur les performances en RI. Cette deuxième étude a pour but de caractériser le modèle mental des utilisateurs après une activité de RI avec ou sans étayage. Il est supposé que la présence d'étayage dans un environnement vidéo aide la RI en fournissant un modèle conceptuel externe et directement utilisable, alors que sans étayage les utilisateurs doivent construire par eux-mêmes leur propre modèle mental, ce qui est coûteux et long mais permet d'avoir une représentation interne de la vidéo. Deux sous-objectifs sont envisagés au cours de cette étude, l'un relatif à l'activité de RI en elle-même, et le second visant à évaluer la qualité du modèle mental des utilisateurs au cours de cette tâche de RI.

1) Le premier sous-objectif de cette étude est de répliquer les précédents résultats observés dans l'étude 1 sur un nouvel échantillon de participants : l'étayage aurait un effet positif sur les performances en RI dans un environnement vidéo. Deux niveaux d'étayage peuvent être distingués. Le micro-étayage renvoie à une segmentation par marqueurs dans la barre de navigation de la vidéo, afin de promouvoir le contrôle de l'utilisateur sur le flot d'informations délivrées et la navigation au sein de la vidéo. Le macro-étayage désigne la structuration de la vidéo en présentant une table des matières, afin de faciliter l'identification de l'information cible. Dans cette étude, seules

²⁴ Cette étude fait actuellement l'objet d'une soumission dans une revue internationale à comité de lecture : Cojean, S., & Jamet, E. (en révision). The role of scaffolding in improving information seeking in videos. *Journal of Computer Assisted Learning*.

deux conditions expérimentales ont été comparées : avec ou sans étayage, la condition avec étayage présentant un étayage à double niveau. Neuf questions dont toutes les réponses se trouvaient dans la vidéo ont été posées aux participants pour permettre une analyse temporelle des résultats.

Plusieurs hypothèses ont été faites :

Hypothèse 1. Réussite à la tâche :

- a) Grâce au modèle conceptuel fourni par l'étayage, les participants dans la condition avec étayage devraient mieux réussir la tâche que les participants sans étayage.
- b) Sans étayage, la performance à chaque question devrait s'améliorer au cours du temps, reflétant ainsi la construction d'un modèle mental pertinent.

Hypothèse 2. Temps de réponse :

- a) Grâce au modèle conceptuel fourni par l'étayage, les participants dans la condition avec étayage devraient passer moins de temps à la RI que les participants sans étayage.
- b) Sans étayage, le temps passé à chercher l'information à chaque question devrait diminuer au cours du temps, reflétant ainsi la construction d'un modèle mental pertinent.

Hypothèse 3. Pertinence du premier clic :

- a) Grâce au modèle conceptuel fourni par l'étayage, les participants dans la condition avec étayage devraient être plus pertinents sur leurs premiers clics (i.e., plus proches du segment cible contenant l'information recherchée) que les participants sans étayage.
- b) Sans étayage, la pertinence du premier clic à chaque question devrait s'améliorer au cours du temps, reflétant ainsi la construction d'un modèle mental pertinent.

Hypothèse 4. Difficulté perçue :

Du fait de la construction de leur modèle mental, les participants sans étayage devraient percevoir la tâche comme étant plus difficile que les participants ayant bénéficié d'étayage.

2) Le second sous-objectif de cette étude consiste à examiner la qualité du modèle mental qu'ont les utilisateurs de la vidéo après une activité de RI avec ou sans étayage. Théoriquement (Borgman, 1986 ; Norman, 1983 ; Staggers & Norcio, 1993), un modèle mental pertinent aiderait à prédire ce qui se passerait lors d'une interaction avec le système (e.g., quelle information apparaîtrait après un clic à tel endroit de la barre de navigation de la vidéo). La présence d'étayage durant une RI peut être considérée comme l'équivalent de la présence d'un modèle conceptuel sur lequel les utilisateurs peuvent se baser pour construire une représentation mentale pertinente de la vidéo. Cette représentation devrait alors faciliter la RI. Dans l'étude précédente, les participants des conditions sans étayage ou avec un étayage partiel (i.e., un seul niveau, micro- ou macro-étayage) ont eu des résultats similaires (i.e., une importante difficulté perçue et l'amélioration de leurs performances au cours de la tâche). Ces résultats observés sont en faveur de l'hypothèse de la construction d'un modèle mental durant la tâche de RI. Cette construction semble être coûteuse cognitivement, mais finalement aussi efficace que la présence d'un étayage double niveau à la fin de la tâche en termes de performance. Quand les deux niveaux d'étayage étaient présents, les utilisateurs pouvaient alors potentiellement s'y référer en tant que modèle conceptuel pertinent, et ont montré de bonnes performances en RI dès le début de la tâche. Cependant, les scores de difficulté perçue n'ont été plus bas que dans la condition où les deux niveaux d'étayage étaient présents. Si le modèle mental se construisait sur la base d'un modèle conceptuel lorsque celui-ci est fourni (ici, l'étayage), les scores de difficulté perçue auraient dû être plus faibles de manière graduelle entre les conditions sans étayage, avec un niveau d'étayage, et avec double niveau d'étayage. En effet, dans les conditions avec un seul niveau d'étayage, les participants ont eu à construire un modèle mental avec un modèle conceptuel plus complet que si l'étayage avait été totalement absent (une difficulté perçue moins importante aurait donc dû être observée). De la même manière, l'étayage double niveau, en fournissant un modèle conceptuel plus complet qu'un étayage à un seul

niveau, aurait dû faire diminuer la difficulté de la tâche perçue par les participants. Cependant, la diminution de la difficulté perçue ne s'est observée que pour la condition avec double étayage, les trois autres conditions ayant eu un score similaire. Ainsi, les résultats de l'étude précédente montrent que la construction d'un modèle mental, que l'étayage soit présent ou non, amène un même score de difficulté perçue. Dans le cadre de la présentation d'un étayage double niveau, le modèle conceptuel présenté serait suffisamment pertinent pour rendre la tâche peu coûteuse cognitivement : il est alors supposé qu'il agirait comme une représentation uniquement externe, et ne soutiendrait pas la construction d'un modèle mental interne. Dans cette nouvelle étude, une tâche de localisation a été proposée aux participants après leur tâche de RI. L'objectif était d'évaluer la qualité de leur modèle mental du contenu de la vidéo. Ce modèle mental était déterminant pour les utilisateurs s'ils voulaient pouvoir retrouver l'emplacement d'une information particulière demandée dans la vidéo, sans avoir accès à aucun étayage ni *feedback* de la vidéo.

Hypothèse 5. Localisation :

Puisque les participants ont pu construire un modèle mental de la vidéo durant la première étape de l'étude (i.e., activité de RI), les participants sans étayage devraient être plus performants dans une tâche de localisation que les participants qui ont bénéficié de l'étayage comme d'un modèle conceptuel resté externe.

2. Méthode

2.1. Participants

50 étudiants (43 femmes, 7 hommes) de l'Université Rennes 2 ont pris part à l'étude. Ils avaient en moyenne 19.31 ans (E.-T. = 1.11). Pour les remercier de leur participation, ils ont tous été gratifiés d'un ticket de cinéma. L'étude s'est déroulée conformément aux principes de la déclaration d'Helsinki.

2.2. Matériel

2.2.1. Tâche de RI

Le matériel utilisé pour cette étude est similaire à celui utilisé pour l'étude 1, à la différence qu'il a été repensé pour que l'ensemble des informations soit présenté sur un seul écran (et non plus sur un écran d'ordinateur, une tablette et des feuilles). Un nouvel environnement vidéo a spécialement été développé pour présenter la vidéo (extraite du site de Canal U²⁵). La vidéo s'intitulait *Du Temps, de l'espace et de l'eau* (Doressoundiram, 2012) et durait environ treize minutes. Comme précisé sur le site Internet, douze chapitres étaient présents dans la vidéo. Le matériel présenté (voir Figure 10) permettait à tous les participants de se déplacer au sein de la vidéo en se servant de la souris autant qu'ils le souhaitaient dans la barre de navigation de la vidéo. Un bouton pause leur permettait également d'arrêter la vidéo à n'importe quel moment. Un enregistreur d'écran²⁶, lancé pour chaque participant, a permis d'analyser plus tard les performances des utilisateurs. Les participants ont été aléatoirement et équitablement répartis entre les deux conditions expérimentales : avec ou sans étayage. Dans la **condition avec étayage** ($n = 25$), une table des matières était présentée à côté de la vidéo et la barre de navigation proposait douze marqueurs correspondant aux douze chapitres dans la vidéo. Les utilisateurs pouvaient avoir accès à une partie spécifique de la vidéo en cliquant sur le chapitre qu'ils voulaient atteindre dans la table des matières (voir Figure 10). Dans la **condition sans étayage** ($n = 25$), il n'y avait ni table des matières ni marqueur dans la barre de navigation, mais uniquement la vidéo et sa barre de navigation. La première étape de cette étude consistait en une tâche de RI similaire à l'étude 1. Les neuf questions pour la tâche de RI (e.g., « Sur quel satellite de Jupiter existerait-il un océan d'eau liquide ? » ; voir Annexe 1) étaient présentées sur l'écran, sur la gauche de la vidéo (en-dessous ou à la place de la table des matières). Un chronomètre indiquait combien de temps il restait au participant pour répondre à la question (voir Figure 10). Les participants avaient cinq minutes pour répondre à chaque question. Une fois leur réponse écrite (toujours numériquement, avec le clavier), ils pouvaient passer à la question suivante. S'ils

²⁵ <http://www.canal-u.tv/>

²⁶ Flashback by Blueberry Software, <https://www.flashbackrecorder.com/fbhome/>

n'avaient pas répondu dans les cinq minutes, la question suivante apparaissait automatiquement. Les neuf questions étaient présentées dans un ordre aléatoire afin de limiter les effets d'ordre (Benton & Daly, 1991).

The figure displays two screenshots of a video-based learning environment. Both screenshots show a question on the left and a video player on the right. The video player shows a crescent moon against a black background. The top screenshot (1) has a question: "L'eau marque la différence de formation entre deux types de planètes: lesquelles ?" and a "Question suivante" button. The bottom screenshot (2) has a question: "Pour quelle raison Vénus est-elle plus chaude que la Terre ?" and a "Question suivante" button. A "Table des matières" (Table of Contents) is visible in the bottom screenshot, listing 12 topics. The video player in the bottom screenshot has a segmented progress bar with 12 segments, and the video title is "Doressoundiram, A. (2012). Du temps, de l'espace et de l'eau. A la recherche de l'eau dans l'Univers."

Figure 10. Environnements vidéo pour la tâche de RI dans les conditions avec (1) ou sans (2) étayage de l'étude 2

2.2.2. Tâche de localisation

La seconde étape de cette étude était identique pour tous les participants, qu'ils aient été dans la condition avec étayage ou la condition sans étayage dans la première partie (i.e., tâche de RI). À la place de la vidéo, un écran noir était présenté, avec en-dessous une barre de navigation similaire à celle de la vidéo (voir Figure 11). En

revanche, en cliquant sur cette barre de navigation, aucun *feedback* n'était délivré (puisque la vidéo n'était plus accessible). Il n'y avait ni table des matières ni marqueur sur la barre de navigation. Des questions (e.g., « Qu'est-ce que la ligne des glaces ? » ; voir Annexe 2) étaient présentées sur le côté gauche de l'écran, et les utilisateurs avaient cinq minutes pour y répondre. Un chronomètre était affiché sous la question. Les participants devaient répondre à huit questions (quatre questions avaient déjà été posées durant la première étape, quatre autres étaient de nouvelles questions). Ils devaient indiquer en cliquant sur la barre de navigation factice où ils pensaient que la réponse se trouvait dans la vidéo (voir Figure 11). Pour les participants qui avaient été pendant la première partie dans la condition sans étayage, la seule différence ici était que la vidéo n'était plus accessible (ni le son, ni l'image). Pour les participants qui avaient passé la première étape dans la condition avec étayage, ils ne pouvaient alors plus bénéficier ni de la vidéo, ni de l'étayage.



Figure 11. Environnement vidéo pour l'étape de localisation de l'étude 2

2.3. Mesures

2.3.1. Intérêt préalable et compétence perçue

Avant de commencer la tâche de RI, les participants devaient répondre à deux questions : une à propos de leur intérêt préalable quant au sujet développé dans vidéo (i.e., l'eau dans l'Univers), et une sur leur compétence perçue quant à ce sujet. L'objectif était de vérifier que les participants étaient équitablement répartis dans les conditions sur ces deux variables. Une échelle de Likert en 11 points a été utilisées

pour mesurer le degré auquel les participants étaient intéressés (« Sur une échelle de 0 à 10, à quel point ce sujet vous intéresse-t-il ? ») et à quel point ils se sentaient compétents (« Sur une échelle de 0 à 10, à quel point vous estimez-vous compétent sur ce sujet ? »).

2.3.2. Réussite à la tâche

La réussite à la tâche a également été évaluée. Pour que la réponse soit considérée comme correcte, les participants devaient écrire l'information recherchée (i.e., la réponse à la question posée) et indiquer à quel moment elle apparaît dans la vidéo (i.e., noter le temps en minutes et secondes du moment où la réponse est explicitement présentée dans la vidéo). Lorsqu'au moins une de ces deux informations était manquante ou erronée, la réponse était considérée comme incorrecte.

2.3.3. Temps de réponse

Les temps de réponse ont été calculés à partir du moment où les participants commençaient à chercher l'information (i.e., à l'apparition de la question), jusqu'au moment où ils trouvaient l'information recherchée (i.e., action sur le bouton pause ou début de l'écriture de la réponse). Les participants avaient cinq minutes par question pour trouver la réponse. L'enregistrement vidéo de l'écran a été utilisé pour analyser chaque RI.

2.3.4. Pertinence du premier clic

La pertinence du premier clic a été évaluée. Pour cela, l'enregistrement vidéo de l'écran a été utilisé pour prendre en note à quel endroit sur la barre de navigation les participants ont réalisé chacun de leurs premiers clics (un premier clic par question). Il y avait douze segments sur la barre de navigation (correspondant aux douze chapitres de la vidéo, affichés dans la condition avec étayage mais invisibles dans la condition sans étayage), il a donc été calculé le nombre de segments entre celui où a eu lieu le premier clic et celui contenant l'information cible (i.e., où se trouvait la réponse recherchée). Le score ainsi obtenu était un taux d'erreur (i.e., distance entre le premier clic et l'information cible), et un faible score était donc synonyme d'une forte pertinence.

2.3.5. Difficulté perçue

Trois questions sur la difficulté perçue (e.g., « Pour moi, effectuer cette recherche d'information était difficile » ; voir Annexe 5) ont été adaptées d'études précédentes (e.g., Kraft et al., 2005 ; Trafimow et al., 2002). Les participants répondaient sur des échelles de Likert en 7 points pour indiquer le degré d'accord avec les phrases proposées. Un alpha de Cronbach a été mesuré pour cette variable, il était de $\alpha = .88$.

2.3.6. Précision de la localisation

Dans cette étude, de précédents travaux réalisés par Marchionini (1989) ont inspiré l'évaluation de la qualité du modèle mental grâce à une tâche de localisation. Il était ainsi demandé aux participants de prédire ce qui se passerait dans une situation spécifique sans possibilité d'avoir un *feedback* du système. En d'autres termes, huit questions similaires à celles posées lors de la tâche de RI étaient présentées une à une. Pour chaque question, le participant devait cliquer sur la barre de navigation factice à l'endroit où il pensait que se trouvait la réponse attendue. Ainsi, pour chacune des huit questions, un score a été calculé. Il s'agissait de la distance, en secondes, entre la réponse donnée par le participant et la localisation correcte de l'information recherchée sur la barre de navigation de la vidéo.

2.4. Procédure

Jusqu'à trois participants pouvaient être accueillis en même temps au cours de la passation. Chacun était installé sur un bureau isolé des autres, avec un casque pour écouter la vidéo lors de la tâche. L'expérimentateur leur proposait d'abord un pré-questionnaire concernant leur intérêt préalable et leur compétence perçue sur le sujet de la vidéo qu'ils verraient plus tard (i.e., l'eau dans l'Univers). Puis l'expérimentateur décrivait aux participants comment l'étude allait se dérouler, donnait la consigne de la RI dans la vidéo (voir Annexe 8), et listait toutes les mesures qui seraient faites durant les différentes tâches. Une fois l'enregistreur d'écran démarré, les participants pouvaient commencer quand ils le souhaitaient et cliquer pour faire apparaître la première question de RI. Un chronomètre sur l'écran indiquait le temps restant pour trouver la réponse. Quand le participant avait répondu aux neuf questions, il pouvait

avoir accès à un post-questionnaire sur la difficulté perçue durant la tâche de RI. Ensuite, l'expérimentateur donnait aux participants la consigne pour la seconde étape, celle de localisation (voir Annexe 9), et les laissait commencer quand ils le souhaitaient. Pour chacune des huit questions, les participants avaient cinq minutes pour y répondre, et pouvaient passer à la question suivante une fois qu'ils avaient répondu. La procédure est résumée en Figure 12. Après cette seconde partie de l'étude, chaque participant recevait un ticket de cinéma pour le remercier de sa participation.

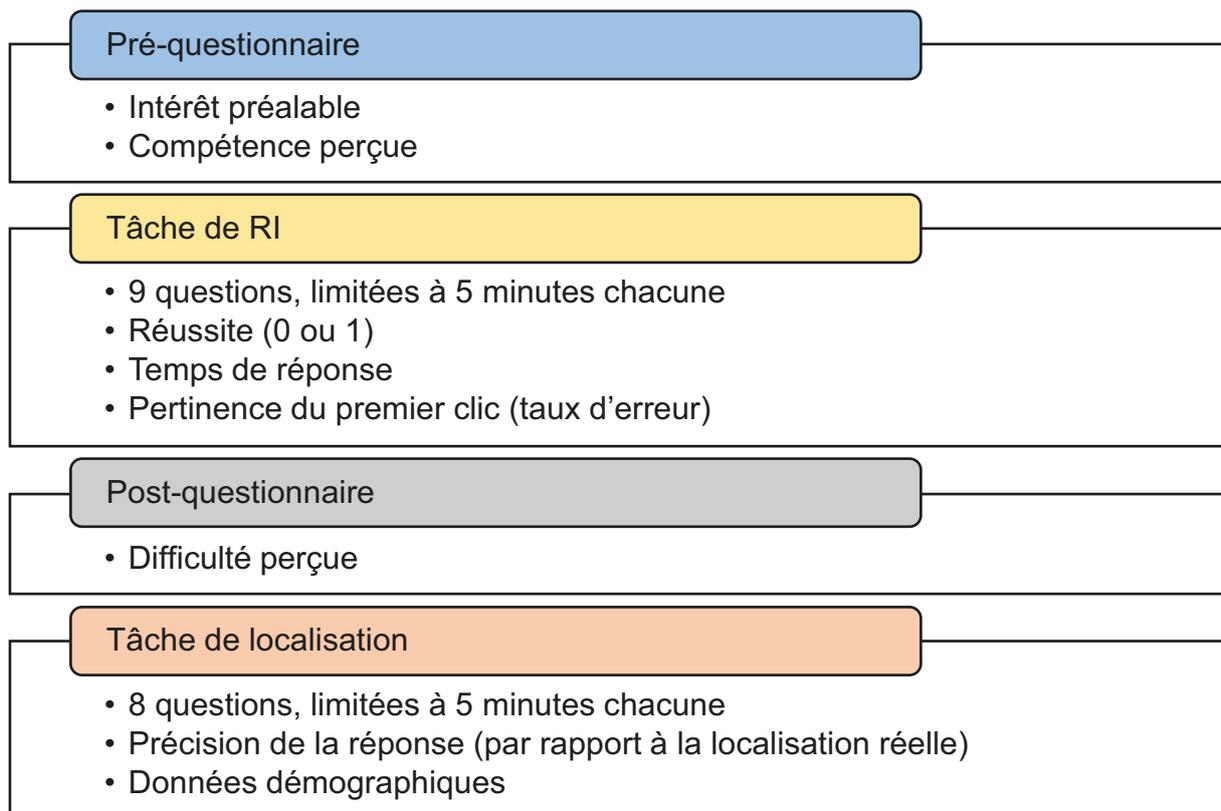


Figure 12. Procédure de l'étude 2

3. Résultats

3.1. Prérequis pour l'analyse statistique

Les questions posées durant le pré-questionnaire ont servi de variables contrôles, afin de s'assurer que les scores des participants des deux conditions expérimentales (avec ou sans étayage durant la première étape) ne différaient pas sur leur intérêt préalable et leur compétence perçue à propos du sujet de la vidéo. Des

analyses de variance (ANOVA) n'ont révélé aucune différence significative entre les conditions sur les variables intérêt préalable ($F(1, 48) = 0.05, p = .819$) et compétence perçue ($F(1, 48) = 1.67, p = .202$). Les données descriptives sont représentées dans le Tableau 6.

Tableau 6. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 2)

Condition	Intérêt préalable		Compétence perçue	
	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	5.76	1.92	2.40	1.63
Avec étayage	5.88	1.76	3.04	1.86

Durant l'activité de RI, les participants devaient répondre à neuf questions. La réussite à la tâche, le temps de réponse et la pertinence du premier clic ont été mesurés pour chaque participant et pour chacune des neuf questions. L'intérêt de ces multiples mesures était d'évaluer l'évolution de la performance au cours du temps. De plus, durant la seconde étape, chaque participant devait répondre à huit questions de localisation. Pour ces raisons, puisque toutes les questions étaient présentées à tous les participants, l'indépendance des données n'est plus justifiable (voir aussi étude 1 ; Field et al., 2012). Afin de prendre en compte la non-indépendance des données, des modèles linéaires mixtes (Gueorguieva & Krystal, 2004) ont été utilisés. Cette méthode statistique permet de tester l'effet d'une variable par comparaison de modèles imbriqués tout en tenant compte du caractère répété des mesures (i.e., un sans la variable et un avec la variable ; e.g., Baayen et al., 2008). Un test de Chi-2 est estimé entre les deux modèles imbriqués. Le seuil de significativité pour les valeurs p a été fixé à $\alpha = .05$. Chaque modèle incluait les effets aléatoires de la question et du participant.

3.2. Tâche de RI

Pour chaque variable dépendante où les modèles linéaires mixtes ont été utilisés (i.e., réussite à la tâche, temps de réponse et pertinence du premier clic), l'objectif était de tester les effets de la condition expérimentale (i.e., étayage) et du rang de la question (i.e., évolution au cours du temps). Ainsi, un modèle nul (M0) a été

comparé à un modèle incluant la variable condition (M1). Le modèle le plus satisfaisant (M0 ou M1) a ensuite été comparé à un modèle incluant un effet additif du rang de la question (M2) et le modèle le plus satisfaisant entre M0, M1 et M2, a finalement été comparé à un modèle incluant un effet d'interaction du rang de la question (M3).

3.2.1. Réussite à la tâche

Afin d'évaluer la réussite à la tâche, les réponses étaient codées soit 0 (réponse incorrecte) soit 1 (réponse correcte). Puisque ces données sont de nature binomiale, des régressions logistiques ont été utilisées (Field et al., 2012), alors que pour les variables temps de réponse et pertinence du premier clic, des régressions linéaires ont été utilisées. Les résultats ont montré un effet significatif de la condition sur la réussite à la tâche, $\chi^2(1, N = 450) = 11.61, p < .001$, mais pas d'effet additif du rang de la question, $\chi^2(1, N = 450) = .33, p = .568$, ni d'effet d'interaction du rang de la question avec la condition, $\chi^2(1, N = 450) = .05, p = .833$. Le modèle M1 a donc été considéré comme étant le plus approprié. Les statistiques descriptives montrent que les participants semblent réussir mieux la tâche quand ils bénéficient d'étayage (voir Figure 13).

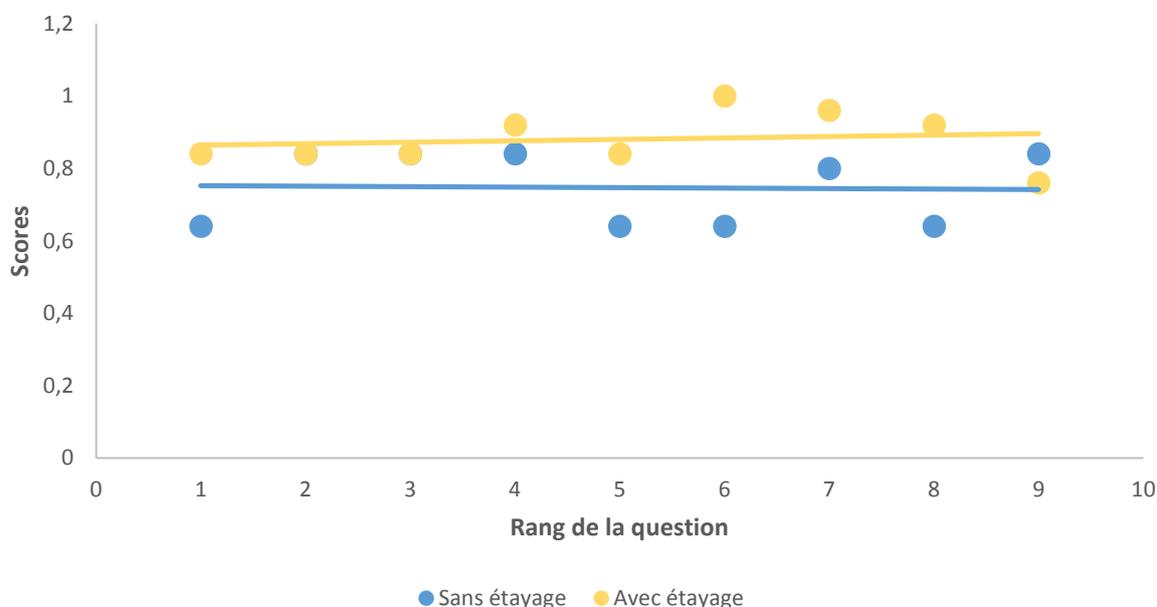


Figure 13. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 2

3.2.2. Temps de réponse

Les temps de réponse à chaque question ont également été étudiés. Comme pour l'étude 1, les échecs ont été ôtés des données pour se concentrer sur les recherches ayant abouti sur une réponse. Un échec est considéré comme tel lorsqu'aucune réponse n'a été apportée à la question posée dans les cinq minutes allouées à la RI par question. Un test du Chi-2 a été réalisé pour déterminer si le nombre d'échecs différait selon les conditions (voir Tableau 7 pour des données descriptives). Les résultats n'indiquent pas de différence significative entre les conditions, $\chi^2(1, N = 450) = 2.71, p = .10$.

Tableau 7. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 2)

Condition	Nombre d'échecs	% (condition)	% (total)
Sans étayage	17	7.56	3.78
Avec étayage	8	3.56	1.78

Les résultats ont montré un effet significatif de la condition sur les temps de réponse, $\chi^2(1, N = 425) = 32.14, p < .001$, ainsi qu'un effet additif, $\chi^2(1, N = 425) = 10.66, p = .001$, et un effet d'interaction du rang de la question avec la condition, $\chi^2(1, N = 425) = 6.52, p = .011$. Le modèle M3 a donc été considéré comme étant le plus approprié. Les statistiques descriptives (voir Figure 14) montrent que les participants qui n'ont pas bénéficié de l'étayage ont passé davantage de temps sur chaque recherche que les participants qui ont bénéficié de l'étayage. De plus, la différence entre les deux conditions expérimentales semble se réduire au cours de la tâche.

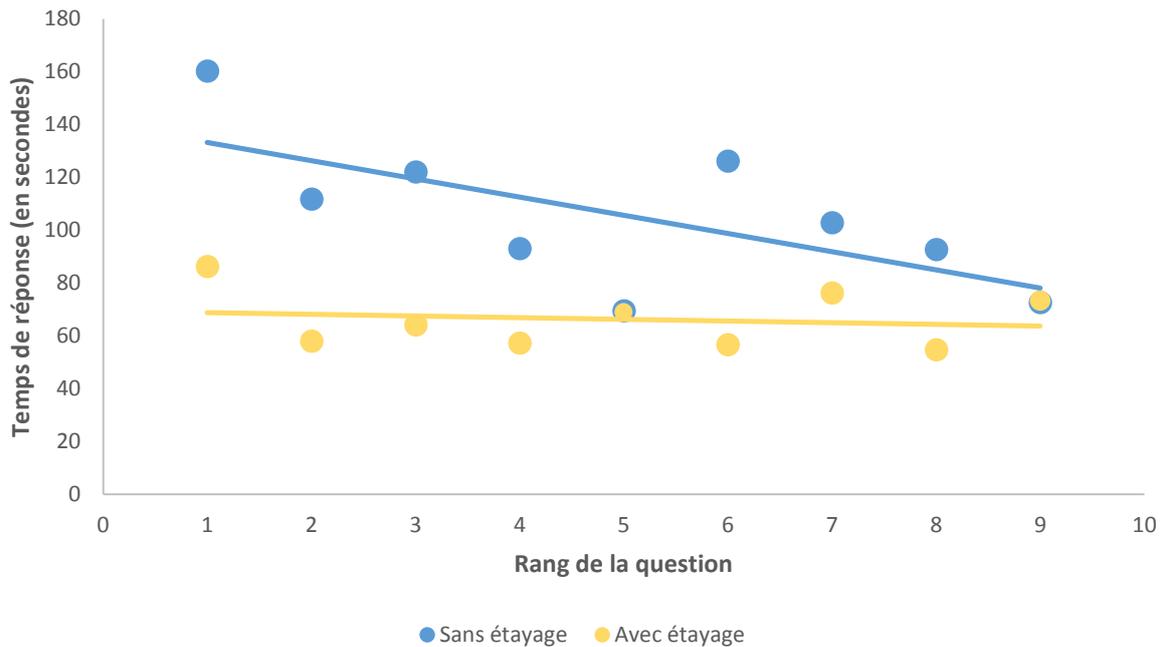


Figure 14. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 2

3.2.3. Pertinence du premier clic

En ce qui concerne la pertinence du premier clic, les analyses ont été réalisées sur les taux d'erreur (i.e., nombre de segments entre celui contenant le premier clic de la recherche et le segment cible contenant la réponse attendue) plutôt que sur un score (inverse) de pertinence. Les résultats ont montré un effet significatif de la condition sur la distance entre les deux segments mesurés, $\chi^2(1, N = 450) = 29.81, p < .001$, ainsi qu'un effet additif, $\chi^2(1, N = 450) = 21.87, p < .001$, et un effet d'interaction du rang de la question avec la condition, $\chi^2(1, N = 450) = 5.10, p = .024$. Le modèle M3 a donc été considéré comme étant le modèle le plus approprié. Les données descriptives montrent que les participants ne bénéficiant pas d'étayage ont fait de plus grandes erreurs sur chaque premier clic que les participants bénéficiant d'étayage. De plus, la différence entre les deux conditions semble se réduire au cours du temps (voir Figure 15).

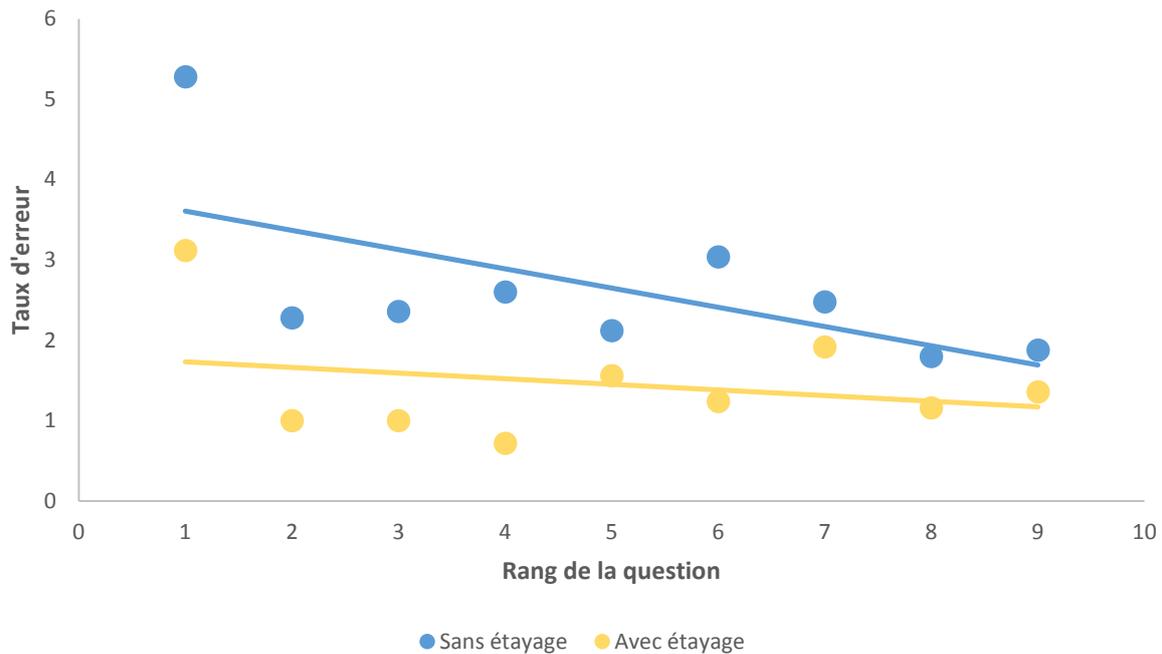


Figure 15. Taux d'erreur moyens des participants en RI sur le premier clic selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 2

3.2.4. Difficulté perçue

Pour la variable de difficulté perçue, une ANOVA a révélé un effet significatif de la condition $F(1, 48) = 4.74$, $p = .034$, $\eta^2_p = .09$ (voir Tableau 8 pour les statistiques descriptives). Les données montrent que les participants qui n'ont pas bénéficié de l'étayage ont perçu davantage de difficulté dans la réalisation de la tâche de RI au sein d'une vidéo que les participants qui ont bénéficié d'étayage.

Tableau 8. Données descriptives pour la variable de difficulté perçue en fonction de la condition expérimentale (étude 2)

Condition	Moy.	E.-T.
Sans étayage	3.19	1.14
Avec étayage	2.53	0.98

3.3. Tâche de localisation

Le score de localisation a été calculé en tant que distance, en secondes, entre la réponse donnée et la localisation de la réponse attendue. Il s'agit donc plutôt d'un

score de distance (ou marge d'erreur) que d'un score de précision. Durant la seconde étape, l'effet supposé du rang de la question, mesuré dans la première étape, n'a plus de raison d'être testé. En effet, l'absence de *feedback* de la vidéo empêche la possibilité d'un apprentissage²⁷. Un modèle nul (M0) a été comparé à un modèle incluant la variable condition (M1). Les résultats ont montré un effet significatif de la condition sur la localisation de la réponse, $\chi^2(1, N = 400) = 5.85, p = .016$. Le modèle M1 a donc été considéré comme étant le plus approprié. Les statistiques descriptives (voir Tableau 9) indiquent que les participants semblent être plus précis dans la localisation (i.e., ils ont fait de plus petites erreurs) quand ils n'ont pas bénéficié d'étayage durant l'étape précédente, par rapport aux participants ayant bénéficié d'étayage.

Tableau 9. Données descriptives pour le score (marge d'erreur) de localisation en fonction de la condition expérimentale (étude 2)

Condition	Moy.	E.-T.
Sans étayage	103.27	33.14
Avec étayage	129.62	37.68

Enfin, l'effet de la condition sur le score de localisation pourrait s'interpréter de manière à penser que si les participants sans étayage ont été plus précis, cela était surtout dû au temps supplémentaire passé sur la vidéo lors de l'étape de RI. Une corrélation a donc été calculée, entre la durée totale d'exposition à la vidéo (i.e., durée totale des temps de réponse durant la première étape) et le score de localisation moyen durant la seconde étape. Le résultat a montré une corrélation significative, mais considérée comme faible, $r = -.30, p = .034$. Cependant, le graphique en nuage de points (voir Figure 16) correspondant semble indiquer qu'il n'existe pas de lien entre le temps de visionnage de la vidéo et le score à la tâche de localisation.

²⁷ Les résultats ont effectivement montré une absence d'effet significatif additif ($\chi^2(1, N = 400) = 0.119, p = .731$) ou d'interaction ($\chi^2(1, N = 400) = 0.071, p = .790$) du rang de la question avec la condition expérimentale.

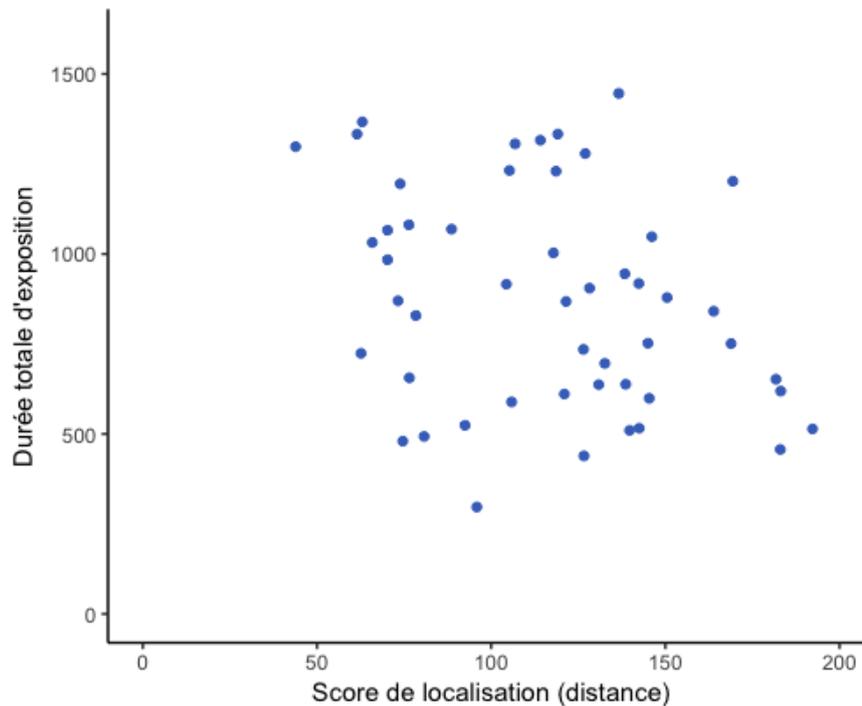


Figure 16. Durée totale d'exposition (en secondes) en fonction du score de localisation (marge d'erreur) pour l'étude 2

4. Discussion

La qualité des représentations mentales et les performances en RI sont fortement liées (He et al., 2008 ; Sharit et al., 2008 ; X. Zhang & Chignell, 2001). En effet, un modèle mental adapté a la particularité de permettre la prédiction de ce qui va se passer au sein du système lors de prochaines interactions (Borgman, 1986 ; Norman, 1983 ; Staggers & Norcio, 1993). L'objectif de cette deuxième étude était de répliquer de précédents résultats sur l'effet positif de l'étaillage lors d'une activité de RI dans un environnement vidéo, et de s'intéresser aux caractéristiques des modèles mentaux des utilisateurs dans ce contexte. La principale hypothèse était que l'étaillage allait aider l'activité de RI, mais ne permettrait pas l'internalisation d'un modèle mental à partir du modèle conceptuel fourni. Sans étaillage, il a été supposé qu'aurait lieu de la part de l'utilisateur une construction longue et cognitivement coûteuse d'un modèle mental interne, et que les performances en RI augmenteraient au fur et à mesure des questions.

Concernant la première étape de l'étude (i.e., tâche de RI), les résultats confirment l'**hypothèse 1** selon laquelle il existe un effet de la condition expérimentale sur le score de réussite à la tâche. Les participants ayant bénéficié d'étayage ont trouvé davantage de réponses correctes dans la vidéo que les participants n'ayant pas bénéficié d'étayage. Cependant, l'absence d'effet d'interaction avec le rang de la question indique que la performance des participants sans étayage est restée en-deçà de celle des participants avec étayage, et ce jusqu'à la fin de la tâche. Pour les temps de réponse, les résultats indiquent que les participants dans la condition sans étayage ont passé davantage de temps à rechercher les informations demandées que les participants profitant de l'étayage, mais cette différence entre les conditions tendait à s'amenuiser au cours du temps. Ces résultats valident l'**hypothèse 2**. La présence d'étayage a également aidé les participants à être plus pertinents durant leurs premiers clics, et cette différence entre conditions tend aussi à disparaître au cours du temps, ce qui est en accord avec l'**hypothèse 3**. À l'issue de la tâche, les participants l'ayant réalisée avec étayage ont reporté de manière significative avoir perçu moins de difficulté que les participants sans étayage, et ce résultat confirme l'**hypothèse 4**. Tous ces résultats sont consistants avec les précédentes études menées sur la RI dans les vidéos, affirmant un effet positif de l'étayage (e.g., Merkt & Schwan, 2014) et une évolution au cours du temps de la performance des utilisateurs qui n'ont pas bénéficié de micro- et/ou macro-étayage (voir étude 1).

L'hypothèse d'une représentation mentale non-intériorisée lors de la présence d'étayage complet (micro- et macro-niveaux) et de la construction d'un modèle mental interne sans ce double niveau a été testée dans la seconde partie de l'étude. Les résultats sur la variable de difficulté perçue dans la première partie sont cohérents avec cette hypothèse. Les résultats de la tâche de localisation ont par ailleurs montré que les participants ayant réalisé la RI (i.e., première étape) sans étayage ont été plus précis (i.e., ils ont fait de plus petites erreurs) que les participants qui avaient pu bénéficier d'étayage. L'**hypothèse 5** est ainsi validée. Ces résultats viennent à l'encontre de la théorie voulant que la présence d'un modèle conceptuel soutienne la construction d'un modèle mental et en soit la base (e.g., Norman, 1983). Dans cette étude, fournir un modèle conceptuel par le biais de l'étayage durant une activité de RI n'a pas favorisé la construction d'un modèle mental pertinent de la vidéo. Lorsque l'étayage a en effet disparu, les participants se sont trouvés moins capables de

localiser une information sans *feedback* de la vidéo que les participants qui n'avaient jamais pu bénéficier de cet étayage auparavant. Ce résultat peut être mis en lien avec l'aspect de la définition du modèle mental, selon lequel un modèle mental pertinent pouvait servir à prédire le comportement futur du système (i.e., de l'environnement vidéo et de son contenu). Savoir quelle information sera présentée en cliquant à tel endroit pourrait renvoyer à une prédiction du comportement de la vidéo. Ce résultat peut également s'interpréter en termes de « difficulté désirable » : lorsque l'étayage est absent, les utilisateurs doivent fournir de plus grands efforts pour parvenir à la même performance que s'ils avaient pu en bénéficier. Ce pourrait être la raison pour laquelle, à la fin de la tâche, ils avaient un meilleur modèle mental du contenu de la vidéo (E. Bjork, Little, & Storm, 2014 ; R. Bjork, 1994 ; Dobson, 2011).

Une limite à cette interprétation pourrait être que le meilleur score à la tâche de localisation par les participants sans étayage durant la première partie serait dû au temps passé à regarder la vidéo. En effet, ce sont également eux (i.e., les participants sans étayage) qui ont passé davantage de temps à naviguer dans la vidéo pour la RI en étape 1 (voir Figure 14). Afin de répondre à cette hypothèse, une corrélation a été calculée entre la durée totale d'exposition au contenu de la vidéo durant la première partie et le score moyen de localisation durant la seconde partie. Il apparaît que cette corrélation est significative, mais faible, et que les données qualitatives (voir Figure 16) ne penchent pas en faveur d'un lien entre ces deux variables. Le temps passé sur la vidéo ne peut donc être mis en lien direct avec la qualité du modèle mental qui s'ensuit, et ce résultat va dans le sens de notre hypothèse principale.

Enfin, il convient de noter que, d'un point de vue pratique et écologique, il n'est pas commun de demander à des étudiants d'effectuer neuf RI au sein de la même vidéo. Ainsi, la possibilité pour eux de construire un modèle mental pertinent de la vidéo est probablement rare dans des conditions classiques, où l'interaction avec le contenu d'une vidéo peut être beaucoup plus courte.

5. Conclusion

Les environnements vidéo, lorsqu'ils sont utilisés pour une tâche de RI, sont plus efficaces s'ils fournissent de l'étayage approprié à l'utilisateur. Cet étayage prend la forme de deux sous-niveaux, correspondant par exemple à une structuration en

chapitres dans une table des matières et une segmentation de la barre de navigation par des marqueurs. Étayer une vidéo permet donc une RI plus rapide, plus précise et moins difficile pour l'utilisateur. Cependant, cette étude montre également que l'étayage fourni reste un modèle conceptuel externe, et que l'utilisateur n'en profite pas pour construire un modèle mental de la vidéo par lui-même. De futures études pourraient être menées sur les effets de l'étayage lors d'une tâche non plus de RI, mais d'apprentissage. En effet, les environnements vidéo sont de plus en plus utilisés dans ce contexte. Les activités de RI et d'apprentissage sont fortement liées, notamment parce que le traitement de l'information en vue d'un apprentissage commence généralement par la localisation de l'information pertinente (Merkt & Schwan, 2014 ; Rouet & Coutelet, 2008). De plus, Merkt et Schwan (2014) ont montré qu'une information facile à trouver est plus aisément rappelée lors de tests d'apprentissage. La présence d'étayage lors d'une tâche de RI aiderait l'utilisateur à améliorer ses performances : en effet, ceux qui bénéficient de l'étayage ont dès la première question la même efficacité qu'ont ceux sans étayage à la neuvième question. Il peut donc être supposé que l'étayage devrait avoir un effet positif sur l'apprentissage. Cependant, cette seconde étude amène à considérer que les utilisateurs, lorsqu'ils utilisent l'étayage pour améliorer leur performance de RI, ne construisent pas de modèle mental de la vidéo aussi pertinent que s'ils n'avaient pas eu d'étayage. De plus, la littérature met justement en avant le fait que l'apprentissage se base sur la qualité des modèles mentaux (e.g., Hegarty, 2014 ; Mayer et al., 2002). L'idée serait alors de se concentrer sur les raisons pour lesquelles l'étayage (i.e., le modèle conceptuel proposé) dans les environnements vidéo n'est pas pertinemment internalisé durant une activité de RI, et sur comment cet étayage pourrait favoriser l'apprentissage.

Selon Mayer (2004), pour acquérir de manière cohérente et organisée des connaissances, les étudiants doivent être engagés dans un processus actif d'apprentissage. Il peut être ainsi supposé que lorsqu'aucune consigne d'apprentissage n'est donnée, ou lorsqu'il n'y a pas de nécessité de construire un modèle mental pertinent (pour réussir une tâche de RI avec étayage par exemple), il n'y aura pas de traitement actif de l'étayage. Lorsque le but d'apprentissage est clairement formulé, l'étayage devrait alors être internalisé et avoir un effet positif sur les performances en apprentissage. Cette hypothèse fera l'objet de l'étude 4.

ÉTUDE 3 : Comment réduire l'effet d'attention divisée pendant la RI au sein d'une vidéo ?

1. Introduction et hypothèses

Les études précédentes de cette thèse se sont attachées à investiguer l'effet de l'étayage en vidéo, et plus particulièrement de l'étayage à double niveau, sur les performances en RI (voir études 1 et 2). Néanmoins, l'ajout d'étayage, et plus encore des deux niveaux conjoints d'étayage, multiplie les sources d'information à disposition des utilisateurs. Ainsi, cette troisième étude vise à évaluer et diminuer les effets négatifs qu'un tel format peut induire en termes d'attention divisée pour l'utilisateur.

L'effet d'attention divisée (Ayres & Sweller, 2014) apparaît lorsque différentes sources d'information sont présentées simultanément à l'utilisateur, et que ces sources ne peuvent être comprises isolément mais qu'elles sont nécessaires à la compréhension globale du document. L'utilisateur doit alors intégrer mentalement le contenu des sources d'information afin qu'elles fassent sens (Ayres & Sweller, 2014 ; Bétrancourt & Bisseret, 1998 ; Erhel & Jamet, 2006 ; Sorden, 2005). Dans le cas par exemple de l'apprentissage d'un document présentant des animations et de l'information textuelle, l'effet spatial d'attention divisée peut apparaître lorsque les apprenants doivent intégrer mentalement les informations de l'image animée avec celles du texte présenté simultanément à l'écran. Un effet temporel d'attention divisée peut également avoir lieu quand les apprenants doivent intégrer mentalement l'information présentée avec la précédente qui n'est plus disponible visuellement ou auditivement. Maintenir l'information provenant d'une première source en mémoire de travail pour l'intégrer avec celle provenant d'une seconde source est l'élément le plus coûteux dans l'effet d'attention divisée. L'utilisateur doit ainsi chercher et faire correspondre des éléments liés entre eux mais éloignés dans le document (Ayres & Sweller, 2014 ; Bétrancourt & Bisseret, 1998 ; Ginns, 2006 ; Sweller et al., 1998). Cette opération d'intégration mentale est particulièrement coûteuse lorsqu'elle a lieu dans

un document dont les éléments sont fortement liés entre eux. En effet, s'il existe une faible interactivité des éléments, un manque d'intégration pourra être potentiellement sans conséquence (Ayres & Sweller, 2014 ; Ginns, 2006).

L'effet d'attention divisée implique par conséquent une charge extrinsèque (i.e., directement liée au format de présentation du document et non aux caractéristiques intrinsèques du contenu) plus élevée, ce qui peut influencer négativement les performances lors d'un apprentissage par exemple (Ayres & Sweller, 2014). Afin de prévenir l'apparition d'un tel effet néfaste, il est recommandé dans la littérature (e.g., Ayres & Sweller, 2014 ; Erhel & Jamet, 2006) de structurer le document de manière à intégrer les différentes sources d'information plutôt que de les laisser séparées. L'intérêt est d'éviter à l'utilisateur d'effectuer cette intégration par lui-même, et donc d'éviter une surcharge cognitive non-nécessaire (Florax & Ploetzner, 2010). Le principe de conception invoqué dans ce cas précis est le principe de contiguïté spatiale. Selon ce principe, un format de document dans lequel les différentes sources d'information sont présentées de manière intégrée plutôt que séparée aidera à l'apprentissage du document (Ayres & Sweller, 2014 ; Florax & Ploetzner, 2010). Il s'agira de l'appliquer entre les deux niveaux d'étayage, parce qu'il est nécessaire de pouvoir relier les informations de l'un avec celles de l'autre. La contiguïté spatiale est à distinguer de la contiguïté temporelle, bien que parfois les deux soient regroupées sous le terme de principe de contiguïté. Le principe de contiguïté temporelle renvoie à un apprentissage favorisé lorsque les informations visuelles et orales sont synchronisées temporellement, c'est-à-dire présentées de manière simultanée et non successive (Moreno & Mayer, 1999).

Afficher les sources d'informations utiles à la compréhension du document en respectant le principe de contiguïté spatiale encouragerait les apprenants à s'engager dans des processus cognitifs d'intégration de manière plus pertinente, ce qui aurait pour conséquence un apprentissage plus efficace (Johnson & Mayer, 2012). Pour rappel, dans le modèle de l'apprentissage présenté par Mayer (2014a ; voir Introduction théorique 2.1.2), l'intégration est un processus clé lors d'une activité d'apprentissage. L'intégration physique des sources d'information au sein du document agirait alors comme un substitut à l'intégration mentale engagée par

l'apprenant en mémoire de travail, et qui est coûteuse cognitivement (Ayres & Sweller, 2014 ; Sweller et al., 1998). Dans un document au format intégré, l'apprenant n'a plus besoin de diviser son attention entre les sources d'information, ce qui lui permet de garder des ressources en mémoire de travail pour traiter l'apprentissage en lui-même (Florax & Ploetzner, 2010).

Mayer et Fiorella (2014) ont cherché à résumer les articles de la littérature présentant un effet positif de la présentation intégrée sur les performances d'apprentissage. Il en ressort que pour 22 expériences menées, les tailles d'effet (d de Cohen²⁸) varient entre 0.22 et 2.06, avec un effet médian de 1.10. L'effet de la contiguïté spatiale paraît ainsi apporter globalement un bénéfice important par rapport à un document dont les sources d'information sont éloignées. Plus largement, une méta-analyse conduite par Ginns (2006) a permis de faire la synthèse de 50 études, dont 37 étudiaient plus spécifiquement un effet de contiguïté spatiale. L'hypothèse d'une efficacité pédagogique de l'intégration spatiale des éléments liés entre eux, notamment lorsque le document est complexe, est soutenue. L'analyse des mouvements oculaires peut fournir des informations supplémentaires sur les stratégies d'apprentissage employées dans différents formats de présentation des documents. Johnson et Mayer (2012) ont effectué une telle analyse dans trois expériences comparant des documents au format intégré ou bien séparé. L'*eye-tracking* sert alors d'outil pour tester le principe de contiguïté spatiale en analysant objectivement les comportements d'exploration visuelle : les personnes confrontées à un format intégré devraient pouvoir effectuer davantage de transitions entre le texte et le diagramme, notamment parce qu'elles seraient plus engagées dans les processus tels que l'intégration. Les résultats des expériences indiquent une supériorité du format intégré par rapport au format séparé pour les scores de compréhension (expériences 1 et 2), mais jamais pour les performances en mémorisation. L'analyse des mouvements oculaires a fait apparaître la mise en place de davantage de transitions entre texte et diagramme associé lorsque le document était présenté dans un format intégré plutôt que séparé (expériences 1, 2 et 3).

²⁸ Il semble admis qu'un effet (mesuré à l'aide d'un d de Cohen) est considéré comme faible à 0.20, moyen à 0.50 et fort à partir de 0.80 (voir notamment Cohen, 1988).

La question de l'effet d'attention divisée entre l'étayage et la vidéo, néfaste pour les performances, a été traitée au cours des précédentes études de cette thèse (voir études 1 et 2) en proposant l'étayage visible en permanence à côté de la vidéo. Mais l'effet d'attention divisée ne se limite pas à l'interaction entre l'étayage et la vidéo. Il peut également survenir à cause des sous-titres de la vidéo, par exemple, parce qu'ils ajouteraient une source d'information supplémentaire à traiter (van der Zee et al., 2017). Concernant les deux niveaux d'étayage possibles et bénéfiques à la RI, l'hypothèse est alors faite qu'un effet d'attention divisé peut apparaître entre les deux niveaux d'étayage possibles (i.e., micro- et macro-étayage). En effet, les informations qu'ils véhiculent doivent potentiellement être intégrées par les individus les unes aux autres pour améliorer leurs performances. Cette troisième étude vise à identifier de quelle manière l'effet d'attention divisée entre les deux niveaux d'étayage lors d'une tâche de RI dans un environnement vidéo peut être évité, ou a minima réduit. Ainsi, il peut sembler nécessaire de s'interroger sur comment intégrer physiquement, en termes de présentation, les deux types d'étayage proposés pour éviter que les utilisateurs ne fassent les liens référentiels entre ces deux sources par eux-mêmes. L'objectif est donc de rendre les liens entre les deux niveaux d'étayage plus **explicites** pour l'utilisateur. Il s'agirait en fait de montrer quels éléments sont liés entre eux et de quelle manière, permettant à l'apprenant d'intégrer plus facilement l'information issue des différentes sources (van Gog, 2014). De précédentes études, conduites sur des tâches d'apprentissage au sein d'environnements textuels et illustrés, laissent à penser que la présence de liens explicites entre les différentes sources d'informations qui sont liées entre elles aide à la construction d'un modèle mental pertinent du document (e.g., Erhel & Jamet, 2006 ; Jamet et al., 2008).

Dans le cas des documents multimédias, une caractéristique spécifique et inhérente à ce format peut permettre d'éviter un effet d'attention divisée en facilitant l'établissement de liens référentiels : l'interactivité. L'interactivité peut se définir comme la possibilité qu'a l'utilisateur, généralement par le biais d'un clic au sein du document, d'influencer les informations auxquelles il a accès (Evans & Gibbons, 2007 ; Scheiter, 2014 ; Schwan & Riempff, 2004). L'utilisateur peut ainsi influencer l'apparition des informations en termes de temporalité (i.e., contrôle du rythme) ou de contenu (Evans

& Gibbons, 2007). Insérer de l'interactivité entre les deux niveaux d'étayage pourrait faire apparaître les liens référentiels qui existent entre eux. L'interactivité pourrait donc aider à l'appropriation du modèle conceptuel (i.e., étayage proposé) de la vidéo par l'utilisateur. Ainsi, l'utilisateur pourrait traiter simultanément en mémoire de travail les différents éléments devant interagir entre eux pour acquérir tout leur sens (Sweller, 2010). La possibilité d'avoir accès directement à une partie spécifique de la vidéo (i.e., utilisation de la segmentation, étayage de micro-niveau), et d'identifier de quelle partie il s'agit au sein de la vidéo dans son ensemble (i.e., utilisation de la structuration, étayage de macro-niveau), devrait donc permettre à l'utilisateur d'avoir de meilleures performances en RI. Cette troisième étude s'attachera à comparer différents formats de présentation des niveaux d'étayage afin d'améliorer l'utilisation de l'outil en facilitant la co-référenciation de la segmentation et de la structuration. Trois conditions seront donc comparées : 1) une condition contrôle sans étayage, 2) une condition avec étayage non-interactif, et 3) une condition avec étayage interactif, c'est-à-dire que l'utilisateur pourra cliquer sur les chapitres présentés dans la table des matières (i.e., macro-étayage) pour accéder directement au moment correspondant dans la vidéo (i.e., micro-étayage).

D'autres études dans la littérature ont été menées quant à une autre manière de réduire l'effet d'attention divisée dans les documents multimédias : l'utilisation de *pop-ups*. D'après leur définition, les informations présentées à travers des *pop-ups* sont liées à des parties spécifiques du contenu du document et apparaissent uniquement après une action de la part de l'utilisateur (Bétrancourt & Bisseret, 1998). Des études précédentes ont montré les bénéfices de la présentation avec des *pop-ups* sur les performances d'apprentissage (Bétrancourt & Bisseret, 1998 ; Erhel & Jamet, 2006). Ce format de présentation, dit intégré, devrait faciliter le traitement de l'information en montrant de manière explicite les connexions entre les différentes parties du document liées entre elles. Cela devrait également permettre d'éviter une éventuelle surcharge cognitive, parce qu'alors seulement quelques parties de l'information sont disponibles à la fois (Erhel & Jamet, 2006). Une quatrième condition expérimentale sera donc comparée aux autres dans cette troisième étude : 4) une condition avec étayage au format *pop-up*, c'est-à-dire que les chapitres de la vidéo (i.e., macro-étayage) seront présentés dans des *pop-ups* accessibles depuis l'endroit correspondant sur la barre de navigation (i.e., micro-étayage).

Cependant, dans l'étude de Merkt et Schwan (2014), l'absence visuelle de la table des matières a potentiellement mené à l'absence d'effet de ce macro-étayage. La première étude de cette thèse, durant laquelle la table des matières était physiquement présente dans l'environnement vidéo tout au long de la tâche de RI, a permis de faire apparaître des effets significatifs du macro-étayage. Une cinquième condition expérimentale sera comparée aux autres dans cette étude : 5) une condition avec étayage au format *pop-up* et avec présence permanente de la table des matières à l'écran et non plus seulement dans les *pop-ups*.

En résumé, cette troisième étude vise à investiguer dans quelle étendue il faut favoriser à la fois les liens référentiels entre les niveaux d'étayage, mais également l'accès visuel permanent à l'information (tout en évitant la surcharge visuelle) lors d'une tâche de RI dans un environnement vidéo (et non pas lors d'une tâche d'apprentissage comme dans les études précédentes de la littérature). L'objectif de cette étude est de déterminer quel format interactif de présentation vidéo est le plus approprié pour améliorer les performances en RI. Le projet LIMAH dans lequel s'inscrit cette thèse est un projet appliqué qui vise à développer un outil facilitant les tâches de RI : il est donc nécessaire d'améliorer l'utilisation de l'environnement vidéo en proposant un format adapté, dont la conception prendra également en compte la présentation optimisée des deux niveaux d'étayage. Un ensemble d'hypothèses a été formulé :

Hypothèse 1. Réussite à la tâche :

- a) Les participants dans les conditions présentant de l'étayage de micro- et macro-niveaux dans la vidéo devraient mieux réussir la tâche de RI que les participants dans la condition sans étayage.
- b) Les participants bénéficiant d'interactivité entre les deux niveaux d'étayage devraient mieux réussir la tâche que les participants sans interactivité.
- c) Les participants effectuant la tâche dans un environnement doté de *pop-ups* au niveau de l'étayage devraient mieux réussir la tâche que les participants sans *pop-up*.
- d) Les participants dans la condition avec *pop-ups* et présentation visuelle permanente des deux niveaux d'étayage devraient mieux réussir la tâche que les participants dans la condition uniquement *pop-up*.

- e) En résumé, le schéma de résultats attendu est le suivant : condition sans étayage < condition avec étayage non-interactif < condition avec étayage interactif < condition *pop-up* < condition *pop-up* et présence visuelle permanente.
- f) Les participants dans la condition sans étayage devraient voir leurs scores de réussite s'améliorer au cours de la tâche, c'est-à-dire au fur et à mesure des questions de RI, alors que les scores de réussite des participants des conditions avec double étayage devraient rester stables tout au long de la tâche (voir étude 1).

Hypothèse 2. Temps de réponse :

- a) Les participants dans les conditions présentant de l'étayage de micro- et macro-niveaux dans la vidéo devraient avoir des temps de réponse plus rapides lors de la tâche de RI que les participants dans la condition sans étayage.
- b) Les participants bénéficiant d'interactivité entre les deux niveaux d'étayage devraient avoir des temps de réponse plus rapides que les participants sans interactivité.
- c) Les participants effectuant la tâche dans un environnement doté de *pop-ups* au niveau de l'étayage devraient avoir des temps de réponse plus rapides que les participants sans *pop-up*.
- d) Les participants dans la condition avec *pop-ups* et présentation visuelle permanente des deux niveaux d'étayage devraient avoir des temps de réponse plus rapides que les participants dans la condition uniquement *pop-up*.
- e) En résumé, le schéma de résultats attendu est le suivant : condition sans étayage > condition avec étayage non-interactif > condition avec étayage interactif > condition *pop-up* > condition *pop-up* et présence visuelle permanente.
- f) Les participants dans la condition sans étayage devraient voir leurs temps de réponse diminuer au cours de la tâche, c'est-à-dire au fur et à mesure des questions de RI, alors que les temps de réponse des participants des conditions avec double étayage devraient rester stables tout au long de la tâche (voir études 1 et 2).

Hypothèse 3. Pertinence du premier clic :

- a) Les participants dans les conditions présentant de l'étayage de micro- et macro-niveaux dans la vidéo devraient avoir des premiers clics plus pertinents (i.e., plus près du segment cible contenant la réponse attendue) lors de la tâche de RI que les participants dans la condition sans étayage.
- b) Les participants bénéficiant d'interactivité entre les deux niveaux d'étayage devraient avoir des premiers clics plus pertinents que les participants sans interactivité.
- c) Les participants effectuant la tâche dans un environnement doté de *pop-ups* au niveau de l'étayage devraient avoir des premiers clics plus pertinents que les participants sans *pop-up*.
- d) Les participants dans la condition avec *pop-ups* et présentation visuelle permanente des deux niveaux d'étayage devraient avoir des premiers clics plus pertinents que les participants dans la condition uniquement *pop-up*.
- e) En résumé, le schéma de résultats attendu est le suivant : condition sans étayage < condition avec étayage non-interactif < condition avec étayage interactif < condition *pop-up* < condition *pop-up* et présence visuelle permanente.
- f) Les participants dans la condition sans étayage devraient voir leur score de pertinence s'améliorer au cours de la tâche, c'est-à-dire au fur et à mesure des questions de RI, alors que les scores de pertinence des participants des conditions avec double étayage devraient rester stables tout au long de la tâche (voir études 1 et 2).

Hypothèse 4. Difficulté perçue :

- a) Les participants dans les conditions présentant de l'étayage de micro- et macro-niveaux dans la vidéo devraient percevoir la tâche de RI comme étant moins difficile que les participants dans la condition sans étayage.
- b) Les participants bénéficiant d'interactivité entre les deux niveaux d'étayage devraient percevoir la tâche comme étant moins difficile que les participants sans interactivité.

- c) Les participants effectuant la tâche dans un environnement doté de *pop-ups* au niveau de l'étayage devraient percevoir la tâche comme étant moins difficile que les participants sans *pop-up*.
- d) Les participants dans la condition avec *pop-ups* et présentation visuelle permanente des deux niveaux d'étayage devraient percevoir la tâche comme étant moins difficile que les participants dans la condition uniquement *pop-up*.
- e) En résumé, le schéma de résultats attendu est le suivant : condition sans étayage > condition avec étayage non-interactif > condition avec étayage interactif > condition *pop-up* > condition *pop-up* et présence visuelle permanente.

2. Méthode

2.1. Participants

100 participants (79 femmes, 21 hommes), étudiants dans la région de Rennes, ont pris part à l'étude de manière volontaire. En moyenne, ils avaient 20.62 ans (E.-T. = 2.31). Les participants ont été recrutés au sein de l'université Rennes 2, notamment par des affiches postées dans les lieux de vie ou de passage (halls, couloirs, cafétérias...). Pour les remercier de leur participation, ils ont tous été gratifiés d'un ticket de cinéma. L'étude s'est déroulée conformément aux principes de la déclaration d'Helsinki.

2.2. Matériel

Le matériel utilisé lors de cette troisième étude implique la même vidéo que celle présentée dans le cadre des études 1 et 2 (Doressoundiram, 2012). Sa durée était d'environ treize minutes, elle se découpait en douze chapitres. Afin de pouvoir présenter la vidéo avec des marqueurs sur la barre de navigation et des chapitres sur le côté, ainsi qu'ajouter de l'interactivité ou encore insérer des *pop-ups*, un environnement vidéo spécifique a été conçu pour présenter aux participants la vidéo selon leur condition expérimentale sur un écran d'ordinateur. Les différentes conditions expérimentales pouvaient présenter ou non un double-étayage, mais également de l'interactivité entre structuration et segmentation (i.e., clic possible dans la table des

matières pour un accès direct au segment correspondant), ou encore une présentation de la structuration au format *pop-up* intégrée dans la segmentation (voir Figure 17).

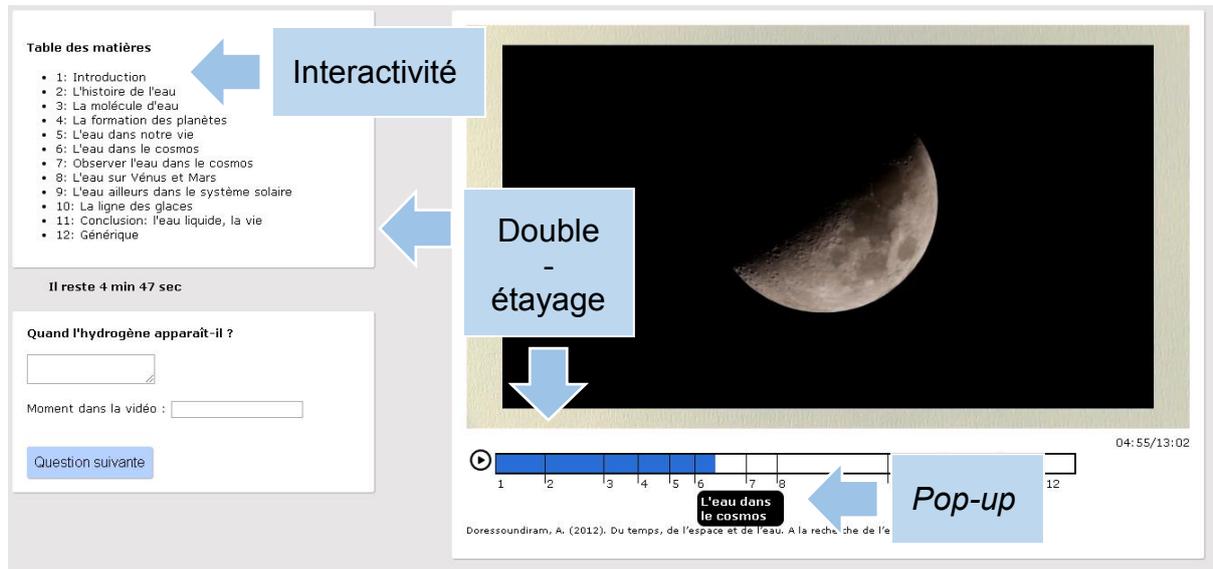


Figure 17. Possibilités d'interaction et d'intégration des deux niveaux d'étayage dans l'environnement vidéo (étude 3)

Une barre de navigation sous la vidéo permettait aux participants de naviguer dans l'ensemble de la vidéo comme ils le souhaitaient. Un bouton pause leur permettait également de stopper la vidéo à n'importe quel moment. Cinq conditions expérimentales ont été comparées dans cette étude : 1) **condition sans étayage** ($n = 21$) : l'environnement vidéo ne présentait pas d'étayage (voir Figure 18) ; 2) **condition avec étayage non-interactif** ($n = 21$) : la vidéo bénéficiait d'étayage de micro- et macro-niveaux (i.e., segmentation de la barre de navigation et table des matières sur le côté), mais sans interactivité, c'est-à-dire que les clics sur la table des matières étaient impossibles et ne renvoyaient donc pas à la partie correspondante dans la barre de navigation (voir Figure 18) ; 3) **condition avec étayage interactif** ($n = 19$) : la vidéo bénéficiait d'un étayage sur deux niveaux, ainsi que d'interactivité entre ces niveaux, c'est-à-dire qu'un clic sur un chapitre précis de la table des matières amenait l'utilisateur à l'endroit correspondant sur la barre de navigation (voir Figure 18) ; 4) **condition pop-up** ($n = 20$) : la vidéo bénéficiait d'un étayage sur deux niveaux, mais le macro-étayage (i.e., table des matières) était présenté uniquement sous la forme de *pop-ups*, accessibles lorsque l'utilisateur passait la souris sur la barre de navigation, c'est-à-dire que chaque nom de chapitre apparaissait

lorsque le chapitre en question était survolé par la souris (voir Figure 18) ; 5) **condition *pop-up* avec présence visuelle permanente** ($n = 19$) : la vidéo bénéficiait d'un étayage sur deux niveaux, et la table des matières était présente à la fois de manière permanente sur le côté de la vidéo et sous la forme de *pop-ups* accessibles lorsqu'un chapitre était survolé par la souris (voir Figure 18).

Pour tous les participants, un enregistreur d'écran²⁹ a été lancé afin de permettre plus tard d'analyser les performances des utilisateurs. Reprenant le même principe de présentation que pour l'étude 2, les neuf questions pour la tâche de RI (e.g., « Sur quel satellite de Jupiter existerait-il un océan d'eau liquide ? » ; voir Annexe 1) étaient présentées à l'écran, à gauche de la vidéo (en-dessous ou à la place de la table des matières). Chaque participant était informé du temps qu'il lui restait pour répondre à chaque question (cinq minutes maximum) grâce à un décompte placé au-dessus de la question. Une fois la réponse du participant écrite dans la case correspondante, il pouvait passer à la question suivante. Si le participant n'avait pas répondu dans les cinq minutes, la question suivante apparaissait automatiquement. Les neuf questions étaient présentées dans un ordre aléatoire afin d'éviter les effets d'ordre (Benton & Daly, 1991).

²⁹ Flashback by Blueberry Software, <https://www.flashbackrecorder.com/fbhome/>

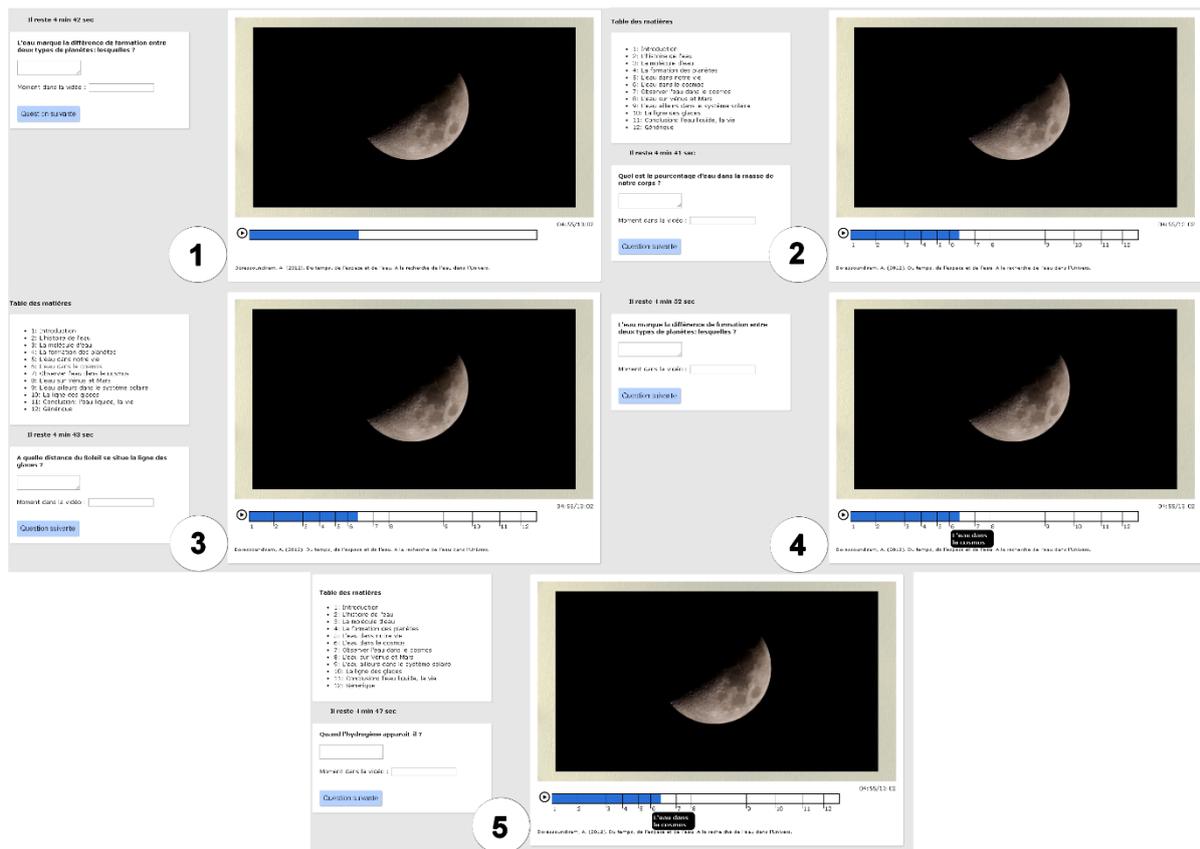


Figure 18. Environnements vidéo dans les conditions expérimentales sans étayage (1), avec étayage non-interactif (2), avec étayage interactif (3), avec pop-ups (4) et avec pop-ups et présence visuelle (5) de l'étude 3

2.3. Mesures

2.3.1. Intérêt préalable et compétence perçue

Avant d'effectuer la tâche de RI, il était demandé aux participants de répondre à deux questions. Il s'agissait de vérifier que leur intérêt préalable quant au sujet développé dans la vidéo (i.e., l'eau dans l'Univers) et leur compétence perçue sur ce sujet étaient équitablement distribués parmi les cinq conditions expérimentales. Les participants devaient indiquer sur une échelle de Likert en 11 points leur degré d'intérêt (« Sur une échelle de 0 à 10, à quel point ce sujet vous intéresse-t-il ? ») et à quel niveau ils situeraient leur compétence sur ce sujet (« Sur une échelle de 0 à 10, à quel point vous estimez-vous compétent sur ce sujet ? »).

2.3.2. Réussite à la tâche

La réussite à la tâche a été évaluée. Ainsi, une réponse était considérée comme correcte quand le participant avait indiqué la bonne réponse à la question posée et le moment dans la vidéo où elle apparaissait. Une réponse était considérée comme incorrecte lorsque la réponse notée par le participant était incorrecte, lorsque le temps indiqué dans la vidéo était incorrect, ou lorsqu'il manquait au moins une de ces deux informations. Lorsque le participant ne trouvait pas la réponse en cinq minutes, la question passait à la suivante, et la réponse était considérée comme incorrecte.

2.3.3. Temps de réponse

Grâce à l'enregistreur d'écran, les temps de réponse ont été calculés à partir du moment où le participant commençait la recherche (i.e., à l'apparition de la question) jusqu'au moment où il trouvait l'information recherchée (i.e., bouton pause activé ou début d'écriture de la réponse). Pour rappel, le temps de réponse était limité à cinq minutes maximum.

2.3.4. Pertinence du premier clic

L'utilisation de l'enregistreur d'écran permettait également de voir dans quel segment (i.e., chapitre) avait lieu le premier clic des participants à chaque question (i.e., RI). La pertinence était calculée selon le nombre de segments entre le segment contenant le premier clic et le segment contenant la réponse attendue. Un score faible (i.e., taux d'erreur) signifie donc que le participant était proche de la réponse attendue, il est donc synonyme de grande pertinence.

2.3.5. Difficulté perçue

Pour mesurer la difficulté perçue, des items ont été adaptés de précédentes études sur ces concepts (Kraft et al., 2005 ; Trafimow et al., 2002). Il s'agit des items présents dans les précédentes études. Trois items de difficulté perçue ont été utilisés (e.g., « Pour moi, effectuer cette recherche d'information était difficile » ; voir Annexe 5). Les participants devaient indiquer sur une échelle de Likert en 7 points (de

1 à 7) le degré avec lequel ils étaient d'accord avec l'affirmation proposée. L'alpha de Cronbach correspondant à cette dimension était de $\alpha = .87$.

2.4. Procédure

Les participants pouvaient être reçus jusqu'à trois en même temps. Ils étaient accueillis puis installés chacun sur un bureau isolé des autres. Un formulaire de consentement a été réalisé pour les informer des conditions de l'étude à partir de cette troisième étude. L'ensemble des mesures effectuées durant l'expérimentation a été décrit, il leur a aussi été précisé qu'ils pouvaient se retirer de l'étude à tout moment. Il leur était demandé de signer ce formulaire en deux exemplaires, d'en remettre un à l'expérimentateur et d'en garder un vers eux. Ensuite, un pré-questionnaire composé de deux questions leur était présenté. Il s'agissait pour le participant d'indiquer son intérêt préalable envers le sujet de la vidéo qui serait présenté lors de l'expérimentation (« L'Eau dans l'Univers »), ainsi que le niveau de compétence qu'il pensait avoir sur ce sujet. L'expérimentateur donnait ensuite la consigne de l'étude (répondre à des questions en cherchant la réponse dans la vidéo ; voir Annexe 10) et décrivait le matériel utilisé (description de l'étayage et des moyens d'interaction ou d'apparition de l'information quand ils étaient présents). L'expérimentateur démarrait ensuite l'enregistreur d'écran, et les participants mettaient un casque audio pour écouter la vidéo. Ils démarraient la tâche quand ils le souhaitaient, en cliquant sur un bouton pour faire apparaître la première question et donc le premier chronomètre (de cinq minutes). Lorsqu'ils avaient trouvé la réponse à la question et l'endroit dans la vidéo où elle était prononcée, les participants notaient ces informations dans les cases correspondantes à l'écran et pouvaient alors passer à la question suivante. Dans le cas où ils ne trouvaient pas la réponse en cinq minutes, la question disparaissait pour laisser place à la suivante et à un nouveau chronomètre de cinq minutes. Il n'était pas possible durant l'entièreté de la tâche de revenir à une question précédente, répondue ou non. Neuf questions se succédaient, après quoi les participants étaient invités à quitter l'environnement vidéo et à appeler l'expérimentateur. L'étape suivante était un questionnaire de difficulté perçue, toujours affiché sur l'écran d'ordinateur. Deux dernières questions demandaient le sexe et l'âge du participant. La procédure est résumée en Figure 19. A la fin de ce post-questionnaire, l'expérimentateur gratifiait les

participants d'un ticket de cinéma chacun pour les remercier de leur participation à l'étude.

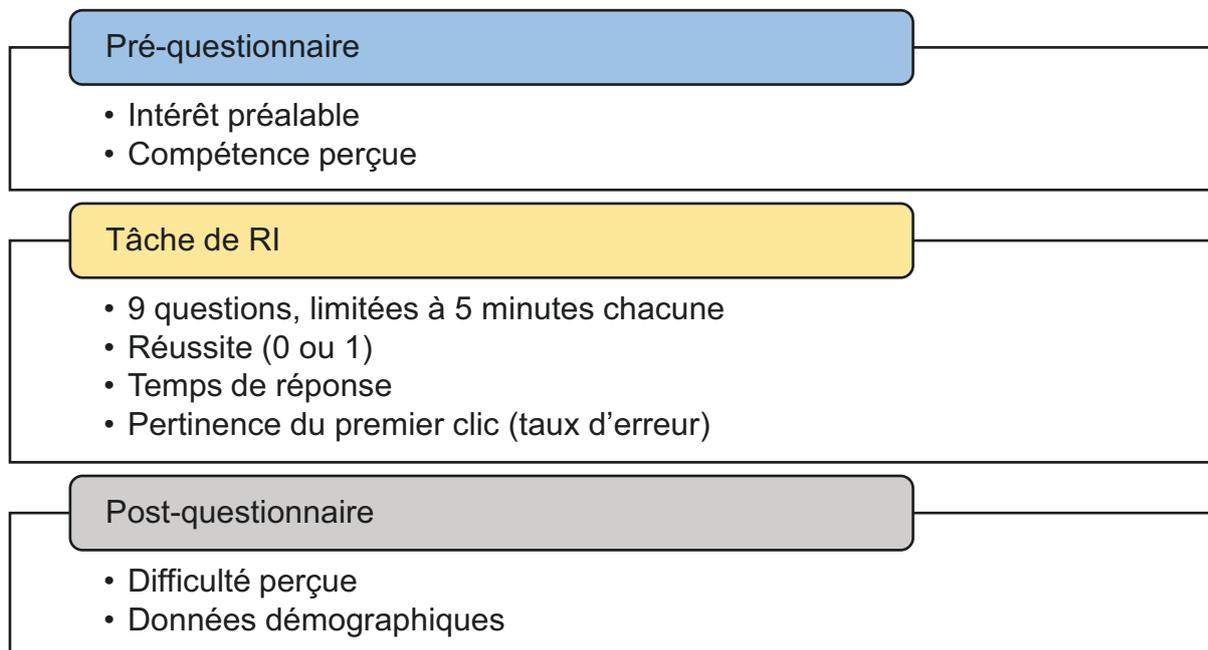


Figure 19. Procédure de l'étude 3

3. Résultats

3.1. Prérequis pour l'analyse statistique

De la même manière que pour les traitements statistiques des données issues des études 1 et 2, des modèles linéaires mixtes ont été utilisés. En effet, lors de la tâche de RI, il était demandé aux participants de répondre à neuf questions. Pour chacune de ces questions, la réussite à la tâche, le temps de réponse, ainsi que la pertinence du premier clic ont été mesurés. L'évolution de la performance au cours des questions, et donc au cours du temps, pouvait ainsi être analysée. Il apparaît donc que, puisque toutes les questions étaient présentées à tous les participants, l'indépendance des données n'est plus justifiable (voir aussi études 1 et 2 ; Field et al., 2012). Les modèles linéaires mixtes permettent de prendre en compte cette non-indépendance des données et de tenir compte de l'aspect répété de la mesure (Gueorguieva & Krystal, 2004). Cette méthode statistique permet de tester l'effet d'une variable par comparaison de modèles imbriqués (i.e., un sans la variable et un avec la variable ; e.g., Baayen et al., 2008). Un test de Chi-2 est mesuré entre les deux

modèles imbriqués. Le seuil de significativité pour les valeurs p a été fixé à $\alpha = .05$. Chaque modèle incluait les effets aléatoires de la question et du participant.

Dans cette étude, cinq conditions étaient comparées entre elles. L'effet de la condition a donc été testé, et s'il apparaît comme significatif cela signifie que les performances des participants se distinguent selon la condition (i.e., selon la présentation de l'environnement vidéo). L'interprétation subséquente des graphiques de données descriptives indique alors quelle(s) condition(s) se distingue(nt) des autres, et dans quel sens. Pour chaque variable dépendante, un modèle nul (M0) a été comparé à un modèle prenant en compte l'effet de la variable condition (M1). Le modèle le plus satisfaisant (entre M0 et M1) a ensuite été comparé à un modèle incluant un effet additif du rang de la question (M2), et le modèle le plus satisfaisant (M0, M1 ou M2) a finalement été comparé à un modèle incluant un effet d'interaction du rang de la question (M3).

3.2. Pré-questionnaire

Des ANOVA ont été réalisées sur les items d'intérêt préalable et de compétence perçue, proposés avant le début de la RI et l'exposition à la vidéo. Il s'agissait de s'assurer que les participants étaient répartis équitablement parmi les cinq conditions expérimentales, c'est-à-dire qu'ils avaient un intérêt préalable et une compétence perçue sur le sujet similaires. Les analyses n'ont révélé aucune différence significative entre les conditions, que ce soit pour l'intérêt préalable ($F(4, 95) = 0.01, p = 1.00$) ou pour la compétence perçue ($F(4, 95) = 0.04, p = .997$). Les données qualitatives sont représentées dans le Tableau 10.

Tableau 10. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 3)

Condition	Intérêt préalable		Compétence perçue	
	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	6.29	1.42	2.81	1.99
Étayage non-interactif	6.29	1.90	2.90	1.70
Étayage interactif	6.32	1.63	3.00	1.56
Pop-up	6.25	2.05	2.95	1.64
Pop-up avec présence visuelle	6.21	1.40	2.87	1.27

3.3. Tâche de RI

3.3.1. Réussite à la tâche

La réussite à la tâche était évaluée, ainsi chaque réponse donnée par le participant était codée soit 0 (réponse incorrecte) soit 1 (réponse correcte). Comme il s'agissait de données binomiales, des régressions logistiques ont été utilisées pour traiter cette variable (Field et al., 2012). Concernant les variables de temps de réponse et de pertinence de la réponse, des régressions linéaires ont été appliquées. Les résultats ont montré un effet significatif de la condition sur la réussite à la tâche, $\chi^2(4, N = 900) = 40.46, p < .001$, mais pas d'effet additif de la condition et du rang de la question, $\chi^2(1, N = 900) = .49, p = .486$, ni d'effet d'interaction de la condition avec le rang de la question³⁰, $\chi^2(4, N = 900) = .54, p = .970$. Le modèle M1 a donc été considéré comme étant le plus approprié pour rendre compte des données observées. Les données descriptives (voir Figure 20) montrent que la seule condition qui semble se détacher des quatre autres est la condition sans étayage. Les participants n'ayant pas bénéficié d'étayage pendant la tâche de RI semblent avoir trouvé moins de réponses correctes que les participants ayant bénéficié de double étayage.

³⁰ Ce résultat est à prendre avec précaution. En effet, un problème de convergence lors de l'estimation est apparu.

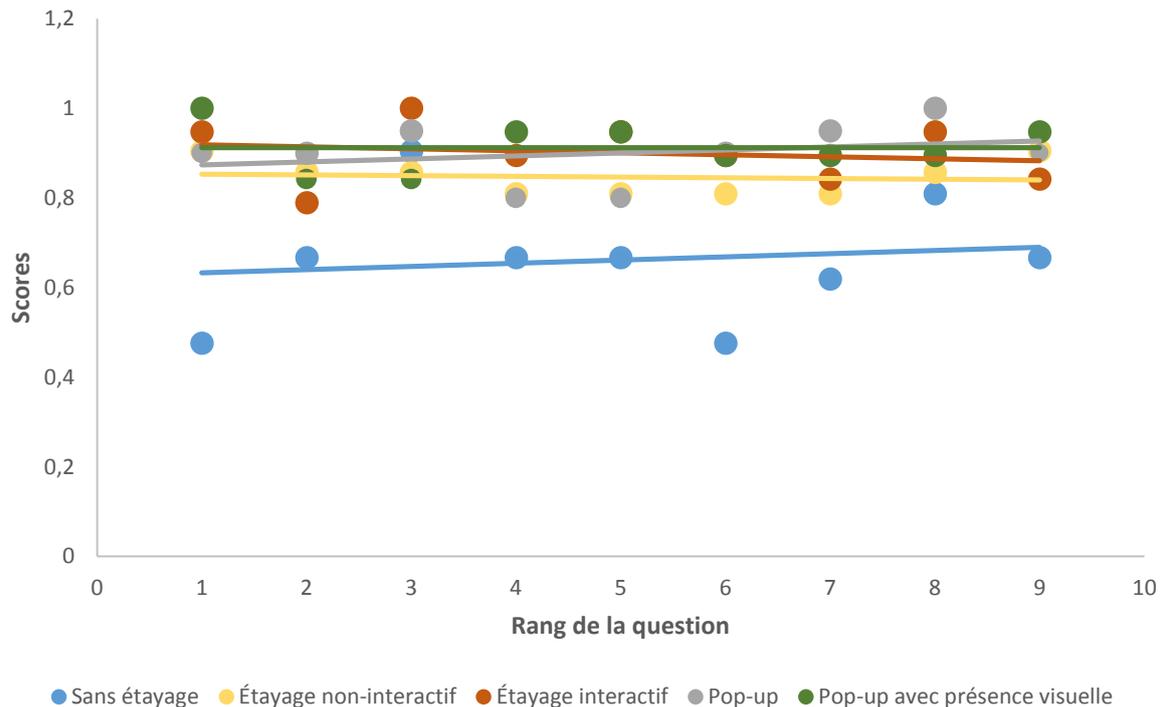


Figure 20. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 3

3.3.2. Temps de réponse

De la même manière que pour les études 1 et 2, les échecs ont été ôtés des données pour l'analyse des temps de réponse des participants. Pour rappel, un échec est un temps de réponse de 300 secondes (i.e., cinq minutes), lorsque le participant n'a pas pu trouver la réponse dans le temps maximal imparti. Un test du Chi-2 a été réalisé pour déterminer si le nombre d'échecs différait selon les conditions (voir Tableau 11 pour des données descriptives). Les résultats indiquent une différence significative, $\chi^2(4, N = 900) = 26.77, p < .001$. Les données descriptives indiquent que les participants ont eu moins d'échecs quand ils se trouvaient dans les conditions avec étayage.

Tableau 11. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 3)

Condition	Nombre d'échecs	% (condition)	% (total)
Sans étayage	18	9.52	2.00
Étayage non-interactif	3	1.59	0.33
Étayage interactif	4	2.34	0.44
Pop-up	3	1.67	0.33
Pop-up avec présence visuelle	3	1.75	0.33

Les résultats ont montré un effet significatif de la condition sur les temps de réponse, $\chi^2(4, N = 869) = 35.00, p < .001$, ainsi qu'un effet additif de la condition et du rang de la question, $\chi^2(1, N = 869) = 44.49, p < .001$, et un effet d'interaction de la condition avec le rang de la question, $\chi^2(4, N = 869) = 16.83, p = .002$. Le modèle M3 a donc été considéré comme étant le plus approprié. Les données descriptives (voir Figure 21) montrent que la seule condition qui semble se détacher des quatre autres est la condition sans étayage. Les participants n'ayant pas bénéficié d'étayage pendant la tâche de RI semblent avoir des temps de réponse plus longs que les participants ayant bénéficié de double étayage, mais ils semblent également améliorer leurs temps de recherche au cours de la tâche.

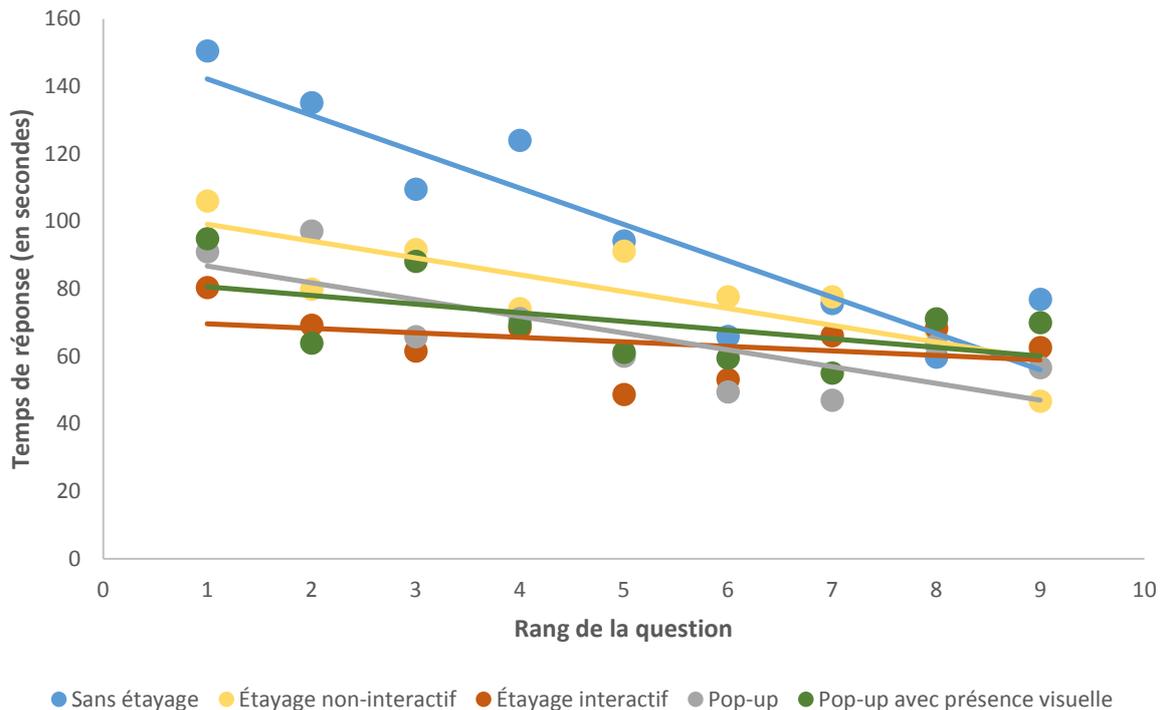


Figure 21. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 3

3.3.3. Pertinence du premier clic

Suivant le même procédé que pour les études 1 et 2, la pertinence de chaque premier clic de chaque question a été mesurée à travers le taux d'erreur, c'est-à-dire le nombre de segments entre celui contenant le premier clic et celui contenant la réponse attendue à la question. Les résultats ont montré un effet significatif de la condition sur la distance entre les deux segments mesurés, $\chi^2(4, N = 900) = 40.30, p < .001$, ainsi qu'un effet additif de la condition et du rang de la question, $\chi^2(1, N = 900) = 68.75, p < .001$, et un effet d'interaction de la condition avec le rang de la question, $\chi^2(4, N = 900) = 20.51, p < .001$. Le modèle M3 a donc été considéré comme étant le modèle le plus approprié. Les données descriptives (voir Figure 22) montrent que la seule condition qui semble se détacher des quatre autres est la condition sans étayage. Les participants n'ayant pas bénéficié d'étayage pendant la tâche de RI semblent avoir des taux d'erreur plus importants (i.e., une plus faible pertinence) que les participants ayant bénéficié de double étayage, mais ils semblent également améliorer la pertinence de leurs premiers clics au cours de la tâche.

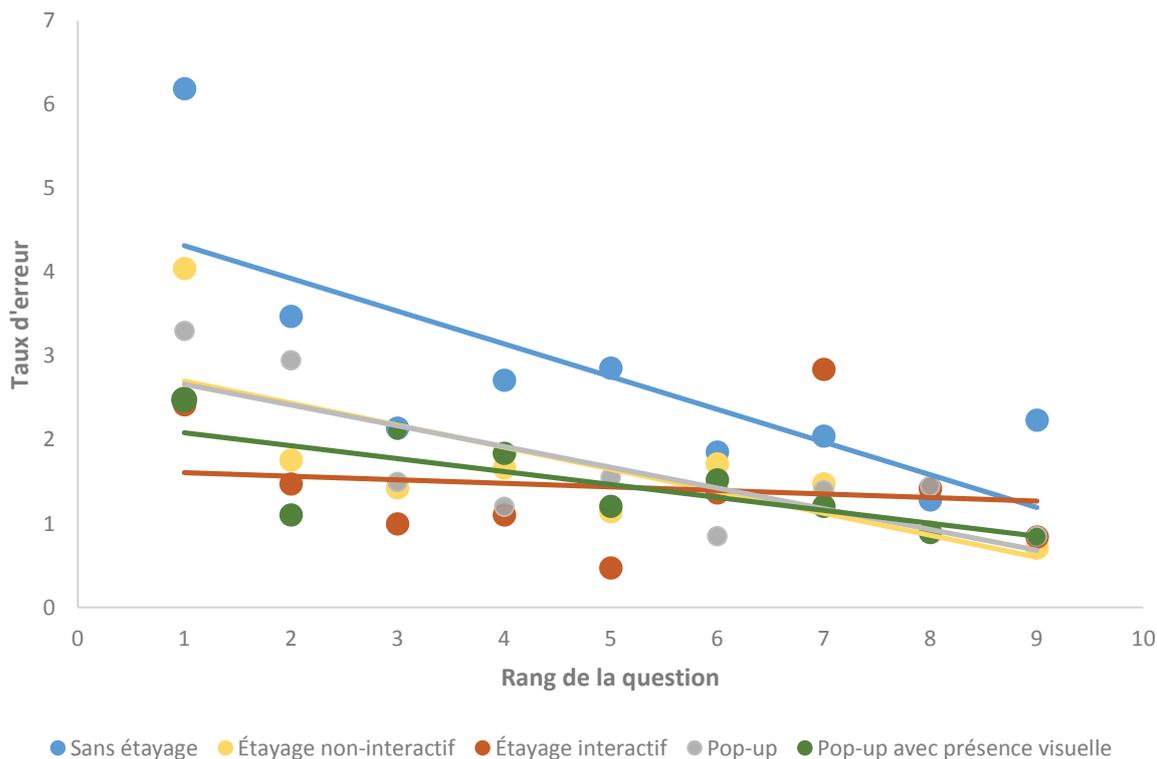


Figure 22. Taux d'erreur moyens des participants sur le premier clic en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 3

3.3.4. Difficulté perçue

Une ANOVA réalisée sur le score moyen total de difficulté perçue a révélé une différence significative entre les conditions ($F(4, 95) = 4.16, p = .004$). Les statistiques descriptives (voir Tableau 12) indiquent que les participants ayant réalisé la tâche de RI sans étayage de la vidéo ont ressenti davantage de difficulté que les participants ayant bénéficié du double étayage. Parmi les participants ayant eu le double étayage, ceux dans la condition *pop-up*, c'est-à-dire avec la segmentation de la barre de navigation (i.e., micro-étayage) associée à l'intégration de la structuration (i.e., macro-étayage) en *pop-up* dans la barre de navigation, ont ressenti moins de difficulté à réaliser la tâche que les participants des autres conditions.

Tableau 12. Données descriptives pour la variable de difficulté perçue en fonction de la condition expérimentale (étude 3)

Condition	Moy.	E.-T.
Sans étayage	3.30	0.98
Étayage non-interactif	2.14	0.99
Étayage interactif	2.16	1.10
<i>Pop-up</i>	1.80	0.66
<i>Pop-up</i> avec présence visuelle	2.11	1.02

4. Discussion

L'apport de l'étayage d'une vidéo sur l'activité de RI a été démontré dans la littérature et dans les études précédentes de cette thèse. L'objectif de cette étude était de comparer **différentes interfaces de présentation** des deux niveaux d'étayage (i.e., micro- et macro-étayages) soutenant l'activité de RI dans un environnement vidéo. Une condition sans étayage présentée (i.e., condition contrôle) était ainsi comparée à une condition de présentation des deux niveaux d'étayage sans interactivité entre eux, une condition permettant l'interactivité des deux niveaux, une condition avec le macro-étayage (i.e., structuration de la vidéo en chapitres) intégré au micro-étayage (i.e., segmentation de la barre de navigation de la vidéo) grâce à des *pop-ups*, et enfin à une condition avec les deux niveaux présentés, interactifs et également intégrés grâce à des *pop-ups*. L'hypothèse principale était que l'ajout d'interactivité entre les deux niveaux d'étayage faciliterait la mise en lien des informations issues de ces deux niveaux, et ainsi améliorerait la performance en RI. De la même manière, l'intégration spatiale des informations des deux niveaux d'étayage par le biais de *pop-ups* devrait également faciliter la mise en relation des informations de structuration et de segmentation. La présence visuelle permanente des deux niveaux d'étayage, associée à une possibilité d'interactivité et d'intégration spatiale de ces deux niveaux, devrait quant à elle être l'interface la plus favorable à une amélioration de la performance en RI et à une réduction de la difficulté perçue à réaliser la tâche.

Selon l'hypothèse 1, le niveau de réussite à la tâche, c'est-à-dire le fait de trouver la réponse attendue à la question par de la RI en vidéo, devrait augmenter

avec les ajouts successifs d'étayage, puis d'interactivité, d'intégration spatiale et enfin la présence continue de toutes ces aides. Il apparaît que la seule condition qui se distingue des quatre autres est la condition sans étayage, dans laquelle les participants réussissent moins bien à trouver les réponses correctes que les participants des autres conditions expérimentales. Ces taux de réussite ne semblent pas évoluer en fonction du temps. L'**hypothèse 1** est donc en partie seulement validée, les tentatives de mise en lien des informations issues des deux niveaux d'étayage ne semblent pas améliorer l'effet déjà bénéfique de l'étayage. Concernant les temps de réponse des participants, c'est-à-dire le temps passé à chaque question à chercher l'information dans la vidéo, un effet de la condition en interaction avec l'évolution temporelle de la tâche a été trouvé. Mais cette fois encore, la seule condition qui semble se distinguer des quatre autres est la condition sans étayage, dans laquelle les participants mettent davantage de temps lors de leur RI, mais semblent s'améliorer au cours de la tâche. Les participants des quatre conditions avec double étayage ont quant à eux le même pattern de résultats : ils sont bons dès le début de la tâche et le restent tout du long. L'**hypothèse 2** est elle aussi partiellement validée, la seule différence notable est celle existant entre les conditions avec double étayage et la condition sans étayage. De la même manière, l'**hypothèse 3** est également partiellement validée. Seuls les participants n'ayant pas bénéficié d'étayage se distinguent des autres : leurs scores d'erreur concernant la pertinence de leurs premiers clics semblent plus importants que ceux avec étayage, mais ils semblent également diminuer au cours de la tâche. Autrement dit, sans étayage, la pertinence du premier clic des participants est moins bonne qu'avec étayage, mais elle s'améliore au fur et à mesure des questions.

Concernant la variable de difficulté perçue, il semble que les participants ayant réalisé la tâche de RI sans étayage l'aient perçue comme étant plus difficile que les participants ayant pu bénéficier du double étayage. Parmi ces participants (i.e., avec double étayage), ceux dans la condition *pop-up*, c'est-à-dire avec le macro-étayage (i.e., structuration de la vidéo en chapitre) intégré au micro-étayage (i.e., segmentation de la barre de navigation de la vidéo), semblent être ceux ayant perçu la tâche comme étant moins difficile que les autres. L'**hypothèse 4** est donc à son tour partiellement validée.

Les différences entre la condition sans étayage et les quatre conditions avec double étayage étaient attendues et avaient déjà fait l'objet de précédentes études (voir études 1 et 2). Les résultats précédemment trouvés ont été répliqués dans cette étude, montrant un effet bénéfique de la présence d'étayage concernant la réussite à la tâche, les temps de réponse en RI et la pertinence des premiers clics de chaque RI. Les participants sans étayage semblent globalement améliorer leur performance tout au long de la tâche, alors que ceux bénéficiant de double étayage ont une performance relativement stable et à son meilleur niveau dès le début de la tâche.

Les hypothèses testées n'ont pu être validées dans leur ensemble, notamment à cause de faibles différences entre les quatre conditions présentant un double étayage. La seule variable permettant de distinguer l'une des conditions des trois autres est la difficulté perçue, les résultats indiquant que les participants en condition *pop-up* ont trouvé la tâche de RI plus facile à réaliser que les autres. Concernant les variables de performance (i.e., réussite à la tâche, temps de réponse et pertinence du premier clic), l'absence de différence significative entre les conditions expérimentales présentant un double étayage peut s'expliquer par l'effet de l'étayage lui-même. En effet, selon les conclusions tirées des précédentes études, et selon les résultats répliqués dans celle-ci, les performances des utilisateurs bénéficiant de l'étayage sont très bonnes dès le début de la tâche et restent stables. Sans étayage, ce n'est qu'au bout de la neuvième question que leurs performances deviennent similaires à celles des participants avec double étayage. On peut ainsi penser que le double étayage permet d'avoir tout au long de la tâche la performance **maximale** que les participants peuvent espérer. Changer la présentation, même si des résultats positifs étaient attendus, pourrait ne pas se traduire par une amélioration de performance tout simplement parce qu'elle était déjà à son meilleur niveau. La seule différence se traduit alors par la difficulté perçue par l'utilisateur à réaliser la tâche de RI, reflétant ainsi la facilité qu'il a eu à mobiliser ces deux niveaux d'étayage pour mettre en lien les informations issues du micro- et du macro-étayage.

Certaines limites sont néanmoins à souligner dans cette étude. Les modifications de format entre les quatre conditions présentant un double étayage

n'étaient peut-être pas assez importantes pour avoir un effet sur les performances en RI. De plus, les performances des utilisateurs semblent avoir été optimales dans ces quatre conditions. L'effet plafond observé, selon lequel les participants avec étayage avaient des performances stables et meilleures que sans étayage, accreditte l'hypothèse d'un effet massif du double étayage. Néanmoins, il pourrait également masquer des différences entre les formats de présentation. Une tâche plus difficile serait à réfléchir pour potentiellement faire apparaître ces différences. De la même manière, les utilisateurs pourraient avoir utilisé les deux niveaux d'étayage selon des stratégies différentes. Une étude utilisant par exemple un *eye-tracker* pourrait permettre d'analyser ces stratégies selon le format de présentation.

5. Conclusion

Selon ces résultats, le format de présentation des deux niveaux d'étayage sous la forme de *pop-ups* semble donc être le plus bénéfique à la RI. L'avantage de ce format peut s'expliquer par le fait que les deux niveaux d'étayage sont intégrés spatialement et l'effet d'attention divisée est évité ou bien, au moins, amoindri. Ces résultats concordent avec ceux de précédentes études menées dans un cadre d'apprentissage, le format *pop-up* était alors également bénéfique sur les performances (Bétrancourt & Bisseret, 1998 ; Erhel & Jamet, 2006). Cependant, contrairement à nos hypothèses, le format *pop-up* associé à une présence visuelle permanente des deux niveaux d'étayage dans l'environnement vidéo n'a pas permis d'améliorer la performance des participants ou de diminuer leur perception de difficulté durant la réalisation de la tâche de RI. On peut alors supposer que, lorsque les deux niveaux d'étayage sont à la fois intégrés grâce au format *pop-up* et présents visuellement en permanence, peut apparaître un effet négatif de surcharge visuelle. Cette problématique de la quantité d'informations présentes dans un document multimédia a déjà été posée dans la littérature (e.g., Bétrancourt et al., 2003 ; Erhel & Jamet, 2006 ; Fleury & Jamet, 2014), et les résultats de cette étude laissent à penser que lorsque trop d'informations, même si elles sont utiles, sont présentées en même temps, l'utilisateur ne peut les traiter simultanément et se perd dans l'exploration du document.

En conclusion, améliorer la mise en coréférence des deux niveaux d'étayage d'une vidéo lors d'une tâche de RI semble possible grâce au format **pop-up**. Ce mode de présentation de l'environnement vidéo permettrait ainsi de garder une performance de qualité tout au long de la tâche, mais également de faire en sorte que l'utilisateur ressente moins de difficulté à la réaliser. La mise en relation des informations issues à la fois des étayages de micro- et macro-niveaux serait ainsi facilitée, notamment grâce à la diminution de l'effet d'attention divisée. Les performances restent cependant similaires entre les conditions présentant un double étayage, peu importe le format de présentation. L'effet massif et positif du double étayage sur les performances en RI dans un environnement vidéo est une fois de plus confirmé.

ÉTUDE 4 : Effet de l'étayage sur la performance en apprentissage dans une vidéo³¹

1. Introduction et hypothèses

Les études précédentes (i.e., études 1, 2 et 3) ont chacune démontré un effet positif de la présence d'un double étayage (i.e., segmentation et structuration) lors d'une activité de RI en vidéo. La localisation et la sélection des informations pertinentes étant des composantes clés dans l'apprentissage (e.g., Cataldo & Oakhill, 2000 ; Mayer, 2014a ; Merkt & Schwan, 2014 ; Puustinen & Rouet, 2009 ; Rouet & Coutelet, 2008), il peut être envisagé que l'étayage, semblant être bénéfique à la RI, pourrait également l'être dans le cadre d'une activité d'apprentissage. Il agirait alors à la manière d'un modèle conceptuel, c'est-à-dire comme un guide à l'élaboration du modèle mental de l'apprenant. Cette quatrième étude vise donc à investiguer l'effet de l'étayage, et plus particulièrement les effets des deux niveaux possibles d'étayage (i.e., micro- et macro-niveaux : segmentation et structuration), sur les performances d'apprentissage au sein d'un environnement vidéo.

Les résultats de l'étude 2 menée précédemment ont indiqué que les participants utilisant l'étayage lors d'une tâche de RI ne semblaient pas l'internaliser. En effet, l'évaluation de la qualité de leur modèle mental a fait apparaître un modèle mental de la vidéo plus pauvre et moins efficace pour les participants avec étayage par rapport aux participants sans étayage. Ce résultat a été discuté en termes d'engagement de l'individu dans la tâche, en lien avec la consigne donnée : dans une tâche de RI, lorsque la consigne indique spécifiquement une activité de RI, il n'est pas nécessaire pour l'individu de concevoir un modèle mental pertinent de la vidéo à partir de l'étayage

³¹ Un article est actuellement en préparation à partir des études 4 et 5 de cette thèse : Cojean, S., & Jamet, E. (en préparation). The effect of scaffolding video-based environment during learning: information seeking as a necessary condition.

(i.e., modèle conceptuel). Ainsi, utiliser l'étayage comme ressource externe apparaît comme suffisant pour une tâche de RI. Dans le cadre d'une activité d'apprentissage, lorsque la consigne est d'améliorer ses connaissances sur le sujet abordé dans la vidéo, il a donc été supposé que l'individu utiliserait les ressources à sa disposition (ici, l'étayage) pour optimiser son apprentissage. Le modèle conceptuel ainsi fourni par l'étayage serait alors traité pour servir de base à la construction d'un modèle mental pertinent (Borgman, 1986 ; Norman, 1983). De plus, selon le modèle développé par Mayer (2014a, voir Introduction théorique 2.1.2), l'une des conditions pour favoriser l'apprentissage est de rendre l'apprenant actif dans le processus. Cette condition s'inspire directement des théories constructivistes, devenues dominantes dans le champ de l'apprentissage. Selon cette approche théorique, l'apprentissage est un processus actif au cours duquel les apprenants créent activement du sens en cherchant par eux-mêmes à construire des connaissances organisées (Hegarty, 2014 ; Mayer, 2004). Les individus mettraient ainsi en place de manière plus adaptée les processus de sélection, d'organisation et d'intégration de connaissances nécessaire à un apprentissage de qualité (Mayer, 2004). Dans le contexte de cette thèse, l'apprentissage actif serait notamment favorisé par les possibilités d'actions offertes par la présence d'étayage en vidéo.

Au sein des documents textuels, les études menées sur l'effet des signaux³² ont pu démontrer globalement un bénéfice de ce type d'étayage sur les performances en apprentissage (e.g., Brooks et al., 1983 ; Hyönä & Lorch, 2004 ; Lorch & Lorch, 1996 ; Sanchez et al., 2001). La présence de ce macro-étayage permettrait alors de mettre en saillance la structure du texte, de faciliter la compréhension de l'organisation de l'information dans le document, et de guider les processus de traitement de l'information par l'apprenant (Baccino & Draï-Zerbib, 2015 ; Lemarié, Lorch, Eyrolle, & Virbel, 2008 ; Lorch, 1989). Cependant, de précédentes études n'ont pas relevé de résultats satisfaisants en faveur de cette hypothèse dans un environnement vidéo (i.e., l'étayage favoriserait un apprentissage de meilleure qualité). Une étude menée par Merkt et ses collaborateurs (2011) a permis de comparer des formats vidéo sur des performances uniquement d'apprentissage. Dans cette étude, trois conditions

³² Outils qui mettent en valeur la structure du document, par exemple les titres correspondant à chaque paragraphe (Sanchez, Lorch, & Lorch, 2001).

expérimentales présentaient soit 1) un texte, 2) une vidéo « ordinaire », c'est-à-dire permettant de faire pause, ou de naviguer dans la vidéo (i.e., activités de micro-niveau), ou bien 3) une vidéo « améliorée », permettant en plus d'utiliser une table des matières et un index (i.e., activités de micro- et macro-niveau). Globalement, les résultats n'ont pas montré de différence d'apprentissage entre les différents types de formats. Cependant, il a été noté que les utilisateurs ayant à leur disposition un étayage de macro-niveau en ont peu fait l'usage. De la même manière, dans l'étude de Merkt et Schwan (2014), en plus des performances en RI, les auteurs se sont également penchés sur l'effet de l'étayage dans une vidéo en comparant les scores d'apprentissage. Pour rappel (voir Introduction théorique 5.2), cette étude comparait quatre conditions expérimentales : 1) un texte seul reprenant la transcription d'une vidéo, 2) une vidéo non-interactive (i.e., sans étayage), 3) une vidéo « ordinaire » (i.e., avec micro-étayage) et 4) une vidéo « améliorée » (i.e., avec micro- et macro-étayage). Les résultats sur les performances d'apprentissage n'ont néanmoins pas mis en relief de différence significative permettant de distinguer les formats de présentation du document. Aucun effet spécifique de la table des matières n'est apparu non plus dans leur tâche de RI. Comme précisé précédemment (voir Introduction théorique 5.2), le format vidéo dans ces deux études présentait une caractéristique non négligeable à notre sens. En effet, la table des matières et l'index n'étaient disponibles qu'en changeant de page à l'écran, donc en quittant momentanément la vidéo. Le processus d'attention divisée qu'implique ce format (Ayres & Sweller, 2014 ; Bétrancourt & Bisseret, 1998 ; Ginns, 2006) peut permettre de relativiser l'absence d'effet du macro-étayage lors d'une activité d'apprentissage. Par ailleurs, les résultats obtenus dans l'étude 1 de cette thèse ont pu montrer des effets distincts, complémentaires et bénéfiques des deux niveaux d'étayage sur les performances en RI. Au cours de cette étude, l'environnement vidéo présentant la vidéo, sa segmentation et sa structuration sur un même affichage : un format de présentation adéquat semble donc permettre l'utilisation de l'étayage de manière pertinente (au moins lors d'une tâche de RI). Ainsi, même si les études précédentes (Merkt & Schwan, 2014 ; e.g., Merkt et al., 2011) n'ont pas démontré d'effet de l'étayage sur les performances en apprentissage, il est probable que l'environnement vidéo alors proposé aux participants n'ait pas permis d'exploiter les apports de ces outils. Il paraît donc intéressant d'étudier l'effet de

l'étayage sur l'apprentissage tout en s'assurant de proposer un environnement vidéo ayant montré des effets positifs en RI.

En résumé, la présence d'étayage de la vidéo présentée durant une tâche d'apprentissage devrait avoir un rôle positif sur les performances en permettant la construction par l'apprenant d'un modèle mental pertinent du contenu de la vidéo. En effet, l'étayage jouerait le rôle d'un modèle conceptuel présenté à l'apprenant, facilitant ainsi l'élaboration du modèle mental (Borgman, 1986 ; Mayer, 1989 ; Seel, 2003). Sans étayage, cette élaboration serait longue et coûteuse. Le système représenté par le modèle conceptuel (i.e., l'étayage) serait alors le contenu de la vidéo. L'étayage, et plus particulièrement un étayage à double niveau, serait donc à internaliser pour aider à la construction d'un modèle mental. Cette élaboration du modèle mental serait alors plus rapide, moins coûteuse et plus pertinente dans le sens où le modèle conceptuel est précisément adapté à la vidéo. Pour rappel, la qualité du modèle mental ainsi construit aurait un rôle dans la mémorisation et la compréhension (voir Introduction théorique 2.2). Il est supposé qu'à la différence d'une tâche de RI, la présentation d'une consigne d'apprentissage permettrait d'engager l'apprenant dans un processus d'internalisation du modèle conceptuel présenté. Enfin, il peut également être supposé que l'apprentissage serait favorisé parce que la phase de révision³³ de la vidéo serait potentiellement plus pertinente en présence d'étayage. En effet, dans un contexte de temps limité d'apprentissage et de format linéaire de la vidéo, l'apprenant devrait rechercher l'information à réviser, et serait ainsi en condition de RI (Dee-Lucas & Larkin, 1995). L'étayage étant bénéfique à la performance en RI, l'hypothèse peut être faite que la phase de révision serait ainsi facilitée.

Les hypothèses suivantes ont ainsi été formulées :

Hypothèse 1. Apprentissage :

Parce que l'étayage jouerait le rôle de modèle conceptuel de la vidéo, il favoriserait la construction d'un modèle mental pertinent par l'apprenant. Ainsi, les

³³ D'après une étude menée par Karpicke, Butler et Roediger (2009), 84 % des apprenants auraient utilisé une stratégie de relecture du document pédagogique, contre 11 % qui auraient mis en place une stratégie de RI.

participants ayant bénéficié de l'étayage de micro- et/ou macro-niveau(x) devraient avoir un score d'apprentissage meilleur que les participants sans étayage. Cet effet devrait être maximal lorsque les deux niveaux d'étayages sont présents conjointement. Plus spécifiquement, cet effet devrait se retrouver dans les scores de mémorisation, c'est-à-dire lors de questions de rappel (Hypothèse 1a) et dans les scores de compréhension, c'est-à-dire avec des inférences (Hypothèse 1b).

Hypothèse 2. Activité de révision :

Les participants bénéficiant de double étayage devraient être plus pertinents durant leur phase de révision que les autres participants. Ces résultats sont liés aux précédents concernant l'effet de l'étayage sur la RI. Plus précisément, le micro-étayage, en favorisant la navigation au sein de la vidéo, devrait faciliter la recherche dans la barre de navigation de l'information à réviser (voir étude 1 ; Hypothèse 2a). Le temps passé à réviser chaque information devrait être plus important lorsque les deux niveaux d'étayage sont présents, parce que cette information serait plus pertinente à réviser (Hypothèse 2b).

Hypothèse 3. Difficulté perçue :

Les participants avec de l'étayage de micro- et/ou macro-niveau(x) devraient percevoir la tâche comme étant moins difficile que les participants ne bénéficiant pas d'étayage. Cet effet devrait être maximal lorsque les deux niveaux d'étayage sont présents conjointement.

Hypothèse 4. Contrôle perçu :

De la même manière que lors d'une tâche de RI (voir étude 1), les participants ayant réalisé la tâche d'apprentissage avec de l'étayage de micro-niveau (i.e., marqueurs sur la barre de navigation) devraient ressentir davantage de contrôle sur la tâche que les participants ne bénéficiant pas de micro-étayage.

Hypothèse 5. Rappel des chapitres :

Les participants bénéficiant de l'étayage de macro-niveau (i.e., table des matières) devraient rappeler correctement davantage de chapitres de la vidéo que les participants ne bénéficiant pas de macro-étayage.

2. Méthode

2.1. Participants

Au total, 77 participants volontaires (62 femmes, 15 hommes) ont pris part à cette étude. Tous étaient étudiants dans la région de Rennes. La moyenne d'âge des participants était de 20.21 ans (E.-T. = 2.27)³⁴. Le recrutement a été réalisé au sein de l'Université Rennes 2, notamment par le biais d'affichages. Pour les remercier de leur participation, tous les étudiants ont été gratifiés d'un ticket de cinéma. L'étude s'est déroulée conformément aux principes de la déclaration d'Helsinki.

2.2. Matériel

L'environnement d'apprentissage présentait une vidéo de dix-huit minutes environ, issue du site Internet de l'Inserm³⁵ et présentant différents aspects de la dyslexie (Moret & Hermann, 2009). Durant cette vidéo, les hypothèses explicatives sur ce trouble étaient détaillées, ainsi que l'accompagnement possible auprès des enfants diagnostiqués. La vidéo, dans son aspect originel, présentait les différents chapitres de manière visuelle (i.e., le titre s'affichait à l'écran). Afin de pouvoir contrôler la présence ou non de macro-structure (i.e., les titres) et de micro-structure (i.e., l'emplacement au sein de la vidéo), la vidéo a été modifiée. Les scènes présentant les titres ont été coupées. De la même manière que pour l'étude 1, un plan factoriel en 2 x 2 a été utilisé, quatre conditions expérimentales étaient donc testées (voir Figure 23) : la vidéo avait ou non des marqueurs sur la barre de navigation (quatorze sections correspondant aux quatorze chapitres, micro-étayage), et elle avait

³⁴ La moyenne d'âge des participants n'a pu être calculée qu'à partir de 75 participants, deux ayant indiqué une date de naissance erronée (e.g., 2017).

³⁵

<https://www.images.inserm.fr/fr/asset/fullTextSearch/search/dyslexie/page/1>

ou non une table des matières présentée sur le côté (macro-étayage). Dans toutes les conditions, les participants pouvaient se déplacer au sein de la barre de navigation en cliquant où ils le souhaitaient.

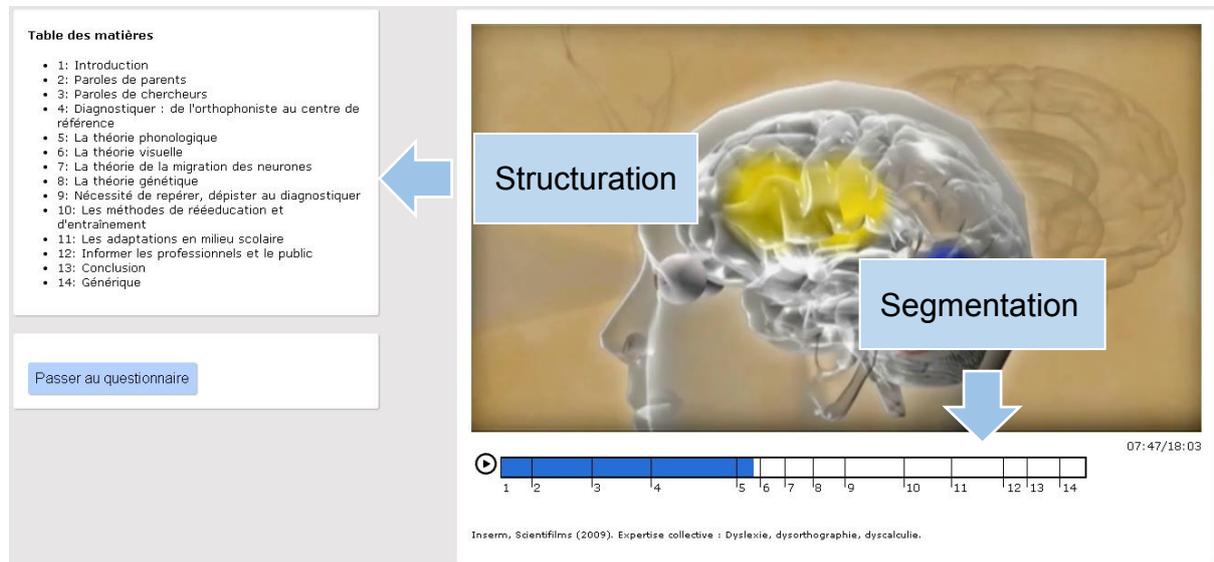


Figure 23. Segmentation et structuration au sein d'un environnement vidéo (étude 4)

L'apprentissage était limité à trente minutes. Durant ce temps, les participants étaient laissés libres dans leur navigation au sein de la vidéo. Un bouton placé à côté de la vidéo leur permettait de passer au questionnaire post-apprentissage dès qu'ils le désiraient. Ce questionnaire proposait des questions de difficulté perçue, de contrôle perçu, ainsi que des questions d'apprentissage et de rappel des chapitres de la vidéo. Les participants ont été aléatoirement répartis dans les quatre conditions expérimentales. Dans la **condition sans étayage** ($n = 19$), la vidéo était présentée avec une barre de navigation dans laquelle les participants pouvaient se déplacer. Un bouton pause était également disponible à côté de la barre de navigation. La durée totale de la vidéo ainsi que le temps de visionnage restant étaient indiqués au-dessus de la barre de navigation. Ni la table des matières, ni la segmentation de la barre de navigation en quatorze chapitres n'étaient présentées (voir Figure 24). Les participants de la **condition micro-étayage** ($n = 20$) étaient confrontés au même matériel que ceux dans la condition sans étayage, à la seule différence que la barre de navigation était segmentée en quatorze parties. Ces segments correspondaient aux chapitres de la vidéo, sans en préciser les titres, mais en permettant d'y accéder directement (voir Figure 24). Les participants de la **condition macro-étayage** ($n = 19$) étaient

confrontés au même matériel que ceux dans la condition sans étayage, à la seule différence qu'une table des matières était présentée à côté de la vidéo sur le même écran. Cette table des matières structurait la vidéo en quatorze chapitres dont les titres étaient indiqués, mais il n'était pas possible de cliquer dessus pour y accéder dans la vidéo (voir Figure 24). Enfin, dans la **condition double-étayage** ($n = 19$), les participants pouvaient se déplacer dans la vidéo avec l'aide à la fois de la segmentation de la barre de navigation (i.e., micro-étayage) et de la table des matières (i.e., macro-étayage). Dans cette condition expérimentale, les participants pouvaient cliquer sur un chapitre dans la table des matières pour y accéder directement dans la vidéo. Les numéros associés à chaque titre de chapitre étaient également associés à des numéros présents sous chaque segment de la barre de navigation (voir Figure 24).

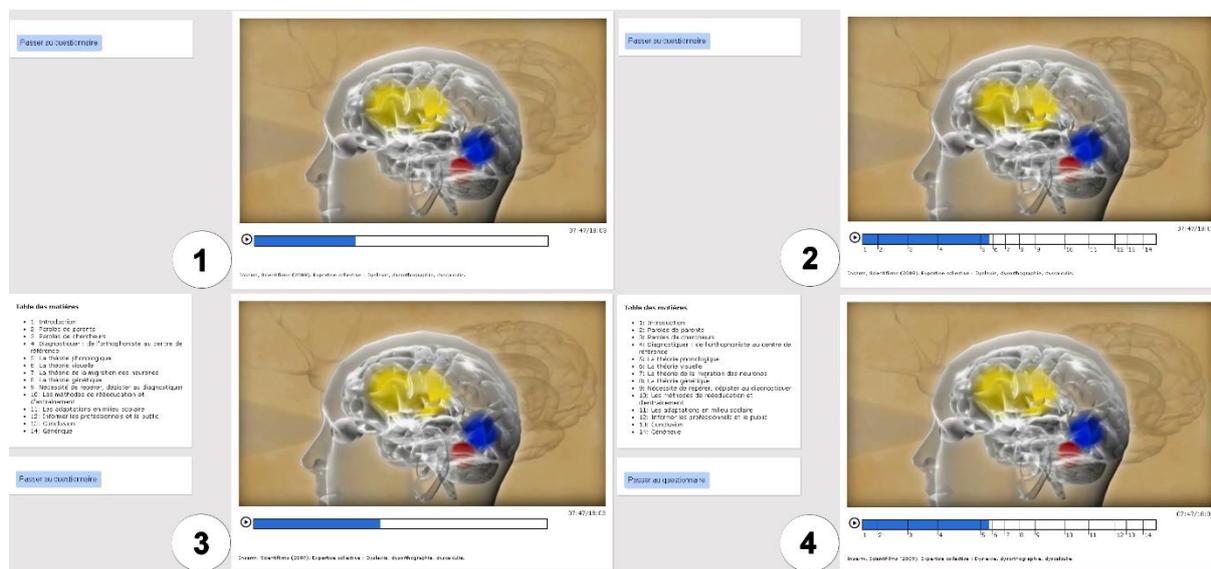


Figure 24. Environnements vidéo dans les conditions sans étayage (1), micro-étayage (2), macro-étayage (3) et double niveau d'étayage (4) de l'étude 4

2.3. Mesures

2.3.1. Intérêt préalable et compétence perçue

Avant la tâche d'apprentissage, le participant était informé que le sujet abordé dans la vidéo était la dyslexie. Il lui était alors demandé d'indiquer son niveau d'intérêt préalable quant à ce sujet, ainsi que sa compétence perçue. Pour cela, il devait se placer sur une échelle de Likert en 11 points selon son degré d'intérêt (« Sur une

échelle de 0 à 10, à quel point ce sujet vous intéresse-t-il ? ») et de compétence perçue (« Sur une échelle de 0 à 10, à quel point vous estimez-vous compétent sur ce sujet ? »). L'objectif était de s'assurer que les scores des individus sur ces deux variables étaient équitablement répartis parmi les quatre conditions expérimentales.

2.3.2. Temps d'apprentissage

Afin de mesurer le plus précisément possible le temps alloué à l'apprentissage par chaque participant, un logiciel enregistreur d'écran³⁶ a été utilisé. Les temps d'apprentissage ont ainsi été calculés à partir du moment où l'apprenant lançait le visionnage de la vidéo, et jusqu'à ce qu'il clique sur le bouton « passer au questionnaire d'apprentissage ». Dans le cas où il aurait cliqué d'abord sur pause dans la vidéo avant de passer au questionnaire, c'est le moment où la vidéo s'est arrêtée qui a compté comme marqueur de fin de l'apprentissage. Pour rappel, le temps maximal d'apprentissage était limité à trente minutes.

2.3.3. Score d'apprentissage

Après consultation du document, les participants devaient répondre à un questionnaire d'apprentissage. Ce questionnaire se décomposait en deux types de questions : des questions de rappel, visant à mesurer la mémorisation des apprenants, et des questions d'inférences, afin de mesurer un score de compréhension. Quinze questions de rappel ont ainsi été proposées (e.g., « Quelles sont les 4 théories sur l'origine de la dyslexie développées dans cette vidéo ? » ; voir Annexe 3). Les réponses à ces questions se trouvaient toutes explicitement exprimées au sein de la vidéo. Cinq questions d'inférences étaient également posées (e.g., « Certains dyslexiques traitent chaque mot comme un nouveau mot en assemblant les lettres une à une. D'autres ne lisent les mots qu'en les adressant directement au lexique et ont beaucoup de mal à lire de nouveaux mots en les assemblant. À quelles théories chacun de ces comportements correspond-il ? » ; voir Annexe 3). Les réponses à ces questions n'étaient pas explicitement dites dans la vidéo. Pour y répondre, l'apprenant devait pouvoir mettre en relation différentes informations de la vidéo.

³⁶ Flashback by Blueberry Software, <https://www.flashbackrecorder.com/fbhome/>

2.3.4. Activité de révision

La pertinence de l'activité de révision par les participants peut s'envisager sous l'angle de deux mesures : le nombre de clics dans la barre de navigation pour retrouver l'information à réviser, et le temps passé à réviser chaque information. Ces données ont pu être extraites grâce aux données collectées avec le logiciel enregistreur d'écran. Les activités de révision étaient considérées comme telles lorsque les apprenants repassaient sur une information déjà visionnée. En d'autres termes, il s'agissait des clics dans la barre de navigation après un visionnage complet de la vidéo une première fois.

2.3.5. Difficulté perçue et contrôle perçu

La difficulté perçue et le contrôle perçu ont été évalués de la même manière que pour les études précédentes (voir études 1, 2 et 3). Les items utilisés ont été adaptés ou inspirés de précédentes études dans la littérature (Kraft et al., 2005 ; Trafimow et al., 2002). Trois items de difficulté perçue ont été utilisés (e.g., « Pour moi, apprendre cette vidéo était difficile » ; voir Annexe 5), ainsi que trois items de contrôle perçu (e.g., « J'ai eu un contrôle total sur mon apprentissage de cette vidéo » ; voir Annexe 6). Pour chaque item, les participants devaient indiquer sur une échelle de Likert en 7 points le degré avec lequel ils étaient d'accord avec l'affirmation proposée. Des alphas de Cronbach ont été réalisés, avec un résultat de $\alpha = .80$ pour les items de difficulté perçue et $\alpha = .82$ pour les items de contrôle perçu.

2.3.6. Rappel des chapitres

Lors du post-questionnaire, il a enfin été demandé aux participants de nommer les quatorze parties composant la vidéo. Pour les individus ayant réalisé la tâche d'apprentissage dans les conditions macro-étayage et double-étayage, il s'agissait d'une tâche de mémorisation, puisque la table des matières avec les noms des chapitres avait été présente tout au long de leur apprentissage, à côté de la vidéo. Pour les participants des groupes micro-étayage et sans étayage, puisqu'ils n'avaient donc jamais vu la table des matières, il s'agissait plutôt d'une tâche d'inférence.

2.4. Procédure

Les participants étaient accueillis et installés chacun sur un poste d'ordinateur. Trois personnes maximum pouvaient venir passer l'expérience en même temps. Une fois installés, les participants remplissaient un formulaire de consentement en deux exemplaires (i.e., un pour eux et un pour l'expérimentateur). Le pré-questionnaire leur était ensuite proposé : les participants répondaient aux questions d'intérêt préalable et de compétence perçue sur le sujet de la dyslexie. Puis, ils lisaient une feuille présentant la consigne de la tâche (voir Annexe 11). Il leur était précisé qu'il s'agissait d'une tâche d'apprentissage, que des questions sur le contenu de la vidéo seraient posées à la fin de cette tâche, et qu'ils disposaient de maximum trente minutes pour apprendre le maximum d'informations présentées au sein de cette vidéo. Après s'être assuré que les participants avaient compris la consigne et n'avaient pas de question, l'expérimentateur lançait le logiciel enregistreur d'écran et indiquait aux apprenants qu'ils pouvaient commencer quand ils le souhaitaient. Les individus devaient porter un casque audio pour entendre la vidéo.

Afin de se situer dans leur temps d'apprentissage, ils pouvaient regarder l'heure affichée sur l'écran d'ordinateur. De plus, au bout de vingt-cinq minutes d'apprentissage, l'expérimentateur venait les informer qu'il leur restait cinq minutes maximum. Lorsque les participants arrivaient à trente minutes d'apprentissage, l'expérimentateur leur indiquait de passer au questionnaire d'apprentissage³⁷. Dans tous les autres cas, les apprenants passaient au questionnaire avant la fin du temps réglementaire.

En cliquant sur le bouton « passer au questionnaire d'apprentissage », un message s'affichait alors à l'écran. Il était demandé au participant de prévenir et d'attendre l'expérimentateur. Celui-ci lançait alors le post-questionnaire. Il était d'abord proposé aux participants de se positionner sur des items de difficulté perçue et de contrôle perçu, affichés dans un ordre aléatoire. Après cela, les questions d'apprentissage étaient proposées. Les vingt questions, comprenant quinze questions de mémorisation et cinq questions de compréhension, étaient proposées dans un

³⁷ Un seul participant est resté durant les trente minutes possibles à apprendre la vidéo.

ordre aléatoire afin d'éviter les effets d'ordre (Benton & Daly, 1991). Il était précisé aux participants d'indiquer une réponse à chaque question, mais que cette réponse pouvait également être « je ne sais pas ».

À la suite du questionnaire d'apprentissage, les participants devaient répondre au questionnaire de rappel des chapitres. La consigne était de proposer ou rappeler le nom des quatorze parties composant la vidéo qu'ils venaient de visionner. Enfin, des questions sur l'âge et le sexe du participant terminaient le post-questionnaire. La procédure est résumée en Figure 25. Après avoir répondu au post-questionnaire, l'expérimentateur gratifiait chaque participant d'un ticket de cinéma pour les remercier de leur participation à l'étude.

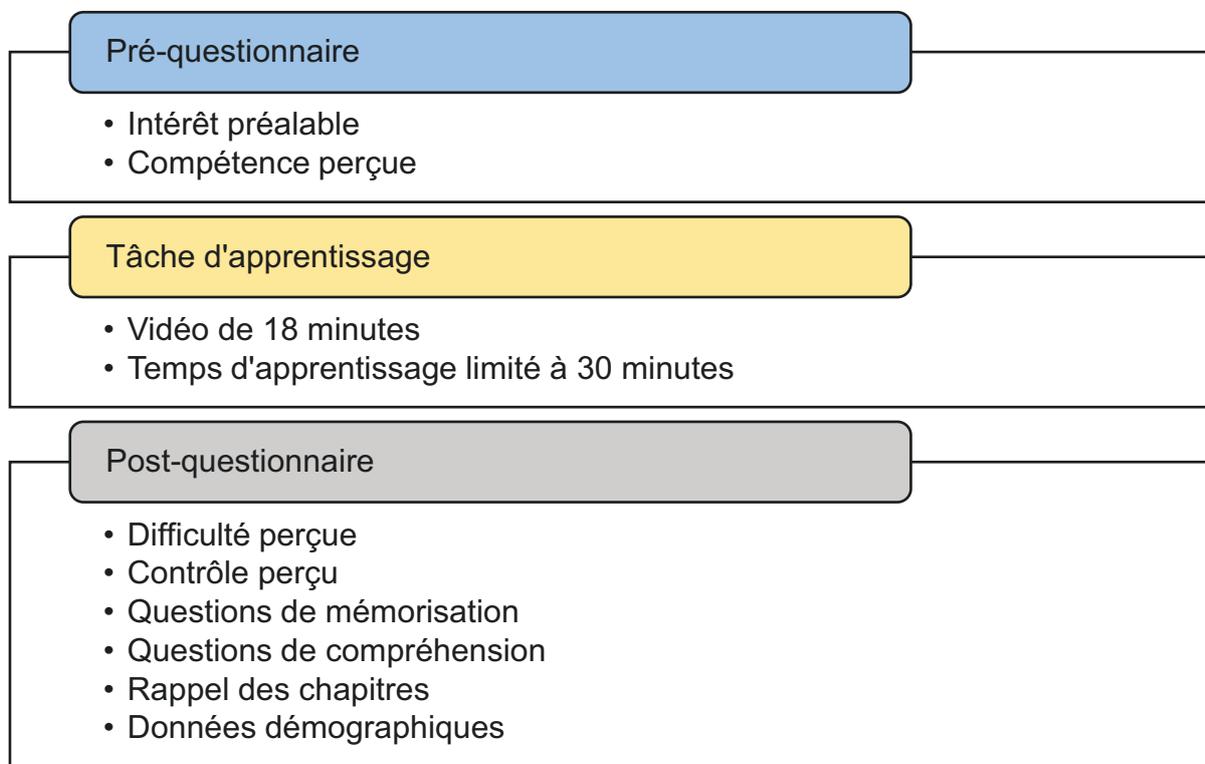


Figure 25. Procédure de l'étude 4

3. Résultats

3.1. Variables contrôles (pré-questionnaire)

Des ANOVA ont été réalisées sur les variables d'intérêt préalable et de compétence perçue, afin de s'assurer que les groupes ne différaient pas entre eux avant le début de l'expérimentation. Les résultats ne montrent aucune différence

significative entre les groupes, que ce soit pour l'intérêt préalable, $F(3, 73) = 0.57$, $p = .634$, $\eta^2_p = .02$, ou pour la compétence perçue, $F(3, 73) = 0.24$, $p = .870$, $\eta^2_p = .01$. Les statistiques descriptives sont présentées dans le Tableau 13.

Tableau 13. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 4)

Condition	Intérêt préalable		Compétence perçue	
	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	7.00	1.37	3.42	1.80
Micro-étayage	6.70	1.53	3.55	1.54
Macro-étayage	6.58	1.35	3.05	2.22
Double niveau d'étayage	7.11	1.45	3.32	2.00

3.2. Tâche d'apprentissage

3.2.1. Temps d'apprentissage

Afin de s'assurer que les temps d'apprentissage ne différaient pas selon les conditions expérimentales, une ANOVA en 2 x 2 a été conduite sur les données recueillies. Les résultats n'ont pas révélé d'effet principal significatif, ni du micro-étayage, $F(1, 73) = 3.49$, $p = .066$, $\eta^2_p = .05$, ni du macro-étayage, $F(1, 73) = 0.98$, $p = .325$, $\eta^2_p = .01$, ni d'effet d'interaction significatif, $F(1, 73) = 0.77$, $p = .384$, $\eta^2_p = .01$ (voir Tableau 14 pour les données descriptives).

3.2.2. Score d'apprentissage

Le score d'apprentissage a été mesuré selon deux types de questions : des questions de rappel et des questions d'inférences. Deux scores ont donc été analysés : le score moyen de mémorisation et le score moyen de compréhension.

Concernant le score de mémorisation, l'ANOVA en 2 x 2 réalisée n'a pas révélé d'effet principal significatif, ni du micro-étayage, $F(1, 73) = 0.31$, $p = .582$, $\eta^2_p = .00$, ni du macro-étayage, $F(1, 73) = 2.41$, $p = .125$, $\eta^2_p = .03$, ni d'effet d'interaction, $F(1, 73) = 0.44$, $p = .511$, $\eta^2_p = .01$ (voir Tableau 14 pour les données descriptives).

Concernant le score de compréhension, l'ANOVA en 2 x 2 n'a pas révélé d'effet principal, ni du micro-étayage, $F(1, 73) = 1.09$, $p = .299$, $\eta^2_p = .02$, ni du macro-étayage, $F(1, 73) = 0.37$, $p = .544$, $\eta^2_p = .01$, ni d'effet d'interaction, $F(1, 73) = 1.07$, $p = .305$, $\eta^2_p = .01$ (voir Tableau 14 pour les données descriptives).

Tableau 14. Données descriptives pour les variables d'apprentissage (temps d'apprentissage et scores de mémorisation et de compréhension) en fonction de la condition expérimentale (étude 4)

Condition	Temps d'apprentissage (en s.)		Score de mémorisation (sur 15)		Score de compréhension (sur 5)	
	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	1397.37	219.52	9.03	2.27	1.58	1.15
Micro-étayage	1267.50	184.41	8.98	2.51	2.10	0.84
Macro-étayage	1401.21	233.48	7.88	2.27	2.00	1.15
Double niveau d'étayage	1354.58	194.67	8.51	1.98	2.00	1.26

3.2.3. Activité de révision

Deux types de données permettaient de se renseigner sur l'activité de révision réalisée par les apprenants : le nombre de clics réalisés dans la barre de navigation après un premier visionnage de la vidéo, et le temps moyen passé à visionner l'information lors de cette tâche (i.e., temps total de révision divisé par le nombre de clics). Pour rappel, les activités de révision ont été comptabilisées comme telles après un visionnage une première fois de la vidéo en entier.

Une ANOVA en 2 x 2 réalisée sur le nombre de clics (i.e., révisions) a montré un effet principal du micro-étayage, $F(1, 73) = 6.46$, $p = .013$, $\eta^2_p = .08$, mais pas d'effet du macro-étayage, $F(1, 73) = 0.18$, $p = .676$, $\eta^2_p = .00$, ni d'effet d'interaction entre les deux niveaux d'étayage, $F(1, 73) = 0.84$, $p = .364$, $\eta^2_p = .01$. Il apparaît que les apprenants ayant été exposés au micro-étayage ont réalisé moins de clics de révision que les apprenants dans les autres conditions (voir données descriptives dans le Tableau 15).

Une ANOVA en 2 x 2 réalisée sur le temps total de révision (après un premier visionnage de la vidéo) n'a pas montré d'effet principal, ni du micro-étayage, $F(1, 73) = 1.81$, $p = .018$, $\eta^2_p = .02$, ni du macro-étayage, $F(1, 73) = 1.61$, $p = .209$, $\eta^2_p = .02$, ni d'effet d'interaction entre les deux niveaux d'étayage, $F(1, 73) = 0.18$, $p = .672$, $\eta^2_p = .00$. En moyenne, l'ensemble des apprenants a passé 4 minutes et 35 secondes sur l'activité de révision (voir données descriptives dans le Tableau 15).

Concernant les temps moyens par révision, l'ANOVA en 2 x 2 a révélé un effet principal du micro-étayage, $F(1, 73) = 7.89$, $p = .006$, $\eta^2_p = .10$, ainsi qu'un effet du macro-étayage, $F(1, 73) = 5.05$, $p = .028$, $\eta^2_p = .06$, et un effet d'interaction entre les deux niveaux d'étayage, $F(1, 73) = 6.98$, $p = .010$, $\eta^2_p = .09$. Les apprenants ayant été exposés à un étayage de double niveau ont en moyenne passé davantage de temps sur chaque révision que les apprenants dans les autres conditions expérimentales (voir données descriptives dans le Tableau 15).

Tableau 15. Nombre de révisions effectuées, temps total et temps moyen par révision en fonction de la condition expérimentale (étude 4)

Condition	Nombre de révisions (clics)		Temps total de révision (en s.)		Temps moyen par révision (en s.)	
	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	19.74	26.82	288.11	231.59	19.23	14.78
Micro-étayage	12.45	20.00	202.30	199.32	20.43	24.96
Macro-étayage	21.84	19.35	328.00	216.60	16.85	16.20
Double niveau d'étayage	6.42	5.29	283.47	204.97	50.01	41.40

3.3. Post-questionnaire

3.3.1. Difficulté perçue

Une ANOVA en 2 x 2 a été conduite sur les scores de difficulté perçue lors de la tâche d'apprentissage. Les résultats ont montré un effet principal du micro-étayage, $F(1, 73) = 5.13$, $p = .027$, $\eta^2_p = .07$, mais pas d'effet du macro-étayage, $F(1, 73) = 0.78$, $p = .382$, $\eta^2_p = .01$, ni d'effet d'interaction entre les deux niveaux

d'étayage, $F(1, 73) = 0.20$, $p = .660$, $\eta^2_p = .00$ (voir données descriptives dans le Tableau 16).

3.3.2. Contrôle perçu

L'ANOVA en 2 x 2 réalisée sur la variable de contrôle perçu durant l'apprentissage en vidéo n'a pas révélé d'effet du micro-étayage, $F(1, 73) = 2.92$, $p = .092$, $\eta^2_p = .04$, ni du macro-étayage, $F(1, 73) = 0.05$, $p = .817$, $\eta^2_p = .00$, ni d'interaction, $F(1, 73) = 0.60$, $p = .441$, $\eta^2_p = .01$ (voir Tableau 16 pour les données descriptives).

3.3.3. Rappel chapitres

Comme pour l'étude 1, concernant les chapitres rappelés sur les quatorze de la vidéo, un score appelé « distance de Levenshtein » (Yujian & Bo, 2007) a été attribué à chaque participant (voir étude 1 pour une définition). Une ANOVA en 2 x 2 a par la suite été conduite sur ces scores pouvant aller de 0 à 14. Les résultats ont montré un effet principal du macro-étayage, $F(1, 73) = 43.54$, $p < .001$, $\eta^2_p = .37$, mais pas d'effet du micro-étayage, $F(1, 73) = 2.57$, $p = .113$, $\eta^2_p = .03$, ni d'effet d'interaction, $F(1, 73) = 1.20$, $p = .278$, $\eta^2_p = .02$. Les données descriptives sont affichées dans le Tableau 16.

Tableau 16. Données descriptives pour les variables de difficulté perçue, contrôle perçu et distance de Levenshtein sur les chapitres rappelés en fonction de la condition expérimentale (étude 4)

Condition	Difficulté perçue		Contrôle perçu		Rappel des chapitres (distance de Levenshtein)	
	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.
Sans étayage	3.32	1.62	5.04	1.33	12.32	1.49
Micro-étayage	2.58	1.13	5.30	0.98	12.00	2.15
Macro-étayage	2.95	0.94	4.75	1.55	9.00	3.25
Double niveau d'étayage	2.46	0.92	5.46	0.98	7.37	3.25

4. Discussion

L'objectif de cette quatrième étude était de préciser les effets des deux niveaux d'étayage (i.e., micro- et macro-niveaux) sur les performances d'apprentissage dans un environnement vidéo. Les bénéfices démontrés dans les études précédentes de l'étayage sur l'efficacité en RI ont suggéré qu'il pourrait en être de même dans une tâche d'apprentissage. L'hypothèse proposée dans cette étude était la suivante : l'étayage de micro- et de macro-niveaux favoriserait la construction d'un modèle mental fidèle à la représentation de la vidéo, et donc améliorerait les performances en apprentissage. Il a également été supposé que chaque niveau d'étayage aurait un effet spécifique : le micro-étayage permettrait à l'apprenant de ressentir davantage de contrôle perçu sur la tâche et de naviguer plus facilement lors des révisions, et le macro-étayage faciliterait le rappel des chapitres de la vidéo.

Les résultats de l'étude ne montrent aucune différence significative en termes de temps d'apprentissage entre les quatre groupes expérimentaux. Ils ont tous globalement passé autant de temps sur la vidéo, en moyenne 22 minutes et 33 secondes. Ce résultat peut s'expliquer par la limite imposée aux apprenants de trente minutes pour la consultation de la vidéo, ainsi que par la durée même de la vidéo (18 minutes) définissant théoriquement le temps minimal de consultation. Peu de variabilité semble donc possible entre ces deux bornes.

Les analyses statistiques n'ont pas révélé non plus de différence significative entre les conditions expérimentales concernant l'apprentissage. Avec ou sans étayage, et quel que soit le niveau d'étayage, les apprenants ont réalisé globalement le même score sur les questions de mémorisation ou de compréhension. L'**hypothèse 1** n'est donc pas validée.

L'activité de révision réalisée par les apprenants a duré globalement le même temps pour tous. Néanmoins, le nombre de clics effectués dans la barre de navigation s'est avéré moindre pour les apprenants bénéficiant d'étayage de micro-niveau. Cela confirme l'**hypothèse 2a**, et apparaît également cohérent avec les résultats précédemment trouvés dans l'étude 1 : le micro-étayage permet de faciliter la navigation au sein de la vidéo en offrant la possibilité de contrôler le flot d'information délivré par la vidéo. Par ailleurs, le temps moyen alloué à chaque révision est significativement plus important lorsque les deux niveaux d'étayage sont réunis. L'**hypothèse 2b** est donc également validée. Cela reflèterait la qualité de l'activité de révision : les apprenants auraient passé davantage de temps sur chaque révision parce qu'elle était plus pertinente. Le double étayage aurait ainsi permis une meilleure navigation (grâce au micro-étayage) et un meilleur ciblage sémantique (grâce au macro-étayage, voir étude 1) de l'information recherchée pour être révisée.

Concernant la difficulté perçue, les résultats indiquent un effet du micro-étayage : lorsqu'ils ont bénéficié de la segmentation de la vidéo (i.e., conditions micro-étayage et double niveau d'étayage), les apprenants ont perçu la tâche comme étant moins difficile que ceux n'ayant pas eu de barre de navigation segmentée. Cet effet est en désaccord avec les résultats de l'étude 1. En effet, lors d'une tâche de RI, les tests statistiques ont révélé un effet d'interaction significatif entre les deux niveaux d'étayage sur la difficulté perçue. L'effet ici seulement du micro-étayage valide donc partiellement l'**hypothèse 3**.

Des effets spécifiques pour chaque niveau d'étayage étaient en revanche attendus concernant les variables de contrôle perçu et de rappel des chapitres. Aucune différence significative entre les groupes expérimentaux n'est apparue pour les scores de contrôle perçu. Autrement dit, quelle que soit la condition, avec ou sans étayage, les participants ont indiqué avoir ressenti le même niveau de contrôle sur la tâche d'apprentissage. Ces résultats vont à l'encontre de l'**hypothèse 4**. Néanmoins,

les apprenants ayant bénéficié de structuration (i.e., groupes macro-étayage et double niveau d'étayage) ont statistiquement rappelé davantage de chapitres correctement que les apprenants n'ayant pas bénéficié de ce niveau d'étayage. L'**hypothèse 5** est ainsi validée.

Globalement, aucun effet significatif de l'étayage n'est observé sur cette tâche d'apprentissage. Le seul effet attesté dans la littérature et retrouvé dans l'étude 1 est l'effet bénéfique de la présence d'une table des matières sur l'organisation sémantique du document, et donc le rappel des chapitres (Amadiou & Salmerón, 2014 ; Sanchez et al., 2001). Cet effet semble suffisamment solide pour s'observer également dans cette étude. L'effet attendu spécifique du micro-étayage (i.e., sur le contrôle perçu) n'a pas été observé, ni aucun effet sur les performances d'apprentissage en elles-mêmes. Selon l'hypothèse de départ, l'individu, parce que la consigne était une consigne d'apprentissage, aurait dû s'engager dans un apprentissage actif et intégrer l'étayage afin de s'en servir comme d'une base pour la construction de son modèle mental. Les résultats de cette étude n'ont mis en évidence aucun bénéfice de l'apport d'étayage sur les performances d'apprentissage. La consigne d'apprentissage n'a donc pas suffi à engager l'apprenant dans un traitement actif de l'étayage pour l'utiliser comme un modèle conceptuel. Ces résultats sont en accord avec l'idée développée dans la littérature selon laquelle les individus confrontés à un document pédagogique ne sont pas toujours conscients des indices organisationnels de textes et structurels (Lazonder & Rouet, 2008). De précédentes études (Lazonder & Rouet, 2008 ; Potocki, Ros, Vibert, & Rouet, 2017 ; Rouet & Coutelet, 2008) ont été menées sur les stratégies employées par des enfants lors d'une activité de RI au sein d'un document écrit. Les auteurs ont globalement démontré que l'utilisation de l'étayage à disposition (une table des matières par exemple) évoluait en fonction de l'âge des enfants et de leur capacité à interpréter ces outils de représentation. Trop jeunes, ils n'auraient pas conscience de l'aide que pourrait apporter l'étayage. Chez les adultes ou étudiants, ces stratégies semblent acquises, et l'étayage serait utile à la performance en RI (voir résultats des études 1, 2, 3). Lors d'une activité d'apprentissage comme au sein de cette quatrième étude, l'utilisation de l'étayage n'est peut-être pas aussi évidente. Les apprenants ne seraient peut-être pas forcément conscients de l'aide apportée par une table des

matières ou une segmentation lors de cette activité. Néanmoins, il est apparu que la présence d'étayage, et plus spécifiquement lorsque les deux niveaux d'étayage ont été utilisés conjointement, a favorisé la qualité de l'activité de révision. En effet, les apprenants avec double étayage ont en moyenne passé davantage de temps à réviser l'information après chaque clic : leur activité de révision semble donc avoir été plus ciblée, plus pertinente. Cette activité peut s'envisager comme une tâche de RI, les individus recherchant l'information qu'ils souhaiteraient réviser au sein de la vidéo : ces résultats sont donc cohérents avec les études précédentes de cette thèse sur les effets de l'étayage en RI. Les comportements de révision avec double étayage apparaissent donc meilleurs, bien qu'aucun effet significatif sur l'apprentissage n'ait pu être observé. Les apprenants auraient ainsi la capacité de réaliser une activité d'apprentissage plus efficace grâce à l'étayage, mais cette aide semble n'être utilisée que tardivement dans l'apprentissage (plus de 70 % des apprenants ne semblent pas l'utiliser lors d'un premier visionnage de la vidéo). La simple présence d'une consigne d'apprentissage ne suffirait donc pas à les rendre suffisamment actifs pour traiter l'information en profondeur en utilisant l'étayage à leur disposition. Contraindre les apprenants à utiliser l'étayage avant leur activité d'apprentissage pourrait potentiellement leur permettre de le traiter activement une première fois avant même le visionnage de la vidéo.

La limite principale de cette étude pourrait venir du fait qu'il s'agisse d'une tâche d'apprentissage : cette situation n'est pas écologique pour le participant. En effet, même si une attention particulière a été portée quant à la construction du matériel (i.e., vidéo sur la dyslexie), afin de le rendre attrayant sur une population d'étudiants, notamment en psychologie, il est difficile de recréer l'enjeu d'une tâche d'apprentissage en contexte réel. Les étudiants n'y ont peut-être pas investi la motivation nécessaire à ce type de tâche, parce qu'elle ne s'inscrit pas dans leur parcours universitaire ou personnel. De la même manière, on peut penser que les individus, lorsqu'ils apprennent à partir d'un document vidéo, hors contexte expérimental, utilisent d'autres ressources : prise de notes, relecture de la vidéo sur plusieurs jours voire semaines, discussion entre pairs, recherche Internet... La prise en compte d'autant de variables est particulièrement compliquée à mettre en œuvre, et cette étude a nécessité de contraindre les apprenants à un type de stratégie, le visionnage de la vidéo uniquement. Les conclusions établies à partir d'une étude qui

paraît relativement éloignée d'une situation réelle d'apprentissage semblent donc à prendre avec précaution. De nouvelles études pourraient ainsi être menées dans un contexte d'apprentissage plus écologique, en postant une vidéo sur une plateforme de MOOC par exemple, avec un test sanctionnant l'apprentissage à la fin du module.

5. Conclusion

L'hypothèse d'un traitement actif de l'étayage dans la tâche d'apprentissage par le simple biais d'une consigne d'apprentissage est donc invalidée dans cette étude. Les participants n'ont en effet pas utilisé l'étayage en tant que modèle conceptuel pour favoriser la construction d'un modèle mental pertinent et ainsi potentiellement améliorer leurs performances en mémorisation et/ou compréhension. Néanmoins, la qualité de l'activité de révision semble améliorée lorsque les deux niveaux d'étayage sont présents conjointement. Il s'agit donc de trouver par quel(s) moyen(s) engager l'apprenant pour qu'il bénéficie du modèle conceptuel de la vidéo proposé par l'étayage. L'une des pistes peut être de trouver comment inciter l'apprenant à traiter l'étayage, afin qu'il se familiarise avec et puisse l'intégrer mentalement. Si l'étayage est utile, lors d'une tâche de RI par exemple, comme démontré dans les études précédentes, on peut alors penser qu'il sera plus facilement traité activement par la suite lors d'une activité d'apprentissage.

ÉTUDE 5 : Effets conjoints de l'étayage et des questions de RI sur l'apprentissage en vidéo³⁸

1. Introduction et hypothèses

Les résultats de la quatrième étude n'ont pas démontré d'effet de l'étayage en vidéo sur les performances d'apprentissage. Une hypothèse explicative a donc été formulée : les participants n'auraient pas pu internaliser le modèle conceptuel fourni (i.e., l'étayage). La consigne d'apprentissage ne suffirait donc pas à rendre l'apprenant suffisamment actif dans son activité au sein de la vidéo. L'étayage ne serait pas traité de manière adaptée pour être bénéfique à l'apprentissage. Plus précisément, l'étayage ne serait appréhendé que tardivement par l'apprenant (i.e., lors de l'activité de révision uniquement). Par conséquent, cette cinquième étude vise à tester l'hypothèse selon laquelle l'ajout d'une tâche spécifique accompagnant l'apprentissage pourrait inciter l'apprenant à explorer et traiter plus efficacement le document avant même son activité d'apprentissage.

L'effet des tests insérés au cours de l'apprentissage a été étudié dans la littérature (e.g., Fernandez & Jamet, 2017 ; Martins, 1995 ; Reynolds et al., 1979). Dans l'étude de Fernandez et Jamet (2017), les participants ayant été confrontés à des tests réguliers ont eu significativement un meilleur score au questionnaire final d'apprentissage que les participants sans test : le bénéfice était de près de 40 %. Selon certains auteurs (Martins, 1995 ; Reynolds et al., 1979), les questions posées pendant l'apprentissage permettraient d'augmenter l'attention de l'apprenant sur le document étudié. De plus, les tests auraient également un effet sur la métacognition des apprenants. En d'autres termes, ils fourniraient des informations aux apprenants sur

³⁸ Un article est actuellement en préparation à partir des études 4 et 5 de cette thèse : Cojean, S., & Jamet, E. (en préparation). The effect of scaffolding video-based environment during learning: information seeking as a necessary condition.

ce qu'ils ont appris, ce qu'ils savent et ce qu'ils ne savent pas (Martins, 1995 ; Roediger III, Putnam, & Smith, 2011). Ainsi, la présentation régulière de tests avec *feedback* permet aux apprenants d'avoir un *monitoring* plus pertinent, c'est-à-dire qu'ils mettraient en place des stratégies plus adaptées à leurs besoins. Les tests permettraient ainsi un meilleur apprentissage des informations testées ou bien même non testées (Roediger III et al., 2011). Une récente méta-analyse conduite par Adesope et ses collaborateurs (2017) a conclu sur une efficacité globale de l'apprentissage avec tests. Globalement, leur effet positif s'expliquerait par le fait qu'ils permettent d'amener les apprenants à s'engager dans un processus de RI, ce qui améliore la rétention de l'information présentée dans le document (Fernandez & Jamet, 2017).

L'effet des tests peut également être bénéfique lorsque les questions sont posées avant la phase d'apprentissage en elle-même. Par exemple, dans une étude menée par Mayer et ses collaborateurs (2003, étude 3), il s'agissait pour les participants d'apprendre le contenu d'un document multimédia interactif qui décrivait le fonctionnement d'un moteur électrique. Deux conditions expérimentales étaient comparées. Dans la première, les apprenants étaient confrontés à des questions avant la phase d'apprentissage. Ils n'étaient pas obligés d'y répondre avant le test final. Dans la seconde condition, les participants apprenaient le document directement, sans question au préalable. Le but des questions présentées avant l'apprentissage était d'engager l'apprenant cognitivement, et de l'orienter pour la phase suivante lors de son apprentissage. Le test de compréhension passé à la fin de l'apprentissage a révélé une différence significative entre les conditions : les participants ayant eu des questions sur le document avant de l'apprendre ont été meilleurs que les autres. La taille d'effet de 0.83 montre un bénéfice important de la présence de questions avant l'apprentissage³⁹. Les auteurs expliquent cet effet notamment par une hypothèse d'auto-explication : les individus apprendraient mieux lorsqu'ils doivent fournir des explications pendant leur apprentissage. Ils seraient aussi fortement incités à manipuler le matériel présenté.

³⁹ Selon les données de cette étude, il s'agit probablement d'un delta de Glass, correspondant à un d de Cohen de 0.92.

Dans le cadre spécifique de l'apprentissage en vidéos, Hannafin et Hughes (1986) soulignaient d'ailleurs déjà il y a plus de vingt ans l'intérêt potentiel d'insérer des questions préalables à l'apprentissage d'un document vidéo. Dans une étude menée par Carpenter et Toftness (2017), l'ajout de questions en rapport avec le matériel présenté en vidéo, mais posées avant la phase d'apprentissage, s'est montré bénéfique pour les performances finales d'apprentissage. Les apprenants devaient répondre aux questions sans visionner la vidéo, et ne recevaient aucun *feedback* sur leurs réponses. Les effets positifs de ces questions se sont retrouvés pour les informations en lien avec les questions, mais également pour les informations non questionnées avant l'apprentissage.

En résumé, l'effet des tests, c'est-à-dire des questions posées à l'apprenant nécessitant une analyse du contenu, est notamment de pouvoir favoriser un traitement actif du matériel présenté en engageant l'apprenant dans un processus de RI (Fernandez & Jamet, 2017). L'insertion d'une activité de RI directement avant l'apprentissage devrait ainsi être bénéfique à l'apprentissage. Par ailleurs, les études précédentes menées sur les effets de l'étayage lors d'une tâche de RI (voir études 1, 2 et 3) ont permis de montrer un bénéfice de la présence conjointe de segmentation et de structuration sur les performances en RI. Cependant, les individus, s'ils étaient meilleurs en RI, ne l'étaient pas par la suite sur une tâche visant à mesurer la qualité de leur modèle mental. L'absence de consigne d'apprentissage était alors formulée comme hypothèse explicative. Dans le cas où une phase d'apprentissage serait annoncée à la suite de la tâche de RI, les apprenants bénéficiant d'étayage devraient pouvoir chercher à internaliser ce modèle conceptuel. De la même manière, les résultats de la précédente étude (voir étude 4) semblent indiquer que les apprenants n'avaient pas conscience de l'utilité de l'étayage. La tâche de RI permettrait alors de mettre en lien l'étayage avec le contenu de la vidéo, et de rendre la présence de l'étayage utile à la réalisation de la tâche. L'apprenant, rendu actif par la RI, serait ainsi plus à même d'utiliser les ressources mises à sa disposition, telles qu'un étayage, afin de s'en servir comme modèle conceptuel qui formerait une base pour la construction d'un modèle mental pertinent. L'apprentissage serait alors de meilleure qualité.

Ainsi, deux effets majeurs sont attendus dans cette étude : la réalisation d'une tâche de RI avant un apprentissage devrait être bénéfique à l'apprentissage, et cet

effet devrait être d'autant plus fort en présence d'étayage. L'interaction de l'étayage et de la RI devrait permettre à l'apprenant de s'engager dans un apprentissage actif, facilité par la présence d'un modèle conceptuel (i.e., étayage) utile pour un modèle mental pertinent. Les hypothèses suivantes ont ainsi été formulées :

Hypothèse 1. Phase de RI :

Conformément aux résultats obtenus lors des études précédentes (voir études 1, 2 et 3), l'activité de RI, lorsqu'elle a lieu avant l'apprentissage, devrait être facilitée lors de la présence d'étayage dans la vidéo. Plus précisément, les participants bénéficiant d'étayage devraient trouver davantage de réponses (Hypothèse 1a), être plus rapides à chaque question de RI (Hypothèse 1b), et enfin plus pertinentes dans le premier clic au sein de la vidéo pour chaque question de RI (Hypothèse 1c).

Hypothèse 2. Apprentissage :

Les participants ayant répondu à des questions de RI avant la phase d'apprentissage devraient être meilleurs dans leur apprentissage que les participants ayant directement réalisé la tâche d'apprentissage. Lorsqu'ils ont bénéficié de questions de RI, les participants avec étayage devraient être meilleurs que les participants ayant appris la vidéo sans étayage. Cet effet devrait se retrouver pour les scores de mémorisation, c'est-à-dire lors des questions de rappel (Hypothèse 2a), et pour les scores de compréhension, c'est-à-dire lors des réponses aux inférences (Hypothèse 2b).

Hypothèse 3. Activité de révision :

De la même manière que pour l'étude 4, un effet de l'étayage sur l'activité de révision est attendu. L'étayage devrait ainsi faciliter l'accès aux informations à réviser (nombre de clics pour y accéder, Hypothèse 3a), et permettre de cibler de manière plus efficace et pertinente les révisions (temps de consultation plus long, Hypothèse 3b).

Hypothèse 4. Difficulté perçue :

Les questions de RI devraient induire des traitements de l'information supplémentaires. Ainsi, les participants ayant répondu à des questions de RI avant la phase d'apprentissage devraient percevoir la tâche comme ayant été plus difficile que les participants ayant directement réalisé la tâche d'apprentissage. Les participants avec étayage devraient percevoir la tâche comme moins difficile que les participants ayant appris la vidéo sans étayage

2. Méthode

2.1. Participants

103 étudiants de la région de Rennes (83 femmes, 20 hommes) ont participé volontairement à cette étude. La moyenne d'âge des participants était de 19.71 ans (E.-T. = 1.41). Le recrutement a été réalisé au sein de l'Université Rennes 2, notamment par le biais d'affichages. Pour les remercier de leur participation, tous les étudiants ont été gratifiés de deux tickets de cinéma, ou bien d'un bon de quinze euros. L'étude s'est déroulée conformément aux principes de la déclaration d'Helsinki.

2.2. Matériel

La vidéo présentée était identique à celle utilisée pour l'étude 4. Il s'agissait d'une vidéo⁴⁰ de dix-huit minutes environ sur la dyslexie (Moret & Hermann, 2009). Un environnement vidéo adapté à l'étude a été créé autour de cette vidéo. Quatre conditions expérimentales, correspondant à quatre environnements différents, ont ainsi été testées. L'environnement vidéo incluait ou non de l'étayage (sans étayage ou bien les deux niveaux d'étayage utilisés conjointement, c'est-à-dire marqueurs sur la barre de navigation et table des matières), et quatre questions de RI sur la vidéo étaient posées ou non avant la phase d'apprentissage (e.g., « Dans le cadre de la théorie phonologique, quelles sont les trois aires cérébrales impliquées dans la lecture ? » ; voir Annexe 4). Chaque question était limitée à cinq minutes et un

⁴⁰

<https://www.images.inserm.fr/fr/asset/fullTextSearch/search/dyslexie/page/1>

chronomètre affichait le temps restant. Dans toutes les conditions, il était possible pour les participants de se déplacer au sein de la barre de navigation comme ils le souhaitent, et ils pouvaient également mettre en pause la vidéo.

La phase d'apprentissage, commune à toutes les conditions expérimentales, durait trente minutes. Les participants étaient laissés libres dans leur navigation au sein de la vidéo. Au premier clic dans la vidéo, c'est-à-dire au moment où elle démarrait pour la phase d'apprentissage, un chronomètre de trente minutes s'affichait à l'écran pour indiquer au participant combien de temps il lui restait. Au bout des trente minutes, si le participant n'était pas encore passé au questionnaire, la vidéo s'arrêtait automatiquement et le questionnaire d'apprentissage apparaissait. Si le participant souhaitait passer au questionnaire avant la fin des trente minutes d'apprentissage, il pouvait cliquer sur un bouton placé à côté de la vidéo, « passer au questionnaire d'apprentissage ». Le questionnaire présentait des items de difficulté perçue ainsi que des questions de mémorisation et de compréhension.

Tous les participants ont été aléatoirement répartis dans les quatre conditions expérimentales. Dans la **condition contrôle** ($n = 25$), la vidéo était présentée avec une barre de navigation que les participants pouvaient utiliser pour se déplacer au sein de la vidéo. Un bouton pause était également présent. La durée totale de la vidéo, ainsi que le temps de visionnage restant étaient indiqués au-dessus de la barre de navigation. Les participants accédaient directement à la phase d'apprentissage limitée à trente minutes (voir Figure 26). Dans la **condition avec étayage** ($n = 26$), les participants accédaient également directement à la phase d'apprentissage de trente minutes, mais ils bénéficiaient de la vidéo accompagnée d'étayage. La barre de navigation de la vidéo était segmentée en quatorze parties, correspondant aux chapitres de la vidéo. Une table des matières présentant les titres de ces parties était disponible sur le côté de la vidéo. Il était possible de cliquer sur les titres des parties pour y accéder directement dans la vidéo (voir Figure 26). Les participants de la **condition avec RI** ($n = 27$) n'étaient pas exposés en premier lieu à la phase d'apprentissage. Avant d'y accéder, ils devaient répondre à quatre questions de RI, dont la réponse se trouvait au sein de la vidéo. Chaque question de RI était limitée à cinq minutes. Si le participant n'avait pas trouvé la réponse dans les cinq minutes, la question suivante apparaissait automatiquement. Pour chaque question, il était

demandé d'indiquer la réponse ainsi que le moment dans la vidéo où cette réponse était formulée. La vidéo était présentée sans étayage (voir Figure 26). Une fois les quatre questions de RI passées, le participant pouvait passer à la phase d'apprentissage de trente minutes. Il était alors de nouveau confronté à la vidéo, et la phase d'apprentissage se déroulait de manière similaire à celle dans la condition sans étayage. La quatrième condition expérimentale était la **condition avec étayage et RI** ($n = 25$). Dans cette condition, les participants étaient confrontés à quatre questions de RI avant de passer à la phase d'apprentissage. Durant les deux phases (RI et apprentissage), la vidéo était présentée avec de l'étayage (voir Figure 26).

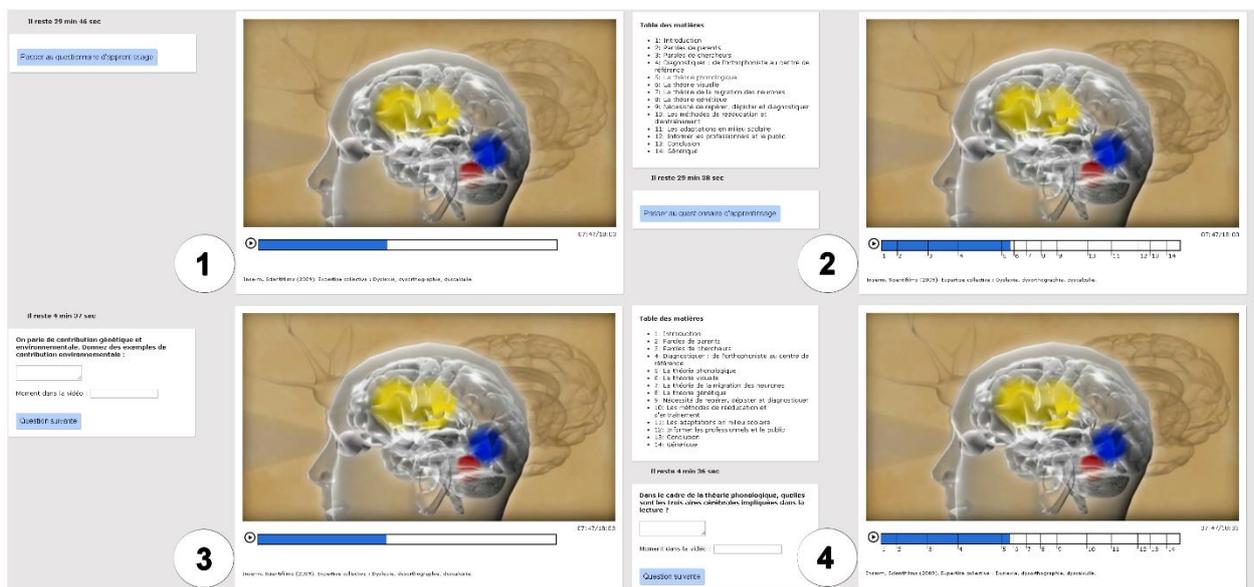


Figure 26. Environnements vidéo dans les conditions contrôle (1), avec étayage (2), avec RI (3) ou avec étayage et RI (4) de l'étude 5

2.3. Mesures

2.3.1. Intérêt préalable et compétence perçue

Avant la tâche d'apprentissage, le participant était informé que le sujet abordé dans la vidéo était la dyslexie. Il lui était alors demandé d'indiquer son niveau d'intérêt préalable quant à ce sujet, ainsi que sa compétence perçue. Pour cela, il devait se placer sur une échelle de Likert en 11 points selon son degré d'intérêt (« Sur une échelle de 0 à 10, à quel point ce sujet vous intéresse-t-il ? ») et de compétence perçue

(« Sur une échelle de 0 à 10, à quel point vous estimez-vous compétent sur ce sujet ? »). L'objectif était de s'assurer que les scores des individus sur ces deux variables étaient équitablement répartis parmi les quatre conditions expérimentales.

2.3.1. Performance en RI

De la même manière que pour les études précédentes centrées sur l'activité de RI (voir études 1,2 et 3), trois variables ont été mesurées concernant la tâche de RI : réussite à la tâche, temps de réponse et pertinence du premier clic. La réussite à la tâche était notée soit 1 si la réponse correcte était trouvée, soit 0 pour une réponse incorrecte ou en l'absence de réponse. Le temps de réponse correspondait au temps mis par le participant avant d'indiquer sa réponse. Pour chacune des quatre questions posées, le temps de recherche était limité à cinq minutes. Enfin, la pertinence du premier clic était mesurée comme le nombre de segments (i.e., chapitres, la vidéo présentant quatorze chapitres) entre le segment cible (où se trouvait la réponse à la question posée) et le segment dans lequel le participant a effectué son premier clic de RI. Le score de pertinence correspondait donc plutôt à un taux d'erreur, le score le plus petit signifiant que le participant était plus proche de la réponse qu'avec un score plus élevé.

2.3.2. Temps d'apprentissage

Le temps d'apprentissage pour chaque participant, c'est-à-dire le temps qu'il a passé sur la vidéo durant la phase d'apprentissage, a été mesuré grâce à un logiciel enregistreur d'écran⁴¹. Avec la limitation de temps de la phase d'apprentissage, le temps maximum possible était donc de trente minutes. Le chronomètre démarrait à partir du moment où le participant lançait la vidéo, et s'arrêtait lorsqu'il cliquait sur le bouton pour passer au questionnaire d'apprentissage. Dans le cas où il aurait cliqué d'abord sur pause dans la vidéo avant de passer au questionnaire, c'est le moment où la vidéo s'est arrêtée qui a compté comme marqueur de fin de l'apprentissage. Pour les participants réalisant une tâche de RI avant la phase d'apprentissage, les trente minutes d'apprentissage débutaient à la fin de la tâche de RI.

⁴¹ Flashback by Blueberry Software, <https://www.flashbackrecorder.com/fbhome/>

2.3.3. Score d'apprentissage

Le questionnaire proposé après la phase d'apprentissage était constitué notamment de quinze questions visant à mesurer la mémorisation du document (e.g., « Quelles sont les 4 théories sur l'origine de la dyslexie développées dans cette vidéo ? » ; voir Annexe 3) et de cinq questions sur la compréhension (e.g., « Certains dyslexiques traitent chaque mot comme un nouveau mot en assemblant les lettres une à une. D'autres ne lisent les mots qu'en les adressant directement au lexique et ont beaucoup de mal à lire de nouveaux mots en les assemblant. À quelles théories chacun de ces comportements correspond-il ? » ; voir Annexe 3). Pour les questions de compréhension, contrairement aux questions de mémorisation, les réponses n'étaient pas explicitement dites dans la vidéo. Pour y répondre, il était nécessaire pour le participant qu'il mette en relation différents éléments de la vidéo.

2.3.4. Activité de révision

De la même manière que pour l'étude 4, la pertinence de l'activité de révision par les participants a été évaluée de deux façons : le nombre de clics dans la barre de navigation pour retrouver l'information à réviser, et le temps passé à réviser chaque information. Ces données ont pu être extraites grâce au logiciel enregistreur d'écran. Les activités de révision correspondaient aux clics dans la barre de navigation après un visionnage complet de la vidéo une première fois.

2.3.5. Difficulté perçue

La variable de difficulté perçue a été évaluée avec les mêmes items que pour les études précédentes (voir études 1, 2, 3 et 4). Ces items ont été adaptés ou inspirés d'études dans la littérature (Kraft et al., 2005 ; Trafimow et al., 2002). Trois items de difficulté perçue ont été utilisés (e.g., « Pour moi, apprendre cette vidéo était difficile » ; voir Annexe 5). Pour chaque item, les participants devaient indiquer sur une échelle de Likert en 7 points le degré avec lequel ils étaient d'accord avec l'affirmation proposée. L'alpha de Cronbach correspondant était de $\alpha = .84$.

2.4. Procédure

La salle d'expérimentation permettait d'accueillir jusqu'à trois personnes en même temps. Chaque participant était installé sur un poste d'ordinateur. Il remplissait alors un formulaire de consentement en deux exemplaires (i.e., un pour eux et un pour l'expérimentateur). Un pré-questionnaire leur était ensuite présenté. Il s'agissait pour le participant de répondre sur son degré d'intérêt préalable sur le sujet de la dyslexie, ainsi que sur son degré de compétence perçue. La consigne distribuée ensuite pour les groupes contrôle et avec étayage (donc sans RI) précisait qu'il allait s'agir d'une phase d'apprentissage de maximum trente minutes, et que des questions leur seraient posées par la suite (voir Annexe 12). Concernant les participants dans les groupes avec RI et avec étayage et RI, la consigne indiquait que la tâche d'apprentissage était précédée d'une tâche de RI (voir Annexe 13). L'expérimentateur s'assurait que les participants avaient correctement compris la consigne, et qu'ils n'avaient pas de question avant de commencer. Puis il lançait l'enregistreur d'écran. Les participants pouvaient débiter l'apprentissage ou la RI dès qu'ils le souhaitaient en cliquant sur le bouton « démarrer » pour lancer les questions de RI, ou bien directement sur la vidéo pour lancer le chronomètre de trente minutes pour la phase d'apprentissage. Pour les participants devant répondre à des questions de RI, le bouton de réponse à la dernière question n'était plus « question suivante » mais « passer à la phase d'apprentissage ». Ils basculaient alors sur le même environnement vidéo que les participants sans RI. Les participants devaient porter un casque audio pour entendre la vidéo.

Lorsqu'ils jugeaient avoir terminé leur apprentissage, les participants pouvaient cliquer sur le bouton « passer au questionnaire d'apprentissage ». Un message s'affichait alors à l'écran, dans lequel il était demandé au participant de prévenir et d'attendre l'expérimentateur. Celui-ci lançait le post-questionnaire, dans lequel les participants devaient répondre à des items de difficulté perçue, puis répondre à des questions de mémorisation et de compréhension. Les vingt questions, comprenant quinze questions de mémorisation et cinq questions de compréhension, étaient proposées dans un ordre aléatoire afin d'éviter les effets d'ordre (Benton & Daly, 1991). Il était précisé aux participants d'indiquer une réponse à chaque question, mais que cette réponse pouvait également être « je ne sais pas ». Enfin, des questions sur l'âge et le sexe du participant terminaient le post-questionnaire. La procédure est résumée

en Figure 27. Après avoir répondu au post-questionnaire, l'expérimentateur gratifiait chaque participant de deux tickets de cinéma, ou d'un bon d'achat de quinze euros, pour les remercier de leur participation à l'étude.

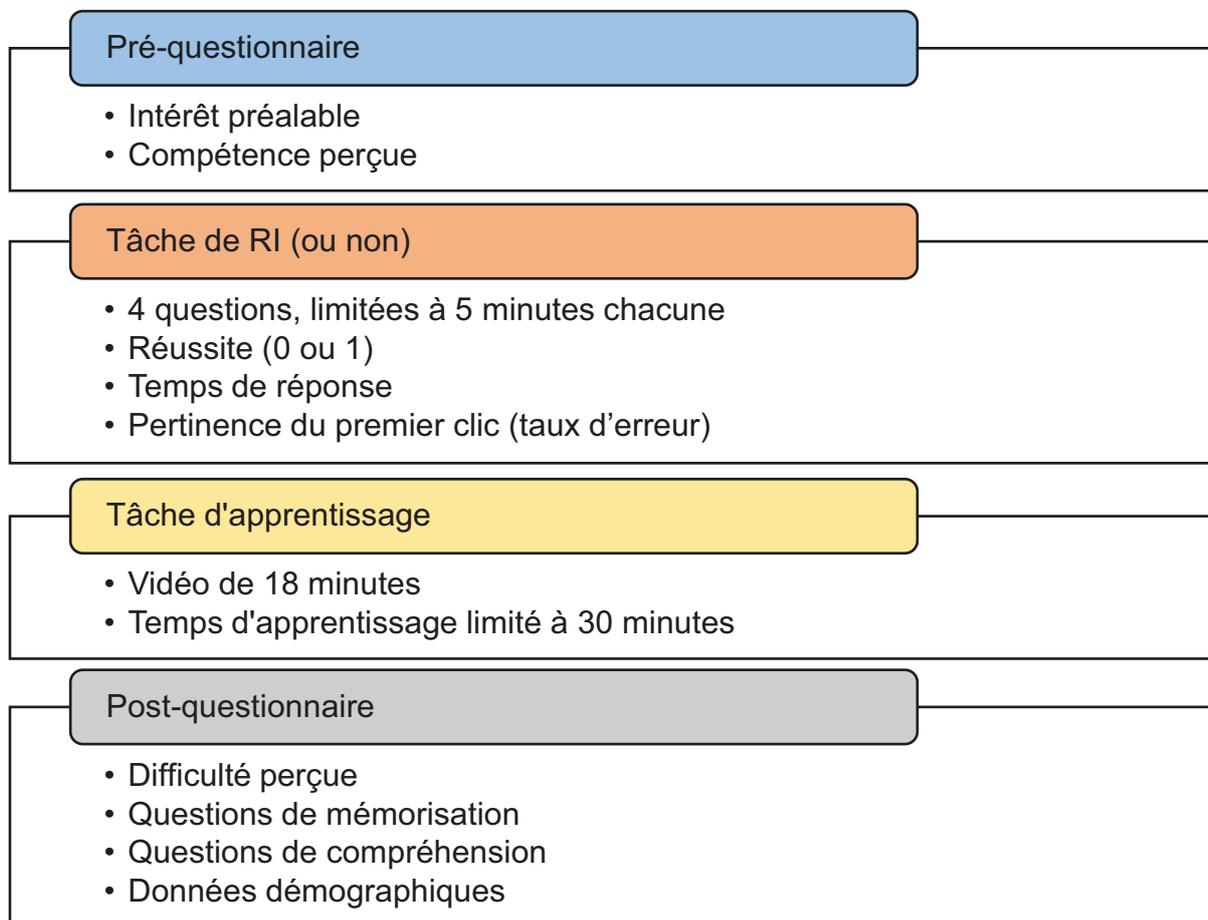


Figure 27. Procédure de l'étude 5

3. Résultats

3.1. Variables contrôles (pré-questionnaire)

Des ANOVA ont été réalisées sur les variables d'intérêt préalable et de compétence perçue, afin de s'assurer que les groupes ne différaient pas entre eux avant le début de l'expérimentation. Les résultats ne montrent aucune différence significative entre les groupes, que ce soit pour l'intérêt préalable, $F(3, 99) = 0.06$, $p = .983$, $\eta^2_p = .00$, ou pour la compétence perçue, $F(3, 99) = 0.06$, $p = .980$, $\eta^2_p = .00$. Les statistiques descriptives sont données dans le Tableau 17.

Tableau 17. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 5)

Condition	Intérêt préalable		Compétence perçue	
	Moy.	E.-T.	Moy.	E.-T.
Contrôle	6.88	1.17	3.16	1.14
Avec étayage	6.73	1.31	3.23	1.88
Avec RI	6.81	1.66	3.33	1.73
Avec étayage et RI	6.76	1.39	3.32	1.82

3.2. Tâche de RI

Comme pour les études précédentes centrées sur l'activité de RI (études 1, 2 et 3), des modèles linéaires mixtes ont été utilisés afin de prendre en compte la non-indépendance des données liée à la multiplicité des questions posées à chaque participant (Field et al., 2012 ; Gueorguieva & Krystal, 2004). Cette méthode statistique permet de tester l'effet d'une variable par comparaison de modèles imbriqués tout en tenant compte du caractère répété de la mesure (e.g., Baayen et al., 2008). Chaque modèle testé incluait les effets aléatoires de la question et du participant.

Pour chacune des variables liées à la RI, un modèle nul (M0) a été comparé à un modèle incluant la variable étayage (M1). Le modèle le plus satisfaisant des deux a ensuite été comparé à un modèle incluant un effet additif du rang de la question posée (M2) et à un modèle incluant un effet d'interaction du rang de la question (M3).

Seuls les participants de deux conditions expérimentales ont été confrontés à une tâche de RI avant l'apprentissage : ceux de la condition avec RI, ainsi que ceux en condition avec étayage et RI. Les réponses de chaque participant à chaque question étant notées soit 0 soit 1, il s'agissait donc de données binomiales et des régressions logistiques ont été utilisées pour traiter la variable de réussite à la tâche (Field et al., 2012). Les variables de temps de réponse et de pertinence du premier clic ont été analysées à l'aide de régressions linéaires.

3.2.1. Réussite à la tâche

Les résultats ont montré un effet significatif de la condition (i.e., de la présence d'étayage) sur la réussite à la tâche, $\chi^2(1, N = 208) = 4.11, p = .043$, mais pas d'effet additif, $\chi^2(1, N = 208) = 3.65, p = .056$, ni d'effet d'interaction du rang de la question avec la condition, $\chi^2(1, N = 208) = 3.16, p = .075$. Le modèle M1 a donc été considéré comme étant le plus approprié pour rendre compte des données observées. Les données descriptives (voir Figure 28) indiquent que les participants de la condition avec étayage semblent avoir globalement mieux réussi, c'est-à-dire trouvé davantage de réponses correctes, que les participants n'ayant pas bénéficié d'étayage durant leur RI.

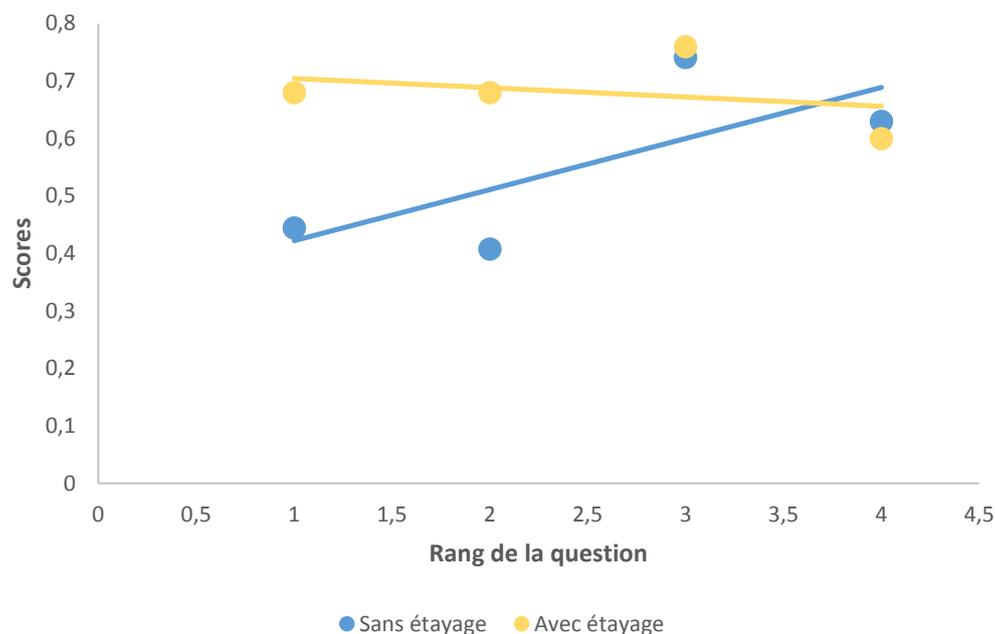


Figure 28. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 4) pour l'étude 5

3.2.2. Temps de réponse

Comme pour les études précédentes (voir études 1,2 et 3), les échecs ont été ôtés des données pour l'analyse des temps de réponse. Les échecs correspondaient aux temps de réponse de 300 secondes, autrement dit lorsque les participants n'avaient pas trouvé la réponse dans la limite de cinq minutes. Un test du Chi-2 a été réalisé pour déterminer si le nombre d'échecs différait selon les conditions (voir

Tableau 18 pour des données descriptives). Les résultats indiquent une différence non significative, $\chi^2(1, N = 208) = 1.01, p = .316$.

Tableau 18. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 5)

Condition	Nombre d'échecs	% (condition)	% (total)
Avec RI	9	8.33	4.33
Avec étayage et RI	4	4.00	1.92

Les résultats ont montré un effet significatif de l'étayage sur le temps de réponse, $\chi^2(1, N = 195) = 4.03, p = .045$, ainsi qu'un effet additif, $\chi^2(1, N = 195) = 9.12, p = .003$, et un effet d'interaction du rang de la question, $\chi^2(1, N = 195) = 5.84, p = .016$. Le modèle M3 a donc été considéré comme étant le plus approprié pour rendre compte des données observées. Les données descriptives (voir Figure 29) indiquent que les participants ayant bénéficié d'étayage sont plus rapides lors des questions de RI, mais que la différence avec les participants sans étayage semble s'atténuer au cours de la tâche.

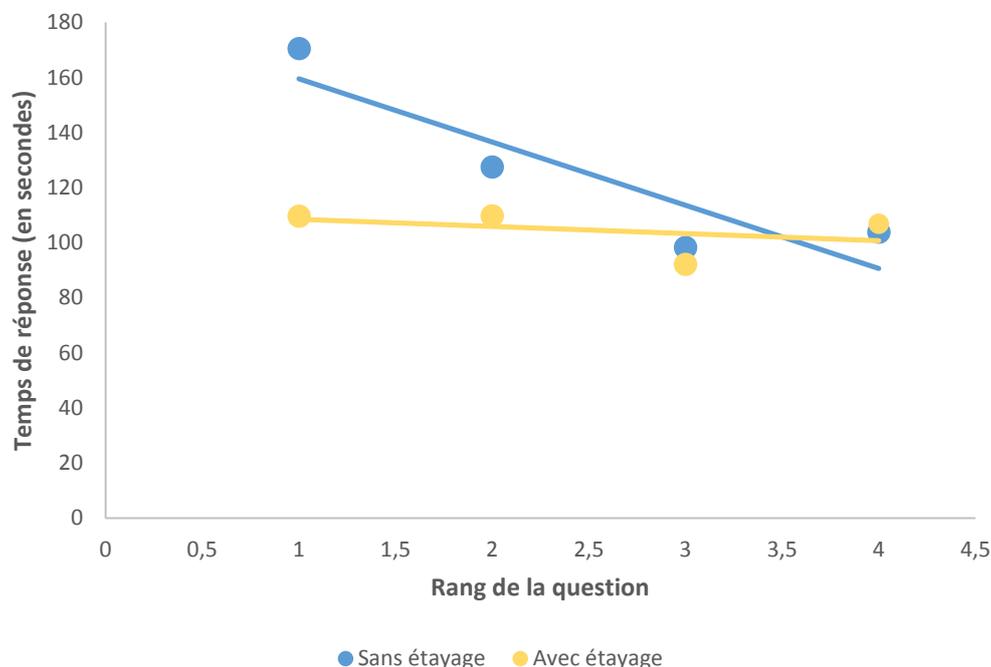


Figure 29. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 4) pour l'étude 5

3.2.3. Pertinence du premier clic

Les résultats ont montré un effet significatif de l'étayage sur la pertinence du premier clic, $\chi^2(1, N = 208) = 7.78, p = .005$, ainsi qu'un effet additif, $\chi^2(1, N = 208) = 9.39, p = .002$, et un effet d'interaction du rang de la question avec la condition, $\chi^2(1, N = 208) = 4.26, p = .039$. Le modèle M3 a donc été considéré comme étant le plus approprié pour rendre compte des données observées. Les données descriptives (voir Figure 30) montrent que les participants avec étayage ont eu des taux d'erreur plus faibles (i.e., ont été plus pertinents) sur leurs premiers clics que les participants sans étayage, mais cette différence semble s'atténuer au fur et à mesure des questions.

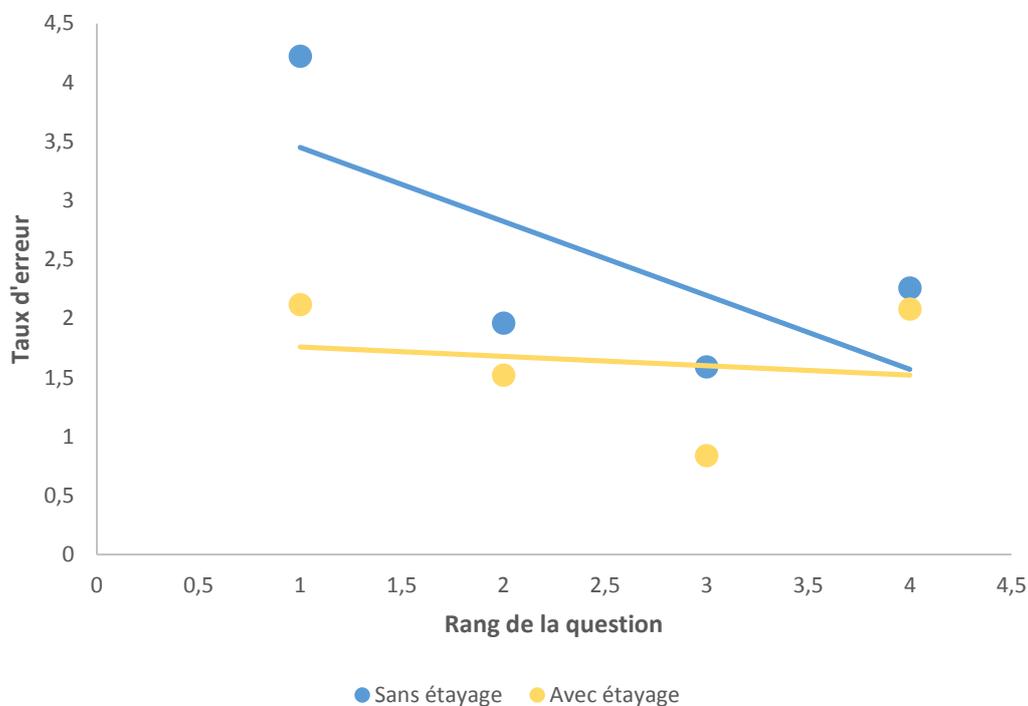


Figure 30. Taux d'erreur moyens des participants en RI sur le premier clic selon la condition expérimentale et le rang de la question (de 1 à 4) pour l'étude 5

3.3. Tâche d'apprentissage

3.3.1. Temps d'apprentissage

Afin de s'assurer que les temps d'apprentissage ne différaient pas selon les conditions expérimentales, une ANOVA en 2 x 2 a été réalisée. Les résultats n'ont pas

révélé d'effet principal, ni de l'étayage, $F(1, 99) = 1.83$, $p = .179$, $\eta^2_p = .02$, ni de la RI, $F(1, 99) = 0.32$, $p = .571$, $\eta^2_p = .00$, ni d'effet d'interaction, $F(1, 99) = 1.46$, $p = .229$, $\eta^2_p = .02$ (voir Tableau 19 pour les données descriptives)⁴².

3.3.2. Score d'apprentissage

Le score d'apprentissage a été mesuré selon deux types de questions : des questions de rappel et des questions d'inférences. Trois scores ont donc été analysés : le score moyen de mémorisation, le score moyen de compréhension, ainsi que le score moyen total d'apprentissage.

Concernant le score de mémorisation, l'ANOVA en 2 x 2 réalisée n'a pas révélé d'effet principal, ni de l'étayage, $F(1, 99) = 1.97$, $p = .163$, $\eta^2_p = .02$, ni de la RI, $F(1, 99) = 2.24$, $p = .138$, $\eta^2_p = .02$, mais a révélé un effet d'interaction significatif, $F(1, 99) = 4.00$, $p = .048$, $\eta^2_p = .04$. Les participants dans la condition avec étayage et RI ont eu de meilleurs scores de mémorisation que les participants des trois autres conditions expérimentales (voir Tableau 19 pour les données descriptives).

Concernant le score de compréhension, l'ANOVA en 2 x 2 n'a pas révélé d'effet principal, ni de l'étayage, $F(1, 99) = 0.49$, $p = .487$, $\eta^2_p = .01$, ni de la RI, $F(1, 99) = 0.59$, $p = .443$, $\eta^2_p = .01$, ni d'effet d'interaction, $F(1, 99) = 1.80$, $p = .183$, $\eta^2_p = .02$ (voir Tableau 19 pour les données descriptives).

⁴² Concernant les participants ayant réalisé une tâche de RI avant la phase d'apprentissage, le temps total d'exposition à la vidéo peut également être calculé. Il s'agit alors d'ajouter au temps d'apprentissage le temps passé en RI (i.e., la somme des temps de réponse aux quatre questions de RI). Pour les participants dans la condition avec RI, le temps d'exposition est en moyenne de 1926.27 secondes (E.-T. = 373.92), et pour les participants dans la condition avec étayage et RI, le temps moyen d'exposition est de 1954.76 secondes (E.-T. = 323.69).

Tableau 19. Données descriptives pour les variables d'apprentissage (temps d'apprentissage et scores de mémorisation et de compréhension) en fonction de la condition expérimentale (étude 5)

Condition	Temps d'apprentissage (en s.)		Score de mémorisation (sur 15)		Score de compréhension (sur 5)	
	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.
Contrôle	1464.52	239.70	8.51	1.97	2.20	0.97
Avec étayage	1470.62	226.49	8.27	1.72	2.06	1.11
Avec RI	1371.00	310.86	8.31	3.00	1.76	1.01
Avec étayage et RI	1505.12	285.58	9.75	1.45	2.18	1.16

3.3.3. Activité de révision

De la même manière que pour l'étude 4, deux types de données permettaient de se renseigner sur l'activité de révision réalisée par les apprenants : le nombre de clics réalisés dans la barre de navigation après un premier visionnage de la vidéo, et le temps moyen passé à visionner l'information lors de cette tâche (i.e., temps total de révision divisé par le nombre de clics).

Une ANOVA en 2 x 2 réalisée sur le nombre de clics a montré un effet principal de l'étayage, $F(1, 99) = 13.74$, $p < .001$, $\eta^2_p = .12$, ainsi qu'un effet de RI, $F(1, 99) = 5.47$, $p = .021$, $\eta^2_p = .05$, mais pas effet significatif d'interaction, $F(1, 99) = 0.41$, $p = .526$, $\eta^2_p = .00$. L'étayage semble permettre de diminuer le nombre de clics, alors que la RI l'augmenterait (voir données descriptives dans le Tableau 20).

Une ANOVA en 2 x 2 réalisée sur le temps total de révision (après un premier visionnage de la vidéo) n'a pas montré d'effet principal, ni de l'étayage, $F(1, 99) = 1.64$, $p = .203$, $\eta^2_p = .02$, ni de la RI, $F(1, 99) = 0.07$, $p = .796$, $\eta^2_p = .00$, ni d'effet d'interaction, $F(1, 99) = 0.78$, $p = .381$, $\eta^2_p = .01$ (voir données descriptives dans le Tableau 20).

Concernant les temps moyens par révision, l'ANOVA en 2 x 2 a révélé un effet principal de l'étayage, $F(1, 99) = 17.30$, $p < .001$, $\eta^2_p = .15$, mais pas d'effet significatif de la RI, $F(1, 99) = 2.33$, $p = .130$, $\eta^2_p = .02$, ni d'effet d'interaction, $F(1, 99) = 0.07$,

$p = .793$, $\eta^2_p = .00$. Les participants ayant réalisé la tâche d'apprentissage en présence d'étayage ont passé davantage de temps lors de chaque révision que les participants sans étayage (voir données descriptives dans le Tableau 20).

Tableau 20. Nombre de révisions effectuées, temps total et temps moyen par révision en fonction de la condition expérimentale (étude 5)

Condition	Nombre de révisions (clics)		Temps total de révision (en s.)		Temps moyen par révision (en s.)	
	Moy.	E.-T.	Moy.	E.-T.	Moy.	E.-T.
Contrôle	25.28	25.58	373.32	239.60	27.03	31.85
Avec étayage	11.15	12.46	393.04	239.57	87.02	108.98
Avec RI	39.30	33.06	343.07	261.57	9.67	9.49
Avec étayage et RI	19.16	17.62	449.68	258.80	62.45	80.22

3.3.4. Difficulté perçue

Une ANOVA en 2 x 2 a été conduite sur les scores de difficulté perçue lors de la tâche d'apprentissage. Les résultats ont montré un effet principal de la RI, $F(1, 99) = 4.94$, $p = .029$, $\eta^2_p = .05$, mais pas d'effet de l'étayage, $F(1, 99) = 0.24$, $p = .624$, $\eta^2_p = .00$, ni d'effet d'interaction entre les deux niveaux d'étayage, $F(1, 99) = 1.18$, $p = .280$, $\eta^2_p = .01$. Les participants ayant dû réaliser une tâche de RI avant la phase d'apprentissage ont indiqué avoir perçu la tâche d'apprentissage comme étant plus difficile que les participants sans RI (voir données descriptives dans le Tableau 21).

Tableau 21. Données descriptives pour la variable difficulté perçue en fonction de la condition expérimentale (étude 5)

Condition	Moy.	E.-T.
Contrôle	3.04	1.36
Avec étayage	2.67	0.88
Avec RI	3.32	1.42
Avec étayage et RI	3.48	1.24

4. Discussion

Cette cinquième étude avait pour objectif d'étudier les effets conjoints de l'étayage en vidéo (i.e., segmentation et structuration) et de la RI lors d'une tâche d'apprentissage. L'hypothèse principale avancée était la suivante : l'ajout d'une tâche de RI avant l'apprentissage engagerait les apprenants dans un traitement actif de la vidéo, et que cet effet serait d'autant plus bénéfique lorsque la vidéo serait présentée avec de l'étayage.

Sur les performances en tâche de RI pour les apprenants qui y ont été confrontés, les résultats indiquent tous un effet positif de l'étayage. Les participants ayant pu en bénéficier ont ainsi trouvé davantage de réponses correctes (Hypothèse 1a), plus rapidement (Hypothèse 1b) et grâce à des premiers clics plus pertinents (Hypothèse 1c). De plus, un effet d'interaction avec l'ordre des questions, c'est-à-dire au fur et à mesure des quatre questions de RI, a été observé pour les variables de temps de réponse et de pertinence du premier clic. Cet effet indique que les participants sans étayage sont moins efficaces que ceux avec étayage, mais que leur performance semble s'améliorer au cours du temps, jusqu'à rejoindre celle des participants avec étayage. L'**hypothèse 1** est donc globalement validée.

Concernant les scores d'apprentissage, les analyses statistiques ont révélé un effet significatif d'interaction entre l'étayage et la RI sur les scores aux quinze questions de mémorisation (i.e., questions de rappel), mais aucun effet significatif n'apparaît sur les scores aux cinq questions de compréhension (i.e., inférences). L'**hypothèse 2a** sur la mémorisation est donc en partie validée, parce que seul l'effet d'interaction est significatif, mais l'**hypothèse 2b** sur la compréhension est quant à elle non validée.

Les résultats de l'étude ne montrent aucune différence significative sur le temps moyen dédié à l'apprentissage vidéo. Pour rappel, le temps était limité à trente minutes. Les participants ont en moyenne passé 24 minutes et 11 secondes sur la tâche d'apprentissage. Pour les mêmes raisons que lors de l'étude 4, la durée de la vidéo de dix-huit minutes ainsi que la limite de temps d'apprentissage ont pu contraindre l'apprentissage entre un temps minimal et un temps maximal d'apprentissage. Cependant, pour les participants ayant réalisé une tâche de RI avant la phase d'apprentissage, le temps total d'exposition est en moyenne de 32 minutes. Ce temps d'exposition plus long n'est toutefois pas forcément associé à un apprentissage de meilleure qualité : les participants avec RI mais sans étayage n'ont pas eu de meilleur score que les participants sans RI. Sans étayage, la RI semble donc être préjudiciable à l'apprentissage, qui devient plus long et moins efficient.

Concernant l'activité de révision, les résultats ont démontré qu'il n'y avait pas de différence significative sur le temps total alloué à la révision selon les conditions expérimentales. En revanche, il est apparu que les participants bénéficiant d'étayage ont réalisé significativement moins de clics dans la barre de navigation pour accéder à l'information qu'ils souhaitaient réviser : leur navigation serait donc plus précise. Un effet de la RI est également ressorti : les apprenants ayant effectué une tâche de RI avant la phase d'apprentissage ont réalisé davantage de clics lors de leur activité de révision. La tâche de RI aurait pu potentiellement amorcer les comportements de RI et donc inciter les apprenants à effectuer davantage de clics. L'**hypothèse 3a** est donc seulement partiellement validée. Pour les temps moyens de consultation lors de l'activité de révision, les résultats ont montré un effet de l'étayage. Les apprenants bénéficiant de segmentation et de structuration ont globalement passé davantage de temps lors de chaque clic dans la barre de navigation à inspecter l'information délivrée par la vidéo. Ces données supposent donc que la qualité (i.e., la pertinence) des révisions a été améliorée grâce à la présence d'étayage. L'**hypothèse 3b** est donc validée, et en cohérence avec les résultats précédents obtenus lors de l'étude 4.

L'analyse des scores de difficulté perçue a révélé un effet significatif de la présence de questions de RI avant la phase d'apprentissage. Les participants ayant effectué les deux tâches (i.e., RI et apprentissage) ont déclaré avoir ressenti davantage de difficulté à les effectuer que les participants n'ayant été confrontés qu'à

la seule tâche d'apprentissage. L'**hypothèse 4** est ainsi validée pour moitié, l'effet de l'étayage n'ayant pas été démontré dans cette étude.

En conclusion, l'effet majeur observé sur l'apprentissage est un effet d'interaction entre étayage et RI. La présence d'étayage ne favoriserait pas à elle seule la mémorisation ni la compréhension (voir également les résultats de l'étude 4). De la même manière, l'ajout d'une phase de RI au sein de la vidéo avant la tâche d'apprentissage n'est pas bénéfique si la vidéo ne présente pas d'étayage. Deux conditions doivent donc être réunies pour observer une amélioration du score d'apprentissage : la vidéo doit présenter un étayage (ici, de micro- et de macro-niveau, c'est-à-dire une segmentation et une structuration du contenu), et les apprenants doivent être confrontés à une tâche de RI avant l'apprentissage. Dans cette étude, quatre questions composaient la tâche de RI, par rapport à neuf dans les études précédentes centrées sur la tâche de RI (voir études 1, 2 et 3). Cette réduction du nombre de questions a été choisie afin que les participants ne soient pas confrontés à une expérimentation trop longue. En effet, la réalisation d'une tâche de RI composée de quatre questions, puis une phase d'apprentissage limitée à trente minutes, et enfin l'écriture des réponses au questionnaire d'apprentissage pouvaient prendre jusqu'à une heure et demie de passation par participant. Doubler le nombre de questions en RI aurait donc encore rallongé ce temps. Néanmoins, les précédentes études de cette thèse montrant un effet d'interaction entre l'étayage de la vidéo et le temps ont été réalisées avec neuf questions, ce qui permettait de voir un pattern intéressant d'amélioration de la performance au cours du temps pour les participants sans étayage. Le fait de passer de neuf à quatre questions aurait pu limiter les effets de l'étayage en RI et masquer cette amélioration au cours du temps. Les résultats de cette étude montrent que le même pattern de performance apparaît bien avec un nombre plus faible de questions.

Un effet bénéfique de la présence d'une tâche de RI avant la phase d'apprentissage était attendu. En effet, les précédents travaux sur le *testing effect* et l'apprentissage actif (Adesope et al., 2017 ; Fernandez & Jamet, 2017 ; Mayer, 2004) laissaient à penser que la manipulation de la vidéo de manière active (i.e., lors d'une tâche de RI par exemple) engagerait l'apprenant dans un traitement plus profond du

contenu du document. Cet effet n'apparaît pas comme systématique dans cette étude. Plus précisément, cet effet ne se retrouve que lorsque la vidéo présente un étayage (ici, de double niveau). La présence de segmentation et de structuration du document semble donc être une condition à l'effet bénéfique de la RI avant l'apprentissage. La réalisation d'une tâche de RI sans étayage n'est donc pas efficace, peut-être car trop coûteuse.

De plus, sans RI, l'étayage de la vidéo n'apporte pas de bénéfice à l'apprentissage (voir également étude 4). L'hypothèse formulée à la suite de l'étude 4 était que la consigne d'apprentissage ne suffisait pas ce que les apprenants s'engagent dans un apprentissage actif, et donc qu'ils puissent internaliser l'étayage présenté (i.e., le modèle conceptuel de la vidéo). La présence d'une tâche de RI devait permettre aux apprenants d'être actifs, et les effets positifs observés dans la condition expérimentale où étayage et RI sont présents vont dans le sens de l'hypothèse formulée. En effet, lorsque l'apprenant est rendu actif dans son apprentissage, en manipulant la vidéo au cours d'une tâche de RI avant l'apprentissage par exemple, il semble internaliser le modèle conceptuel lorsqu'il est présent, et cette internalisation favorise les performances d'apprentissage. Les résultats issus de l'analyse des données sur les stratégies de révision des apprenants soutiennent également cette hypothèse. En effet, sans RI, l'étayage permet bien de favoriser des comportements de révision plus pertinents (voir également étude 4), mais seule la condition proposant étayage et RI conjointement a permis d'améliorer les performances d'apprentissage. Le traitement de l'information serait ainsi rendu plus profond grâce à la tâche de RI, et accompagnerait les comportements de révision rendus efficaces avec l'étayage.

L'effet de l'étayage et de la RI sur l'apprentissage se retrouve de manière plus détaillée sur les questions de mémorisation, mais pas de compréhension. L'hypothèse d'un modèle mental favorisé grâce à l'internalisation du modèle conceptuel est donc globalement soutenue. En effet, le modèle mental jouerait un rôle dans les performances en mémorisation en facilitant le stockage et la récupération des informations issues de la vidéo. L'absence d'effet sur les inférences signifierait que l'objectif d'améliorer la compréhension de la vidéo n'a pas été atteint. La compréhension serait en effet liée à un niveau d'élaboration différent de la mémorisation : elle dépend de la qualité de la représentation mentale, associée

notamment à l'utilisation des connaissances antérieures de l'apprenant (voir modèle de Mayer, 2014a ; Introduction théorique 2.1.2). Une autre explication peut être apportée quant à l'absence d'effet sur la compréhension au contraire de la mémorisation. Dans l'étude menée par Waern (1981), deux groupes expérimentaux étaient comparés. Dans le premier, la consigne était d'apprendre le document par cœur (i.e., consigne de mémorisation). Dans le second groupe, l'objectif donné était de comprendre le document. Les résultats ont montré que les performances en mémorisation et compréhension différaient dans chaque groupe, chaque facette de l'apprentissage étant favorisée lorsque la consigne la mettait en avant. Les résultats de l'étude présentée ici laissent donc à penser que lorsque la consigne énoncée est une consigne générale d'apprentissage, les apprenants se focaliseraient uniquement sur la mémorisation du document.

Cette étude présente cependant quelques limites. En effet, les participants ayant effectué une tâche de RI avant l'apprentissage ont passé en moyenne davantage de temps à naviguer dans la vidéo. Les temps d'apprentissage sont sensiblement similaires entre les quatre conditions expérimentales, mais dans deux conditions, les participants ont également passé du temps à chercher des réponses aux questions de RI, ils ont donc pu visionner la vidéo plus longtemps. Cependant, l'absence d'effet sur les scores d'apprentissage de la RI en condition « avec RI » permet de penser que ce n'est pas le temps total passé sur la vidéo qui compte, mais davantage le temps d'apprentissage. À temps d'apprentissage relativement égaux, seule la condition « avec étayage et RI » se distingue des trois autres.

De plus, la mesure de l'apprentissage, si elle semble faire consensus dans la littérature, peut difficilement être mise en œuvre de façon idéale. Des items visant à mesurer la mémorisation et la compréhension ont été créés, mais l'absence d'effet sur les items d'inférences peut s'expliquer par une difficulté à proposer des items demandant de réelles inférences pour y répondre. Les items de compréhension sont supposés nécessiter un transfert de connaissances de la part de l'apprenant, depuis la situation apprise dans la vidéo jusqu'à une situation nouvelle. De fait, les réponses aux questions d'inférences ne doivent pas avoir été formulées dans la vidéo. Cinq

questions de compréhension ont été créées pour cette étude, elles pourront être repensées lors d'une reproduction de l'étude.

Enfin, comme pour l'étude 4, le caractère écologique de la tâche d'apprentissage en vidéo avec ou sans étayage pourrait être perfectionné par une nouvelle série d'études : effet de la prise de note, apprentissage à long terme, insertion de la vidéo dans le programme d'un cours réel proposé dans le cursus des étudiants, etc.

5. Conclusion

En conclusion, l'étayage dans une vidéo ne semble pas avoir d'effet sur l'apprentissage lorsqu'il est ajouté seul. De la même manière, proposer aux apprenants une tâche de RI avant la phase d'apprentissage n'améliore pas leur performance. Il apparaît cependant dans cette étude que l'association d'une phase de RI à une vidéo proposant un double étayage soit bénéfique à la mémorisation de l'information. Concrètement, une recommandation pédagogique peut être formulée à la suite de cette étude : lors de la création de matériel vidéo pédagogique, le concepteur devrait veiller à insérer un étayage de double niveau (i.e., segmentation et structuration), et lors de la phase d'apprentissage, l'enseignant devrait inciter les apprenants à s'engager activement dans la tâche en leur proposant une activité de RI. Cette activité peut être courte, quatre questions semblent suffisantes pour une vidéo d'une vingtaine de minutes. De nouvelles pistes expérimentales peuvent être envisagées en tenant compte de contextes plus appliqués, tels que l'apprentissage en vidéos sur des plateformes de MOOC.

Discussion générale

L'utilisation de supports vidéo semble être devenue de plus en plus fréquente dans les environnements pédagogiques, notamment au sein des MOOC (Delen et al., 2014 ; Yadav et al., 2015). Deux activités majeures peuvent se dérouler au sein d'une vidéo pédagogique : de l'apprentissage et/ou de la RI. Ces activités semblent liées, la localisation de l'information intervenant en général avant le processus d'apprentissage (Merkt & Schwan, 2014 ; Rouet & Coutelet, 2008). Dans les deux cas, la qualité du modèle mental de l'utilisateur/apprenant joue un rôle fondamental dans les performances. Lors d'une activité de RI, selon le modèle développé par Sharit et ses collaborateurs (2008), la représentation mentale de la vidéo serait la première étape du processus de localisation, avant la planification et l'exécution des actions de navigation et de recherche. Lors d'un apprentissage, selon le modèle de l'apprentissage multimédia de Mayer (2014a), l'apprenant sélectionne, organise et intègre les informations qu'il juge pertinentes au sein d'un modèle mental unifié et prenant en compte les connaissances antérieures.

La construction d'une représentation mentale utile peut être aidée avec la présence d'un modèle conceptuel. Celui-ci va fournir une représentation du système étudié, ici la vidéo, qui va servir de base à l'élaboration d'un modèle mental pertinent (Borgman, 1986 ; Carmel et al., 1992 ; Norman, 1983 ; Seel, 2003). L'une des manières de proposer un modèle conceptuel d'une vidéo est d'étayer le document au niveau de deux contraintes spécifiques à ce format : l'aspect transitoire de l'information (Höffler & Leutner, 2007 ; Merkt et al., 2011) et le manque de représentation de structure (D. Zhang et al., 2006). En premier lieu, la **segmentation** de la barre de navigation de la vidéo (i.e., étayage de micro-niveau) devrait permettre à l'utilisateur/apprenant de contrôler le rythme de la vidéo et de naviguer au niveau de la micro-structure de la vidéo (Merkt et al., 2011). La **structuration** (i.e., étayage de macro-niveau), quant à elle, consiste à présenter les grandes parties de la vidéo, et permettrait la navigation au niveau de la macro-structure de celle-ci (Merkt et al., 2011). La présence de ces deux niveaux d'étayage dans une vidéo devrait ainsi favoriser la représentation mentale qu'ont les utilisateurs/apprenants de la vidéo. Cela

devrait alors aider à améliorer les performances à la fois lors des activités de RI et d'apprentissage.

1. L'effet de l'étayage vidéo lors d'une activité de RI

Une première partie des études menées au cours de cette thèse concernait l'effet de l'étayage vidéo lors d'une activité de RI. L'hypothèse majeure était que la présence d'**étayage** (i.e., segmentation et/ou structuration) serait favorable aux performances des utilisateurs en RI. Un questionnement a également émergé sur la qualité du modèle mental des utilisateurs lors d'une activité de RI avec ou sans étayage. Si l'étayage peut agir à la manière d'un modèle conceptuel qui guide le processus de localisation, aide-t-il pour autant à la construction d'un **modèle mental** plus pertinent du document étudié, à savoir la vidéo ? Enfin, il s'agissait de comparer différents formats de présentation de l'étayage, pour éviter un effet d'attention divisée entre les deux niveaux d'étayage. L'idée était notamment de favoriser les **liens référentiels** à l'aide d'interactivité (i.e., possibilité de cliquer sur l'un des niveaux pour accéder au deuxième correspondant) ou encore de *pop-ups*.

Les résultats des trois études menées indiquent globalement un effet **positif** de l'étayage sur la RI. Cet effet se retrouve sur la réussite à la tâche (i.e., les participants avec étayage trouvent davantage de réponses correctes aux questions de RI), les temps de réponse (i.e., les participants avec étayage sont plus rapides pour trouver les réponses aux questions de RI), la pertinence du premier clic (i.e., les participants avec étayage ont des premiers clics plus proches de la réponse attendue lors de chaque question de RI), et enfin sur la difficulté perçue de la tâche (i.e., les participants avec étayage trouvent la tâche de RI moins difficile à réaliser que les participants sans étayage). Un effet du temps au cours de la tâche a également été observé lors des trois études concernant les variables de temps de réponse et de pertinence : les participants sans étayage semblent moins performants au début de la tâche (i.e., à la première question de RI) qu'à la fin (i.e., à la neuvième question). La différence entre les conditions avec ou sans étayage semble s'atténuer au cours de la tâche, au point que les performances pourraient être considérées comme similaires lors de la dernière question. Ces résultats complètent ceux de Merkt et Schwan (2014), qui avaient

échoué à démontrer un effet positif de l'étayage vidéo sur les performances en RI. Dans leur étude, les auteurs ont considéré, au sein d'une vidéo pédagogique, l'ajout d'une table des matières et d'un index comme un étayage de macro-niveau, et les possibilités d'arrêter ou de se déplacer dans la barre de navigation comme un étayage de micro-niveau. Les résultats ont échoué à démontrer un effet de l'étayage de micro-niveau ainsi que de la table des matières sur les performances en RI. Néanmoins, la barre de navigation n'était pas segmentée dans cette étude, et la table des matières n'était pas affichée en même temps que la vidéo.

La première étude de cette thèse apporte un éclairage supplémentaire sur les effets spécifiques de chaque niveau d'étayage testé (i.e., micro- ou macro-niveau, c'est-à-dire segmentation ou structuration). L'effet bénéfique de la présence de l'étayage n'apparaît que lorsque les deux niveaux d'étayage sont présentés conjointement. La présence d'un seul niveau, sans le second, ne suffirait pas pour améliorer les performances en RI. Dans cette étude, la seule condition qui se détache des trois autres est la condition avec un double niveau d'étayage, et les participants en bénéficiant ont des performances stables (et meilleures que dans les autres conditions) tout au long de la tâche. Des effets spécifiques à chaque niveau ont tout de même été observés, en adéquation avec les hypothèses formulées. En effet, d'après la littérature, la segmentation (i.e., micro-étayage) devrait permettre de contrôler le flot d'informations délivrées par la vidéo en naviguant plus précisément au sein du document (e.g., Merkt et al., 2011 ; Spanjers et al., 2012 ; Wouters et al., 2007). La structuration de la vidéo (i.e., macro-étayage), quant à elle, devait permettre de faire apparaître l'organisation sémantique du contenu du document (Amadiou & Salmerón, 2014). Les résultats de cette première étude ont confirmé l'effet positif de la segmentation sur le contrôle perçu par les participants sur la tâche de RI en vidéo, ainsi que le bénéfice de la structuration sur le rappel des chapitres constituant le document vidéo. En d'autres termes, la segmentation aiderait la **navigation** au sein de la vidéo étudiée et la structuration renseignerait sur le contenu **sémantique** de la vidéo. Ces deux aspects semblent ainsi complémentaires, et leur interaction permettrait une amélioration des performances en RI au sein d'un environnement vidéo.

La deuxième étude avait pour objectif d'évaluer la qualité du modèle mental des utilisateurs après une activité de RI avec ou sans étayage. Un modèle mental pertinent est censé aider l'utilisateur à prédire les comportements du système étudié (Borgman, 1986 ; Norman, 1983 ; Staggers & Norcio, 1993). Ici, il était demandé de localiser des informations dans la vidéo précédemment visionnée sans avoir de *feedback*. Les résultats ont montré que les participants qui n'avaient pas bénéficié d'étayage lors de la phase de RI avaient été moins performants en RI mais meilleurs lors de la tâche impliquant le modèle mental. Leur score de localisation s'est avéré être globalement plus précis que celui des participants qui avaient réalisé la RI avec étayage. L'étayage lorsqu'il est présent agirait donc comme un guide externe à la réalisation des tâches de RI mais ne serait pas automatiquement internalisé pour aider à la construction d'un modèle mental plus pertinent. À l'inverse, sans étayage, les performances en RI sont plus faibles, mais les utilisateurs construiraient eux-mêmes un modèle mental adapté.

Par ailleurs, l'étude 1 laisse à penser que l'étayage n'est bénéfique à la RI que si les deux niveaux sont utilisés conjointement. La présentation de la vidéo avec segmentation et structuration peut être envisagée de différentes manières, et ce sont plusieurs formats qui ont été comparés. L'objectif de la troisième étude était de définir un format de présentation qui améliorerait peut-être encore les performances en RI, ou bien qui serait déjà moins coûteux cognitivement pour l'utilisateur. Une présentation avec de l'interactivité entre les deux niveaux d'étayage, ainsi qu'un format en *pop-ups* avec ou sans présence permanente du sommaire ont été comparés à des formats plus classiques avec ou sans étayage. Dans la littérature, l'apport du format *pop-up* a notamment fait l'objet d'études en apprentissage (e.g., Bétrancourt & Bisseret, 1998 ; Erhel & Jamet, 2006), et a par ailleurs montré un aspect positif sur les performances. Si le format sans étayage se distingue une nouvelle fois des quatre autres avec double niveau d'étayage, aucune différence significative n'est ressortie entre les conditions avec étayage sur les performances en RI. Il est à supposer que le double étayage améliore, peu importe le format de présentation, les performances en RI (i.e., réussite à la tâche, temps de réponse, pertinence du premier clic) de la manière la plus optimale possible. Les performances, stables dans le temps au fur et à mesure des questions, atteindraient toutes le seuil maximal de performance possible sur une telle tâche. Néanmoins, le format *pop-up* se distingue des autres formats sur la variable de difficulté perçue : les participants ayant bénéficié de l'étayage présenté de cette

manière ont perçu la tâche de RI comme étant moins difficile que les autres participants. Ces résultats sont en accord avec les précédents dans la littérature portant sur des tâches d'apprentissage (e.g., Bétrancourt & Bisseret, 1998 ; Erhel & Jamet, 2006). Le format *pop-up* semble donc préférable aux autres, en ce qu'il permet une performance similaire mais réalisée avec moins de difficulté.

En conclusion, la présence d'étayage est bénéfique aux performances en RI en vidéo en termes de réussite à la tâche, temps de réponse et pertinence des premiers clics de recherche. Cet effet n'apparaît que lorsque les deux niveaux d'étayage sont utilisés conjointement, c'est-à-dire lors de la présence de segmentation et de structuration. Des effets spécifiques (i.e., aide à la navigation et identification du contenu sémantique) à ces deux niveaux existent et apparaissent complémentaires. Le pattern de réponse principal se dégagant est une stabilité dans l'efficacité de l'étayage sur les performances en RI, bien que cet avantage ne mène pas à une internalisation du modèle conceptuel présenté pour participer à la construction d'un modèle mental particulièrement fiable. Sans étayage, ou bien lorsqu'un seul des niveaux d'étayage est présent, les performances sont plus faibles, mais elles s'améliorent au cours de la tâche, potentiellement grâce à la construction – coûteuse cognitivement – d'un modèle mental pertinent. Ces résultats ont été consistants sur les quatre études impliquant une tâche de RI (i.e., études 1, 2, 3 et 5), et surtout auprès de deux vidéos différentes (la vidéo de l'étude 5 n'étant pas la même que celle utilisée dans les trois premières études). D'autre part, le format de présentation des deux niveaux d'étayage ne semble pas modifier l'effet sur les performances. Cependant, dans le cas d'une tâche de RI, la modalité de format en *pop-ups* réduirait la difficulté perçue de la tâche par les utilisateurs.

Dans le modèle décrivant les processus impliqués dans la localisation de l'information proposé par Sharit et ses collaborateurs (2008), la base d'une RI performante est la qualité de la représentation mentale du contenu. En d'autres termes, un modèle mental pertinent du contenu de la vidéo serait essentiel à la mise en œuvre des processus impliqués dans la localisation des informations. Les résultats des études menées dans le cadre de cette thèse permettent d'envisager que la base du processus de localisation ne soit pas un modèle mental – forcément interne – de la

vidéo, mais plutôt une représentation qui peut tout aussi bien rester externe. L'étayage proposé viendrait jouer le rôle de représentation du contenu de la vidéo, et aiderait alors les processus subséquents de planification et d'exécutions des actions de RI.

Puisque les processus engagés lors d'une activité de RI peuvent être impliqués dans la mise en œuvre d'une activité d'apprentissage, il semble par la suite nécessaire de s'intéresser aux effets de l'étayage en vidéo lors d'un apprentissage. L'étayage, lorsqu'il est présent, n'est pas utilisé comme base pour un modèle mental de meilleure qualité lors d'une activité de RI. La modification de la consigne, orientée non plus vers un objectif de RI, mais vers un objectif d'apprentissage, devait permettre d'engager les apprenants dans un processus actif d'internalisation du modèle conceptuel proposé.

2. L'effet de l'étayage vidéo lors d'une activité d'apprentissage

La seconde partie des études conduites au cours de cette thèse concernait l'effet de l'étayage vidéo lors d'une activité d'apprentissage. L'hypothèse majeure était que l'étayage en vidéo était bénéfique à l'apprentissage, car l'objectif d'apprentissage, à la différence d'un objectif de RI, permettrait à l'apprenant de s'engager dans un **traitement actif** de l'information. L'étayage agirait alors en tant que modèle conceptuel à la base de la construction d'un modèle mental pertinent du contenu de la vidéo. Dans le cas où la simple présence d'une consigne d'apprentissage ne suffirait pas à permettre un apprentissage actif bénéfique aux performances, une nouvelle hypothèse a été formulée : l'introduction d'**une phase de RI** avant la phase d'apprentissage renforcerait globalement l'apprentissage et ferait apparaître les effets positifs de l'étayage en vidéo.

Deux études ont été menées dans cette optique. Les résultats de la première n'indiquent aucun effet significatif de l'étayage sur l'apprentissage. Les conditions expérimentales présentant l'environnement vidéo accompagné de micro-étayage (i.e., segmentation), de macro-étayage (i.e., structuration), d'étayage de double niveau ou encore sans étayage, n'ont pas conduit à des différences quant aux performances en apprentissage (incluant des questions de mémorisation et de compréhension). Néanmoins, des effets de l'étayage ont été démontrés concernant la qualité des

activités de révision. La segmentation permettrait de faciliter la navigation au sein de la vidéo afin de retrouver l'information à réviser (Merkt et al., 2011 ; Wouters et al., 2007), et une utilisation conjointe de la segmentation et de la structuration permettrait une révision globalement plus pertinente. L'apport de la structuration peut être envisagé comme une aide sémantique au ciblage de la révision (Amadiou & Salmerón, 2014). Mais l'activité de révision, si elle est facilitée par l'étayage, ne semble pas suffire à améliorer l'apprentissage. L'hypothèse a donc été faite que la présence d'une consigne d'apprentissage ne suffisait pas à engager les apprenants dans un processus actif de traitement de l'information. De fait, la seconde étude réalisée au sein de cette partie indique un effet d'interaction entre la phase de RI avant l'apprentissage et l'étayage de la vidéo. La mémorisation a été améliorée lorsque les apprenants ont bénéficié à la fois d'étayage et d'une phase de RI. Il semblerait donc que le modèle mental des apprenants ayant bénéficié d'étayage et de RI ait été plus pertinent et mieux organisé : la récupération des informations en mémoire s'en est trouvée facilitée. La segmentation et la structuration (i.e., le modèle conceptuel présenté) de la vidéo auraient ainsi agi comme des guides pour l'organisation de l'information entrante au sein du modèle mental. Au niveau du modèle d'apprentissage proposé par Mayer (2014a), l'étayage optimiserait les processus de sélection mais surtout d'organisation de l'information entrante au niveau du modèle mental (Amadiou & Salmerón, 2014). Enfin, l'analyse des activités de révision ont permis de constater un effet de l'étayage et de la RI sur la navigation au sein de la vidéo, ainsi qu'un effet positif de l'étayage sur la qualité des révisions. Ces résultats confirment l'apport de l'étayage dans l'activité de révision, et plus globalement la nécessité d'un traitement actif supplémentaire pour traiter l'information à un niveau plus profond.

Au sein de la littérature, des études ont été précédemment réalisées sur les relations entre RI et compréhension du document (e.g., Cataldo & Cornoldi, 1998 ; Cataldo & Oakhill, 2000). Néanmoins, il n'est pas précisé par les auteurs de ces études quelles questions ont été posées afin d'évaluer la compréhension du texte par les lecteurs. S'agissait-il de compréhension au niveau du sens de la phrase (i.e., comprendre ce qui a été présenté, saisir le sens) ou bien de la capacité des lecteurs à effectuer des inférences (i.e., réutiliser les connaissances acquises dans un nouveau contexte) ? Ici, les résultats de l'étude 5 ne montrent pas d'effet de la RI sur la compréhension (i.e., score aux questions d'inférence) de la vidéo. Par rapport à

l'étude 4, l'ajout d'une phase de RI permet en revanche d'améliorer les performances en mémorisation, mais la RI semble dans ce cas avoir plutôt agi comme une condition nécessaire à l'utilisation (i.e., l'internalisation) de l'étayage. La RI ne serait pas dans ce contexte la variable à l'origine de l'amélioration des performances d'apprentissage.

En conclusion, la consigne d'apprentissage ne suffit pas à elle seule à engager l'apprenant dans un traitement actif de la vidéo et à faire apparaître les effets bénéfiques de l'étayage sur l'apprentissage. L'ajout d'une tâche spécifique, ici de la RI, avant celle d'apprentissage semble nécessaire au traitement actif du contenu de la vidéo pédagogique. Lors de cette tâche de RI, l'étayage jouerait un rôle bénéfique sur les performances et serait donc bien utilisé par l'apprenant (voir études 1, 2, 3 et 5). Une nouvelle hypothèse pourrait alors être formulée, selon laquelle l'apprenant, lors de la phase d'apprentissage suivant celle de RI, aurait déjà traité l'étayage et constaté son utilité. Cela favoriserait l'effet de l'étayage lorsqu'il est présent et que l'apprentissage a été accompagné d'une tâche de RI. Sans phase de RI, l'hypothèse peut être faite que l'étayage n'est pas utilisé, car aucun avantage à le traiter n'a pu être décelé par l'utilisateur. Finalement, il pourrait s'agir d'une aide métacognitive donnant un *feedback* par exemple sur l'accessibilité de l'information : cela permettrait de fait de réduire le sentiment de confiance trop élevé et néfaste à l'apprentissage (Carpenter & Toftness, 2017 ; Finn & Tauber, 2015).

3. Limites de la thèse

Ce travail, composé de cinq études différentes, comporte néanmoins certaines limites. Trois aspects, en particulier, méritent réflexion : la population étudiée, le matériel d'expérimentation utilisé, ainsi que les choix de mesure de certaines variables.

3.1. Population étudiée

L'ensemble des participants aux cinq études de cette thèse a été recruté au sein de la population étudiante de la ville de Rennes. Les expérimentations ayant pour sujet principal la RI et l'apprentissage dans un contexte pédagogique, le choix a été fait de ne recruter que des étudiants. Il a en effet été supposé que les étudiants avaient l'avantage d'être encore dans une dynamique d'apprentissage scolaire/universitaire,

et parce qu'il s'agit d'une part importante de la population cible des plateformes de MOOC. Néanmoins, la généralisation des résultats issus des études sur une population aussi homogène doit être envisagée avec précaution. Deuxièmement, le choix d'une population étudiante a imposé une contrainte majeure dans la mise en place des études : le recrutement des participants. Il a été difficile de recruter des étudiants volontaires pour participer à des études durant une heure à une heure et demie, qui de surcroît n'avaient pas participé aux études précédentes. En effet, la vidéo présentée étant la même au sein des trois études de RI, et une seconde ayant été utilisée pour les deux études d'apprentissage, il était nécessaire que les participants n'aient pas été déjà confrontés au matériel. De même, une courte explication sur les raisons de l'étude était donnée en fin de passation aux participants qui le désiraient. Ils étaient alors mis au courant des hypothèses sur les effets de l'étayage. Les étudiants ayant participé à l'une des études de RI ne pouvaient donc pas être recrutés à nouveau pour une étude d'apprentissage, les variables mesurées étant similaires. De manière globale, il a été compliqué d'avoir accès à un nombre suffisant d'étudiants pour compléter les expériences. Cela a donc affecté la taille des groupes expérimentaux, qui dans certaines études comportent parfois moins de vingt participants. Troisième contrainte concernant la population d'étudiants, ces derniers peuvent être considérés comme des experts en RI. Il est imaginable en effet qu'ils aient recours à cette pratique régulièrement dans un cadre universitaire et/ou personnel. Il pourrait être intéressant de tester les effets de l'étayage sur des enfants par exemple, qui n'auraient pas encore été confrontés à des demandes de RI dans un environnement vidéo. Enfin, il est également à noter que le recrutement d'étudiants exclusivement amène une nouvelle contrainte : le manque de motivation des participants. Même s'ils étaient volontaires, la situation de passation d'expérience n'est pas la plus écologique possible. Dans un contexte d'utilisation d'une plateforme MOOC ou plus généralement de vidéo pédagogique, les étudiants choisissent la vidéo qu'ils vont regarder, ou bien le moment de le faire, ou encore ils le font parce que cela est nécessaire dans le cadre de leurs études. Les vidéos proposées dans les études de cette thèse ne s'incluaient pas dans le cursus universitaire des participants. Les thèmes abordés, l'eau dans l'Univers et la dyslexie, ont été choisis pour intéresser le plus grand nombre, mais rien n'obligeait les étudiants à traiter de la manière la plus profonde possible le matériel. S'intéresser aux processus d'apprentissage

nécessiterait donc également de prendre en compte la motivation des apprenants à s'engager dans une telle tâche (Mayer, 2014b).

3.2. Matériel

Les deux environnements vidéo différents proposés ont été construits spécifiquement pour ces études. Cependant, pour des raisons de contrôle expérimental, ils n'étaient pas comparables à de réels MOOC. Sur les plateformes dédiées, l'apprentissage est conçu pour être étalé dans le temps, sur plusieurs semaines en général, et les participants sont amenés à visionner plusieurs vidéos. De plus, la qualité de l'apprentissage n'a été mesurée qu'à court terme, c'est-à-dire au cours de la même session que le visionnage de la vidéo. Il aurait été intéressant de mesurer l'apprentissage à long terme dans les études 4 et 5, mais le risque était de constater une importante mortalité expérimentale. Par ailleurs, pour mesurer un apprentissage à long terme, il aurait fallu s'assurer de reproduire un contexte d'apprentissage le plus proche possible d'une situation réelle : les participants auraient dû être autorisés par exemple à prendre des notes lors de la confrontation à la vidéo, et à les consulter de nouveau plus tard. Enfin, seulement deux environnements vidéo différents ont été utilisés pour l'ensemble des cinq études réalisées. Même si les résultats montrent un pattern de réponse similaire concernant l'activité de RI au sein des deux vidéos (voir les études 1-2-3 et 5), la généralisation de l'effet positif de l'étyage sur la RI à l'ensemble des vidéos nécessite de nouvelles études complémentaires.

3.3. Mesures

Au cours des études, différentes variables ont pu être mesurées, et certaines par le biais de questionnaires. Lorsqu'il s'agissait de variables telles que la difficulté perçue, ou bien le contrôle perçu, des échelles de Likert en 7 points ont été utilisées. Cependant, pour mesurer la qualité de l'apprentissage, le choix a été fait de laisser une majorité de questions ouvertes. Deux questions sur 20 impliquaient de donner une réponse parmi trois modalités possibles (i.e., vrai, faux, je ne sais pas), les autres demandant au participant de rédiger l'entièreté de sa réponse. La difficulté dans la rédaction des questions d'apprentissage a surtout été de créer des questions dont la

réponse demandait à l'apprenant de faire une inférence (voir discussion étude 5). Les questions de mémorisation, dont la réponse était explicitement donnée dans la vidéo, ont été plus évidentes à proposer. Pour les questions de compréhension, la réponse ne devait pas avoir été dite dans la vidéo, et impliquer de la part de l'apprenant la mise en relation de plusieurs idées apprises, afin de réutiliser les connaissances dans un nouveau contexte. Cinq questions de compréhension ont été proposées aux apprenants lors des études 4 et 5, mais la pertinence de ces questions peut être remise en question. Impliquent-elles nécessairement un niveau de compréhension élevé de la part de l'apprenant ? Mesurent-elles ce niveau de compréhension ? Discriminent-elles les apprenants entre eux ? Autrement dit, l'absence d'effet significatif sur cette variable signifie-t-il que les apprenants ont un même niveau de compréhension final de la vidéo, ou bien que la différence de niveau de compréhension des apprenants n'a pu être appréhendée correctement ?

Par ailleurs, parmi les variables retenues, le choix a été fait de mesurer la difficulté perçue par les participants après la tâche, et non la charge cognitive ressentie. Trois types de charge cognitive sont généralement distingués (Chanquoy, Tricot, & Sweller, 2007 ; Paas & Sweller, 2014) : la charge cognitive intrinsèque, relative à la complexité de l'information présentée dans le document, la charge cognitive extrinsèque (ou charge inutile), relative à la complexité de la présentation de l'information dans le document, et enfin la charge germane (ou charge pertinente/efficace), relative aux ressources disponibles en mémoire de travail. La charge cognitive intrinsèque n'est pas modifiable à moins de changer le contenu du document. L'objectif recherché dans un contexte de conception de documents multimédias, et donc par exemple de vidéo, est donc de réduire la charge cognitive extrinsèque (i.e., associée au format de présentation du document), afin d'augmenter la charge germane (i.e., les ressources en mémoire de travail pour traiter l'information) pour gérer au mieux la charge cognitive intrinsèque (i.e., le contenu du document). Un tel paradigme aurait pu être utilisé dans les études développées au cours de cette thèse : il aurait été intéressant de mesurer les effets de la présence d'étayage micro-et/ou macro-niveau sur les taux de charge cognitive. Cependant, aucun questionnaire validé de ces dimensions ne semble disponible en français à l'heure actuelle. De plus,

de nombreuses limites conceptuelles et méthodologiques ont été relevées dans la littérature sur la charge cognitive. Parmi ces limites, le problème du manque de clarté du concept a été soulevé, la distinction entre charge cognitive, charge mentale et effort mental étant parfois très ténue (Moreno, 2010). La mesure de la charge cognitive s'avère ainsi compliquée (Chanquoy et al., 2007), et l'une des mesures les plus répandues de la charge cognitive est d'ailleurs un item davantage d'effort mental perçu (de Jong, 2010 ; Paas, 1992). Enfin, le concept de charge cognitive est parfois critiqué pour son aspect difficilement interprétable (Chanquoy et al., 2007) et non-falsifiable, des explications post-hoc seraient en effet possibles quel que soit le sens des résultats (de Jong, 2010). Les questions de difficulté perçue utilisées lors des études de cette thèse ont été inspirées de travaux précédents dans la littérature (Kraft et al., 2005 ; Trafimow et al., 2002). La consistance de cette mesure a été éprouvée lors de chaque étude, et les alphas de Cronbach ont toujours été satisfaisants (entre .80 et .87).

De plus, des hypothèses ont été formulées au cours des différentes études sur la qualité du modèle mental des utilisateurs ou apprenants. Cependant, la difficulté de mesure et d'appréhension du modèle mental est importante. En effet, le modèle mental, par nature, n'est pas directement observable et doit être inféré à travers les scores sur une autre variable (He et al., 2008 ; Rowe & Cooke, 1995 ; Staggers & Norcio, 1993 ; X. Zhang & Chignell, 2001). Dans la littérature, différentes manières de mesurer les modèles mentaux ont été utilisées, s'appuyant par exemple sur les verbalisations des stratégies ou les connaissances sur le sujet étudié (e.g., Azevedo et al., 2004 ; Chi et al., 1994 ; Slone, 2002). Deux questionnements peuvent également être soulevés concernant les modèles mentaux abordés au cours de thèse : s'agit-il de modèles mentaux de la vidéo, ou bien du contenu de la vidéo ? Dans le premier cas, cela reviendrait à considérer la vidéo comme un système. Un modèle mental pertinent permettrait de prédire les états futurs de ce système, c'est-à-dire de percevoir son fonctionnement interne. C'est cette approche qui a été utilisée dans les études portant sur la RI (études 1, 2 et 3), et la prédiction des états futurs consistait en la pertinence de la localisation d'informations au sein de la vidéo : savoir où cliquer pour afficher telle information, etc. Dans les études se focalisant sur l'activité

d'apprentissage (études 4 et 5), le modèle mental pertinent ne serait pas un modèle du système qu'est la vidéo, mais davantage du contenu de la vidéo. Les informations extraites et organisées par l'apprenant seraient alors des informations présentées au sein de la vidéo, et non plus relatives au fonctionnement de celle-ci. L'unicité du concept de modèle mental pose ainsi question, et il serait intéressant d'envisager une approche unifiée afin de prendre en compte tous les aspects et contextes d'utilisation du modèle mental.

Enfin, l'analyse des activités de révision au cours des apprentissages (voir études 4 et 5) ne permet pas d'évaluer de manière satisfaisante les stratégies employées par les apprenants. En effet, seules les données concernant le nombre de clics effectués, le temps total de révision et donc le temps moyen par clic, ont pu être recueillies. L'hypothèse que ces données nous informent sur la qualité de l'activité de révision est donc forte et mériterait d'être accompagnée de mesures complémentaires. Par exemple, l'usage de protocoles de pensée à voix haute (i.e., verbalisations par les apprenants de ce qu'ils pensent en simultané avec leur apprentissage ; Greene et al., 2015) pourrait permettre d'avoir accès au but recherché par l'apprenant lors de chaque clic dans la barre de navigation (e.g., Fernandez & Jamet, 2017 ; Hinostroza, Ibieta, Labbé, & Soto, 2018) : cherche-t-il une information particulière ? Navigue-t-il sans objectif défini ? Sait-il ce qui lui est nécessaire de réviser pour améliorer son apprentissage ? Utilise-t-il l'étayage à sa disposition ? En d'autres termes, des données relatives à la métacognition des apprenants permettraient d'interpréter au mieux les stratégies de révision. En ce sens, évaluer la métacompréhension des individus (e.g., demander à l'individu d'indiquer s'il se sent capable de répondre à une question précise ; Redford, Thiede, Wiley, & Griffin, 2012) fournirait également des éléments pour tester l'effet des questions de RI avant apprentissage en tant que *feedbacks* sur la qualité de leur représentation de la vidéo. En effet, une phase de RI avant apprentissage pourrait permettre à l'apprenant de se positionner quant à ce qu'il connaît ou non de la vidéo, et de réduire le sentiment néfaste de sur-confiance sur ses connaissances (Carpenter & Toftness, 2017 ; Finn & Tauber, 2015). D'autres études (e.g., Anderson & Thiede, 2008 ; Fernandez & Jamet, 2017 ; Redford et al., 2012) dans la littérature ont déjà été menées sur les apports de cartes conceptuelles, tests ou

résumés sur les processus métacognitifs des apprenants. De la même manière, une analyse des mouvements oculaires (*eye-tracking*) permettrait d'étudier la navigation entre les sources d'information (e.g., étayage, contenu de la vidéo) et les stratégies interindividuelles (e.g., Hannus & Hyönä, 1999 ; Hyönä, Lorch, & Kaakinen, 2002 ; Johnson & Mayer, 2012 ; Potocki et al., 2017 ; van Gog & Scheiter, 2010). Des études (Hannus & Hyönä, 1999 ; Hyönä et al., 2002) ont ainsi pu montrer, grâce aux temps de consultation de certaines zones d'intérêt dans l'environnement pédagogique, que les individus pouvaient ne pas traiter les titres d'un texte ou bien passer très peu de temps sur une illustration censée favoriser l'apprentissage. L'usage d'un *eye-tracker* pourrait donc aider à étudier l'utilisation de l'étayage en situation de RI ou d'apprentissage, à approfondir les résultats sur les effets d'attention divisée entre les niveaux d'étayage et à consolider ceux sur l'appréhension de l'étayage avant la phase de révision au cours d'un apprentissage.

4. Perspectives et travaux futurs

Les études réalisées au cours de cette thèse ne permettent pas d'interroger de manière exhaustive les effets de l'étayage vidéo en situation de RI ou d'apprentissage. Les effets spécifiques de chaque type d'étayage (i.e., segmentation et structuration) ont été investigués dans l'étude 4, mais pas dans l'étude 5. Il serait intéressant de comparer des conditions sans étayage, avec étayage de micro-niveau, avec étayage de macro-niveau, ou encore avec double étayage dans un contexte de traitement actif de la vidéo grâce à de la RI avant l'apprentissage. En RI, la comparaison de ces quatre conditions a révélé un effet conjoint des deux niveaux d'étayage, mais également des aspects spécifiques à chaque niveau. Lors d'une activité d'apprentissage, le même pattern de résultats pourrait être attendu. De la même manière, en réponse aux études menées sur la RI, et pour les transférer dans un contexte d'apprentissage, il serait intéressant de tester plusieurs formats de présentation de l'étayage. En RI, le format *pop-up* s'est révélé préférable. Une étude comparant les mêmes conditions expérimentales que dans l'étude 3, mais lors de tâches de RI puis d'apprentissage, pourrait ainsi être menée. Enfin, l'étude 5 teste l'effet de l'étayage associé à une phase de RI avant apprentissage. Si les avantages supposés de la phase de RI peuvent s'apparenter à un *testing effect*, l'effet de l'étayage associé à une phase de RI après apprentissage pourrait être également étudié. En effet, les questions portant sur le

document sont régulièrement posées après la phase d'apprentissage, et dans le cas d'une phase de révision après les questions, celles-ci guideraient le comportement de relecture (e.g., Dirkx, Thoma, Kester, & Kirschner, 2015 ; Fernandez & Jamet, 2017).

Par ailleurs, le projet LIMAH dans lequel s'inscrit cette thèse prévoit de développer un outil transcrivant automatiquement le discours présent dans une vidéo en texte écrit. La présence d'une transcription totale des informations déjà transmises par voix vidéo fait émerger deux axes de réflexion. Premièrement, la RI pourrait être facilitée grâce au texte écrit. De nombreuses études ont déjà été menées sur l'effet d'un macro-étayage au sein d'un document écrit (e.g., Brooks et al., 1983 ; Lorch, 1989 ; Rouet, Vidal-Abarca, Erbou, & Millogo, 2001b ; Sanchez et al., 2001), et il apparaît notamment que l'avantage d'un tel format tiendrait au fait que l'individu peut avoir une vue globale du contenu écrit (van der Meij & van der Meij, 2014). L'organisation de l'information serait alors plus évidente, et la RI aidée grâce à la partie transcrite. Néanmoins, en second lieu, la même information étant présente à la fois en vidéo et à l'écrit, un effet de redondance néfaste à l'apprentissage pourrait apparaître (e.g., Jamet & Le Bohec, 2007 ; Kalyuga et al., 2004).

De plus, l'effet de l'étayage en vidéo démontré dans cette thèse sur l'organisation des informations délivrées peut amener une réflexion sur la qualité de la prise de notes des apprenants. Les travaux précédents centrés sur cette pratique ont permis de montrer qu'il s'agit d'un comportement spontané de la part des apprenants (e.g., Caspi, Gorsky, & Privman, 2005 ; Fiorella & Mayer, 2017). L'effet de génération de notes semble bénéfique à l'apprentissage, mais néanmoins globalement modéré (Bui, Myerson, & Hale, 2013 ; Kobayashi, 2005). En général, les stratégies de prise de notes employées apparaissent différentes dans des documents écrits ou vidéo : un support papier permet notamment à l'apprenant d'écrire directement dans les marges ou à des endroits particuliers du texte qu'il lit (Caspi et al., 2005). La présence d'étayage dans un environnement vidéo pourrait ainsi guider l'apprenant dans une prise de notes organisée et de meilleure qualité. Les effets attendus seraient donc positifs lors d'un apprentissage vidéo, dans un contexte plus écologique en permettant la réalisation de ce comportement spontané.

Par ailleurs, si le format de présentation des deux niveaux d'étayage a été interrogé au cours de l'étude 3 de la présente thèse, d'autres aspects restent à étudier. En s'inspirant des travaux menés par Arguel et Jamet (2009), différents niveaux de granularité de l'étayage pourraient être comparés. L'hypothèse d'un étayage trop fin, c'est-à-dire d'une vidéo découpée en de trop nombreuses parties, qui serait néfaste à l'apprentissage demanderait à être testée.

Enfin, les études conduites au cours de cette thèse mériteraient d'être répliquées dans un nouveau contexte : durant de réels cours proposés sur une plateforme de MOOC par exemple. La mise en place et le contrôle expérimental seraient compliqués à organiser, mais cela permettrait néanmoins d'accéder à un nombre beaucoup plus important de participants et à un contexte d'apprentissage nettement plus écologique.

5. Conclusion

5.1. Le rôle de l'étayage en vidéo

En résumé, l'étayage (i.e., segmentation et structuration) des vidéos pédagogiques peut être bénéfique à la fois aux performances en RI et en apprentissage. Lors d'une activité de RI, l'étayage va agir à la manière d'un modèle conceptuel, c'est-à-dire qu'il servira de représentation externe du contenu de la vidéo. Cette représentation sera la base des processus impliqués dans la réalisation de la RI. Dans une situation d'apprentissage, l'étayage sera également un modèle conceptuel qui cette fois guidera l'organisation en mémoire de travail de l'information entrante. Il aidera alors la récupération en mémoire de cette même information. L'effet positif de l'étayage vidéo sur la mémorisation nécessite cependant un traitement actif de la part de l'apprenant. Ce traitement actif peut être favorisé par le biais d'une tâche de RI avant la phase d'apprentissage par exemple.

Le rôle du modèle conceptuel serait de servir d'organisateur de connaissances (Staggers & Norcio, 1993). L'étayage organiserait donc l'information présentée à l'individu. Ce serait une représentation externe du contenu de la vidéo. Il agirait également comme un modèle conceptuel en présentant à l'apprenant comment son modèle mental doit être organisé pour que l'information puisse être récupérée

aisément. En RI, l'information est récupérée **dans la vidéo** donc l'étayage reste une représentation **externe**. Lors d'un apprentissage, l'information est récupérée **en mémoire** donc l'étayage est **internalisé** dans le modèle mental.

5.2. Implications pédagogiques

L'ensemble des études menées au cours de cette thèse a permis d'identifier quelques préconisations de conceptions des environnements vidéo pédagogiques. Premièrement, la présence d'étayage associé à la vidéo a un effet bénéfique sur les performances en RI des utilisateurs (voir études 1, 2, 3 et 5). Cependant, cet effet n'apparaît que si les deux niveaux d'étayage sont présents : en d'autres termes, si la segmentation et la structuration de la vidéo sont utilisées conjointement. La présence d'un seul de ces niveaux d'étayage a un effet spécifique (sur le contrôle perçu ou le rappel des chapitres) mais la tâche de RI n'est facilitée qu'en présence des deux niveaux d'étayage. D'autre part, il n'y a pas d'internalisation automatique du modèle conceptuel de la vidéo proposé par l'étayage (voir étude 2). En l'absence de consigne d'apprentissage, les individus ne vont pas bénéficier de l'apport d'étayage, celui-ci restera une représentation externe du contenu vidéo.

Lorsque la consigne est une consigne d'apprentissage, l'utilisation à bon escient de la segmentation et de la structuration se fait sous condition : l'apprenant doit être engagé dans un traitement actif de l'information. Il doit finalement prendre conscience de la présence et de l'intérêt de l'étayage. Ajouter une tâche de RI, même si elle est courte, avant de demander un apprentissage permettrait d'inciter l'apprenant à traiter l'étayage en tant que modèle conceptuel de la vidéo (voir études 4 et 5). L'étayage est bénéfique à l'apprentissage dans ce contexte (voir étude 5). Enfin, l'apport des deux niveaux d'étayage sur la RI semble maximal lorsque la structuration de la vidéo rejoint la segmentation sous un format *pop-up*. Le format le plus efficient lors d'une tâche d'apprentissage reste à investiguer.

Les vidéos pédagogiques proposées à l'heure actuelle, sur les plateformes de MOOC par exemple, semblent souffrir d'un manque d'unité de conception (voir

Introduction théorique 1.2.2). La préconisation essentielle qui se dégage des travaux menés au cours de cette thèse serait de structurer en une table des matières et de segmenter la barre de navigation selon ces différentes parties chaque vidéo proposée. Enfin, dans ce contexte, l'ajout d'une phase de RI avant un apprentissage a montré des bénéfices sur les performances en mémorisation des apprenants.

Références

- Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: a meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659-701. <https://doi.org/10.3102/0034654316689306>
- Allen, I. E., & Seaman, J. (2016). Online Report Card: Tracking Online Education in the United States. *Babson Survey Research Group*.
- Amadiou, F., & Salmerón, L. (2014). Concept maps for comprehension and navigation of hypertexts. Dans D. Ifenthaler & R. Hanewald (Éds), *Digital knowledge maps in education* (pp. 41–59). New York, NY : Springer.
- Amadiou, F., Salmerón, L., Cegarra, J., Paubel, P.-V., Lemarié, J., & Chevalier, A. (2015). Learning from concept mapping and hypertext: An eye tracking study. *Educational Technology & Society*, 18(4), 100–112.
- Anderson, M. C. M., & Thiede, K. W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica*, 128(1), 110-118. <https://doi.org/10.1016/j.actpsy.2007.10.006>
- Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior*, 25(2), 354-359. <https://doi.org/10.1016/j.chb.2008.12.014>
- Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge handbook of multimedia learning* (2^e éd., pp. 206–226). (S.I.) : Cambridge University Press.

- Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, 29(3), 344-370. <https://doi.org/10.1016/j.cedpsych.2003.09.002>
- Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition – Implications for the design of computer-based scaffolds. *Instructional Science*, 33(5-6), 367-379. <https://doi.org/10.1007/s11251-005-1272-9>
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development*, 56(1), 45–72.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Baccino, T., & Draï-Zerbib, V. (2015). *La Lecture numérique: 2e édition*. (S.I.) : PUG.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Benton, J. E., & Daly, J. L. (1991). A question order effect in a local government survey. *The Public Opinion Quarterly*, 55(4), 640–642.

- Bétrancourt, M., & Bisseret, A. (1998). Integrating textual and pictorial information via pop-up windows: An experimental study. *Behaviour & information technology*, 17(5), 263–273.
- Bétrancourt, M., Dillenbourg, P., & Montarnal, C. (2003). Computer technologies in powerful learning environments: the case of using animated and interactive graphics for teaching financial concepts. Dans E. De Corte, L. Verschaffel, N. Entwistle, & J. J. G. Van Merriënboer (Éds), *Unravelling basic components and dimensions of powerful learning environments* (pp. 143–157). Oxford, UK : Elsevier.
- Biard, N., Cojean, S., & Jamet, E. (2017). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2017.12.002>
- Bjork, E., Little, J. L., & Storm, B. C. (2014). Multiple-choice testing as a desirable difficulty in the classroom. *Journal of Applied Research in Memory and Cognition*, 3(3), 165-170. <https://doi.org/10.1016/j.jarmac.2014.03.002>
- Bjork, R. (1994). Memory and metamemory considerations in the training of human beings. Dans J. Metcalfe & A. Shimamura (Éds), *Metacognition: Knowing about Knowing* (pp. 185-205). (S.I.) : Cambridge, MA: MIT Press.
- Borgman, C. L. (1986). The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of man-machine studies*, 24(1), 47–64.
- Borgman, C. L., Hirsh, S. G., Walter, V. A., & Gallagher, A. L. (1995). Children's searching behavior on browsing and keyword online catalogs: the Science

- Library Catalog project. *Journal of the American Society for information Science*, 46(9), 663.
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior*, 21(3), 487–508.
- Brooks, L. W., Dansereau, D. F., Spurlin, J. E., & Holley, C. D. (1983). Effects of headings on text processing. *Journal of Educational Psychology*, 75(2), 292.
- Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology*, 105(2), 299-309. <https://doi.org/10.1037/a0030367>
- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98(1), 182-197. <https://doi.org/10.1037/0022-0663.98.1.182>
- Cagiltay, K. (2006). Scaffolding strategies in electronic performance support systems: Types and challenges. *Innovations in education and Teaching International*, 43(1), 93–103.
- Carmel, E., Crawford, S., & Chen, H. (1992). Browsing in hypertext: A cognitive study. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(5), 865–884.
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational psychology review*, 14(1), 5–26.
- Carpenter, S. K., & Toftness, A. R. (2017). The effect of prequestions on learning from video presentations. *Journal of Applied Research in Memory and Cognition*, 6(1), 104-109. <https://doi.org/10.1016/j.jarmac.2016.07.014>

- Carr-Chellman, A., & Duchastel, P. (2000). The ideal online course. *British Journal of Educational Technology*, 31(3), 229–241.
- Caspi, A., Gorsky, P., & Privman, M. (2005). Viewing comprehension: Students' learning preferences and strategies when studying from video. *Instructional Science*, 33(1), 31–47.
- Cataldo, M. G., & Cornoldi, C. (1998). Self-monitoring in poor and good reading comprehenders and their use of strategy. *British Journal of Developmental Psychology*, 16(2), 155–165.
- Cataldo, M. G., & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? An investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of educational psychology*, 92(4), 791.
- Chanquoy, L., Tricot, A., & Sweller, J. (2007). *La charge cognitive: théorie et applications*. Paris : Armand Colin.
- Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, 80, 108-121. <https://doi.org/10.1016/j.compedu.2014.08.015>
- Chen, S. Y., Fan, J.-P., & Macredie, R. D. (2006). Navigation in hypermedia learning systems: experts vs. novices. *Computers in Human Behavior*, 22(2), 251–266.
- Chi, M. T., Leeuw, N., Chiu, M.-H., & LaVanher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, 18(3), 439–477.
- Chun, D. M., & Plass, J. L. (1996). Facilitating reading comprehension with multimedia. *System*, 24(4), 503–519.

- 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(2), 315–324.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J : L. Erlbaum Associates.
- David, P., Song, M., Hayes, A., & Fredin, E. S. (2007). A cyclic model of information seeking in hyperlinked environments: The role of goals, self-efficacy, and intrinsic motivation. *International Journal of Human-Computer Studies*, 65(2), 170–182.
- de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, 38(2), 105-134.
<https://doi.org/10.1007/s11251-009-9110-0>
- Dee-Lucas, D., & Larkin, J. H. (1995). Learning from electronic texts: Effects of interactive overviews for information access. *Cognition and instruction*, 13(3), 431–468.
- Delen, E., Liew, J., & Willson, V. (2014). Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. *Computers & Education*, 78, 312-320.
<https://doi.org/10.1016/j.compedu.2014.06.018>
- DeStefano, D., & LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior*, 23(3), 1616-1641.
<https://doi.org/10.1016/j.chb.2005.08.012>

- Dinet, J., Chevalier, A., & Tricot, A. (2012). Information search activity: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 62(2), 49-62. <https://doi.org/10.1016/j.erap.2012.03.004>
- Dinet, J. (2014). *La recherche d'information dans les environnements numériques*. (S.I.) : ISTE Editions.
- Dirkx, K. J. H., Thoma, G.-B., Kester, L., & Kirschner, P. A. (2015). Answering questions after initial study guides attention during restudy. *Instructional Science*, 43(1), 59-71. <https://doi.org/10.1007/s11251-014-9330-9>
- Dobson, J. L. (2011). Effect of selected « desirable difficulty » learning strategies on the retention of physiology information. *AJP: Advances in Physiology Education*, 35(4), 378-383. <https://doi.org/10.1152/advan.00039.2011>
- Doolittle, P. E., Bryant, L. H., & Chittum, J. R. (2015). Effects of degree of segmentation and learner disposition on multimedia learning. *British Journal of Educational Technology*, 46(6), 1333–1343.
- Doressoundiram, A. (2012, juin). *Du temps, de l'espace et de l'eau. A la recherche de l'eau dans l'Univers*. Repéré à https://www.canal-u.tv/video/canal_uved/du_temps_de_l_espace_et_de_l_eau.9857
- Doyle, J. K., & Ford, D. N. (1998). Mental models concepts for system dynamics research. *System dynamics review*, 14(1), 3–29.
- Dreher, M. J., & Guthrie, J. T. (1990). Cognitive processes in textbook chapter search tasks. *Reading Research Quarterly*, 323–339.

- Eisenberg, M., Johnson, D., & Berkowitz, B. (2010). Information, communications, and technology (ICT) skills curriculum based on the Big6 skills approach to information problem-solving. *Library Media Connection*, 28(6), 24–27.
- Eitel, A., Scheiter, K., & Schüler, A. (2013). How inspecting a picture affects processing of text in multimedia learning. *Applied Cognitive Psychology*, 27(4), 451–461.
- Erhel, S., & Jamet, E. (2006). Using pop-up windows to improve multimedia learning: Spatial integration of textual and illustrated information. *Journal of Computer Assisted Learning*, 22(2), 137-147. <https://doi.org/10.1111/j.1365-2729.2006.00165.x>
- Evans, C., & Gibbons, N. J. (2007). The interactivity effect in multimedia learning. *Computers & Education*, 49(4), 1147-1160. <https://doi.org/10.1016/j.compedu.2006.01.008>
- Fernandez, J., & Jamet, E. (2017). Extending the testing effect to self-regulated learning. *Metacognition and Learning*, 12(2), 131-156. <https://doi.org/10.1007/s11409-016-9163-9>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. (S.I.): SAGE Publications. ISBN-13.
- Finn, B., & Tauber, S. K. (2015). When confidence is not a signal of knowing: How students' experiences and beliefs about processing fluency can lead to miscalibrated confidence. *Educational Psychology Review*, 27(4), 567-586. <https://doi.org/10.1007/s10648-015-9313-7>

- Fiorella, L., & Mayer, R. E. (2017). Spontaneous spatial strategy use in learning from scientific text. *Contemporary Educational Psychology*, 49, 66-79. <https://doi.org/10.1016/j.cedpsych.2017.01.002>
- Fleury, S., & Jamet, E. (2014). Facilitating the comparison of multiple visual items on screen: The example of electronic architectural plan correction. *Applied Ergonomics*, 45(3), 601-607. <https://doi.org/10.1016/j.apergo.2013.08.004>
- Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity. *Learning and Instruction*, 20(3), 216-224. <https://doi.org/10.1016/j.learninstruc.2009.02.021>
- Flores, S., Bailey, H. R., Eisenberg, M. L., & Zacks, J. M. (2017). Event segmentation improves event memory up to one month later. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(8), 1183.
- Frank, B. M., Garlinger, D. K., & Kiewra, K. A. (1989). Use of embedded headings and intact outline with videotaped instruction. *The Journal of Educational Research*, 82(5), 277-281.
- Furlough, C. S., & Gillan, D. J. (2018). Mental models: Structural differences and the role of experience. *Journal of Cognitive Engineering and Decision Making*, 155534341877323. <https://doi.org/10.1177/1555343418773236>
- Ganier, F., & de Vries, P. (2016). Are instructions in video format always better than photographs when learning manual techniques? The case of learning how to do sutures. *Learning and Instruction*, 44, 87-96. <https://doi.org/10.1016/j.learninstruc.2016.03.004>

- Giannakos, M. N., Jaccheri, L., & Krogstie, J. (2014). Looking at MOOCs rapid growth through the lens of video-based learning research. *International Journal of Emerging Technologies in Learning*, 9(1).
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15(4), 313–331.
- Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, 16(6), 511–525.
- Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C., & McVea, M. (2015). Domain-specificity of self-regulated learning processing in science and history. *Contemporary Educational Psychology*, 42, 111-128. <https://doi.org/10.1016/j.cedpsych.2015.06.001>
- Gueorguieva, R., & Krystal, J. H. (2004). Move over anova: Progress in analyzing repeated-measures data and its reflection in papers published in the archives of general psychiatry. *Archives of general psychiatry*, 61(3), 310–317.
- Guthrie, J. T., & Mosenthal, P. (1987). Literacy as multidimensional: Locating information and reading comprehension. *Educational Psychologist*, 22(3-4), 279-297. <https://doi.org/10.1080/00461520.1987.9653053>
- Hannafin, M. J., & Hughes, C. W. (1986). A framework for incorporating orienting activities in computer-based interactive video. *Instructional Science*, 15(1), 239-255. <https://doi.org/10.1007/BF00139613>
- Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low-and high-ability children. *Contemporary educational psychology*, 24(2), 95–123.

- Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, 21(6), 713-729. <https://doi.org/10.1002/acp.1345>
- He, W., Erdelez, S., Wang, F.-K., & Shyu, C.-R. (2008). The effects of conceptual description and search practice on users' mental models and information seeking in a case-based reasoning retrieval system. *Information Processing & Management*, 44(1), 294-309. <https://doi.org/10.1016/j.ipm.2007.03.008>
- Hegarty, M. (2014). Multimedia learning and the development of mental models. Dans R. E. Mayer (Éd.), *The Cambridge handbook of multimedia learning* (2^e éd., pp. 673–701). (S.I.) : Cambridge University Press.
- Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, 12, 45-58. <https://doi.org/10.1016/j.edurev.2014.05.001>
- Hinostroza, J. E., Ibieta, A., Labbé, C., & Soto, M. T. (2018). Browsing the internet to solve information problems: A study of students' search actions and behaviours using a 'think aloud' protocol. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-018-9698-2>
- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17(6), 722-738. <https://doi.org/10.1016/j.learninstruc.2007.09.013>
- Hyönä, J. (1995). An eye movement analysis of topic-shift effect during repeated reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(5), 1365.

- Hyönä, J., & Lorch, R. F. (2004). Effects of topic headings on text processing: evidence from adult readers' eye fixation patterns. *Learning and Instruction, 14*(2), 131-152. <https://doi.org/10.1016/j.learninstruc.2004.01.001>
- Hyönä, J., Lorch, R. F., Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*(1), 44-55. <https://doi.org/10.1037//0022-0663.94.1.44>
- Ilioudi, C., Giannakos, M. N., & Chorianopoulos, K. (2013). Investigating differences among the commonly used video lecture styles. Dans *WAVE 2013 The Workshop on Analytics on Video-based Learning (Vol. i, pp. 21–26)*.
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of documentation, 52*(1), 3–50.
- Jamet, E. (2008). *La compréhension des documents multimédias: de la cognition à la conception*. Marseille : Solal.
- Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction, 18*(2), 135-145. <https://doi.org/10.1016/j.learninstruc.2007.01.011>
- Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology, 32*(4), 588-598. <https://doi.org/10.1016/j.cedpsych.2006.07.001>
- Janosky, B., Smith, P. J., & Hildreth, C. (1986). Online library catalog systems: an analysis of user errors. *International Journal of Man-Machine Studies, 25*(5), 573–592.

- Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, 18(2), 178.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. (S.I.) : Harvard University Press.
- Joseph, A. M., & Nath, B. A. (2013). Integration of Massive Open Online Education (MOOC) system with in-classroom interaction and assessment and accreditation: An extensive report from a pilot study. Dans *Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE)* (p. 105). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
- Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors*, 46(3), 567–581.
- Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, 17(4), 471-479. <https://doi.org/10.1080/09658210802647009>
- Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, 28(3), 820-831. <https://doi.org/10.1016/j.chb.2012.01.011>

- Khacharem, A., Zoudji, B., Kalyuga, S., & Ripoll, H. (2013). The expertise reversal effect for sequential presentation in dynamic soccer visualizations. *Journal of Sport and Exercise Psychology, 35*(3), 260–269.
- Kim, M. C., & Hannafin, M. J. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. *Computers & Education, 56*(2), 403–417.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. (S.I.) : Cambridge university press.
- Kintsch, W., & Yarbrough, J. C. (1982). Role of rhetorical structure in text comprehension. *Journal of educational psychology, 74*(6), 828.
- Kitajima, M., Blackmon, M. H., & Polson, P. G. (2000). A comprehension-based model of web navigation and its application to web usability analysis. Dans S. McDonald, Y. Waern, & G. Cockton (Éds), *People and Computers XIV — Usability or Else!* (pp. 357-373). London : Springer London.
- Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*(2), 242-262.
<https://doi.org/10.1016/j.cedpsych.2004.10.001>
- Koutropoulos, A., Gallagher, M. S., Abajian, S. C., de Waard, I., Hogue, R. J., Keskin, N. Ö., & Rodriguez, C. O. (2012). Emotive vocabulary in MOOCs: Context & participant retention. *European Journal of Open, Distance and E-Learning, 15*(1).
- Kraft, P., Rise, J., Sutton, S., & Røysamb, E. (2005). Perceived difficulty in the theory of planned behaviour: Perceived behavioural control or affective attitude? *British*

Journal of Social Psychology, 44(3), 479-496.

<https://doi.org/10.1348/014466604X17533>

Krug, D., George, B., Hannon, S. A., & Glover, J. A. (1989). The effect of outlines and headings on readers' recall of text. *Contemporary Educational Psychology*, 14(2), 111–123.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science*, 42(5), 361.

Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science*, 25(2), 117–131.

Lazonder, A. W., Biemans, H. J., & Wopereis, I. G. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, 51(6), 576–581.

Lazonder, A. W., & Rouet, J.-F. (2008). Information problem solving instruction: Some cognitive and metacognitive issues. *Computers in Human Behavior*, 24(3), 753–765. <https://doi.org/10.1016/j.chb.2007.01.025>

Lemarié, J., Lorch, R. F., Eyrolle, H., & Virbel, J. (2008). SARA: A text-based and reader-based theory of signaling. *Educational Psychologist*, 43(1), 27-48. <https://doi.org/10.1080/00461520701756321>

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *ECTJ*, 30(4), 195–232.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational psychology*, 75(3), 402.

- Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational psychology review*, 1(3), 209–234.
- Lorch, R. F., Lemarié, J., & Grant, R. (2011). Signaling hierarchical and sequential organization in expository text. *Scientific Studies of Reading*, 15(3), 267-284.
<https://doi.org/10.1080/10888431003747535>
- Lorch, R. F., & Lorch, E. P. (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology*, 88(1), 38-48.
<https://doi.org/10.1037/0022-0663.88.1.38>
- Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge handbook of multimedia learning* (2^e éd., pp. 227–246). (S.I.) : Cambridge University Press.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European journal of psychology of education*, 14(2), 225–244.
- Lowe, R. K., & Schnotz, W. (2014). Animation principles in multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge Handbook of Multimedia Learning* (2^e éd., pp. 513–546). (S.I.) : Cambridge University Press.
- Marchionini, G. (1989). Making the transition from print to electronic encyclopaedias: adaptation of mental models. *International journal of man-machine studies*, 30(6), 591–618.
- Marchionini, G. (1997). *Information seeking in electronic environments*. (S.I.) : Cambridge university press.
- Marchionini, G., & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *Computer*, 21(1), 70–80.

- Martins, D. (1995). Influence des connaissances et de l'intérêt sur la compréhension de textes. *L'année psychologique*, 95(2), 201-217. <https://doi.org/10.3406/psy.1995.28820>
- Mayer, R. E. (1989). Models for understanding. *Review of educational research*, 59(1), 43–64.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American psychologist*, 59(1), 14.
- Mayer, R. E. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. Dans R. E. Mayer (Éd.), *The Cambridge handbook of multimedia learning* (pp. 183–200). (S.I.) : Cambridge University Press.
- Mayer, R. E. (2014a). Cognitive theory of multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge Handbook of Multimedia Learning* (2^e éd., pp. 43–71). (S.I.) : Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Mayer, R. E. (2014b). Incorporating motivation into multimedia learning. *Learning and Instruction*, 29, 171-173. <https://doi.org/10.1016/j.learninstruc.2013.04.003>
- Mayer, R. E. (2014c). Introduction to multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge Handbook of Multimedia Learning* (2^e éd., pp. 1-24). (S.I.) : Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of educational psychology*, 93(2), 390.

- Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95(4), 806-812. <https://doi.org/10.1037/0022-0663.95.4.806>
- Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraining and signaling. *Journal of Educational Psychology*, 76(6), 1089.
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. *The Cambridge Handbook of Multimedia Learning*, 279-315.
- Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, 11(4), 256.
- Mayer, R. E., Mathias, A., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pre-training: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied*, 8(3), 147-154. <https://doi.org/10.1037//1076-898X.8.3.147>
- Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and instruction*, 12(1), 107–119.
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. Dans

The Cambridge Handbook of Multimedia Learning (2^e éd., pp. 316-344). (S.I.) : Cambridge University Press.

Merkt, M., Ballmann, A., Felfeli, J., & Schwan, S. (2018). Pauses in educational videos: Testing the transience explanation against the structuring explanation. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2018.01.013>

Merkt, M., & Schwan, S. (2014). How does interactivity in videos affect task performance? *Computers in Human Behavior*, 31, 172-181. <https://doi.org/10.1016/j.chb.2013.10.018>

Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2011.03.004>

Michas, I. C., & Berry, D. C. (2000). Learning a procedural task: Effectiveness of multimedia presentations. *Applied Cognitive Psychology*, 14(6), 555–575.

Moreno, R. (2010). Cognitive load theory: more food for thought. *Instructional Science*, 38(2), 135-141. <https://doi.org/10.1007/s11251-009-9122-9>

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of educational psychology*, 91(2), 358.

Moret, P., & Hermann, P. (2009, janvier). *Expertise collective : Dyslexie, dysorthographe, dyscalculie*. Inserm/Scientifilms. Repéré à <https://www.youtube.com/watch?v=fFQ1vmkebNA&feature=youtu.be>

Mulder, Y. G., Bollen, L., de Jong, T., & Lazonder, A. W. (2016). Scaffolding learning by modelling: The effects of partially worked-out models. *Journal of Research in Science Teaching*, 53(3), 502-523. <https://doi.org/10.1002/tea.21260>

- Narayanan, N. H., & Hegarty, M. (1998). On designing comprehensible interactive hypermedia manuals. *International journal of human-computer studies*, 48(2), 267–301.
- Norman, D. A. (1983). Some observations on mental models. Dans D. Gentner & A. L. Stevens (Éds), *Mental models* (pp. 7–14). Hillsdale, N.J : Erlbaum.
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429-434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Paas, F. G., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge Handbook of Multimedia Learning* (2^e éd., pp. 27-42). Cambridge : Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.004>
- Paivio, A. (1990). *Mental representations: A dual coding approach*. (S.I.) : Oxford University Press.
- Potelle, H., & Rouet, J.-F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58(3), 327-345. [https://doi.org/10.1016/S1071-5819\(03\)00016-8](https://doi.org/10.1016/S1071-5819(03)00016-8)
- Potocki, A., Ros, C., Vibert, N., & Rouet, J.-F. (2017). Children's visual scanning of textual documents: Effects of document organization, search goals, and metatextual knowledge. *Scientific Studies of Reading*, 21(6), 480-497. <https://doi.org/10.1080/10888438.2017.1334060>

- Puustinen, M., & Rouet, J.-F. (2009). Learning with new technologies: Help seeking and information searching revisited. *Computers & Education*, 53(4), 1014-1019. <https://doi.org/10.1016/j.compedu.2008.07.002>
- Quintana, C., Krajcik, J., & Soloway, E. (2002). A case study to distill structural scaffolding guidelines for scaffolded software environments. Dans *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 81–88). ACM.
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria : R Foundation for Statistical Computing. Repéré à <https://www.R-project.org/>
- Redford, J. S., Thiede, K. W., Wiley, J., & Griffin, T. D. (2012). Concept mapping improves metacomprehension accuracy among 7th graders. *Learning and Instruction*, 22(4), 262-270. <https://doi.org/10.1016/j.learninstruc.2011.10.007>
- Reiser, B. J. (2002). Why scaffolding should sometimes make tasks more difficult for learners. Dans *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community* (pp. 255–264). International Society of the Learning Sciences.
- Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, 13(3), 273–304.
- Renkl, A., & Atkinson, R. K. (2007). Interactive learning environments: Contemporary issues and trends. An introduction to the special issue. *Educational Psychology Review*, 19(3), 235-238. <https://doi.org/10.1007/s10648-007-9052-5>

- Reutemann, J. (2016). Differences and Commonalities—A comparative report of video styles and course descriptions on edX, Coursera, Futurelearn and Iversity. M. Khalil, M. Ebner, M. Kopp, A. Lorenz y M. Kalz, *Proceedings of the European Stakeholder summit on experiences and best practices in and around MOOCs (EMOOCs 2016)*, 383–392.
- Reynolds, R. E., Standiford, S. N., & Anderson, R. C. (1979). Distribution of reading time when questions are asked about a restricted category of text information. *Journal of Educational Psychology*, 71(2), 183.
- Rieh, S. Y., Collins-Thompson, K., Hansen, P., & Lee, H.-J. (2016). Towards searching as a learning process: A review of current perspectives and future directions. *Journal of Information Science*, 42(1), 19–34.
- Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. Dans *Psychology of Learning and Motivation* (Vol. 55, pp. 1-36). (S.I.) : Elsevier. <https://doi.org/10.1016/B978-0-12-387691-1.00001-6>
- Rouet, J.-F. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with computers*, 15(3), 409–428.
- Rouet, J.-F., & Coutelet, B. (2008). The acquisition of document search strategies in grade school students. *Applied Cognitive Psychology*, 22(3), 389–406.
- Rouet, J.-F., & Tricot, A. (1998). Chercher de l'information dans un hypertexte: vers un modèle des processus cognitifs. *Les hypermédias, approches cognitives et ergonomiques*, 57–74.

- Rouet, J.-F., Vidal-Abarca, E., Erbou, A. B., & Millogo, V. (2001a). Effects of Information Search Tasks on the Comprehension of Instructional Text. *Discourse Processes*, 31(2), 163-186.
https://doi.org/10.1207/S15326950DP3102_03
- Rouet, J.-F., Vidal-Abarca, E., Erbou, A. B., & Millogo, V. (2001b). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes*, 31(2), 163–186.
- Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological bulletin*, 100(3), 349.
- Rowe, A. L., & Cooke, N. J. (1995). Measuring mental models: Choosing the right tools for the job. *Human resource development quarterly*, 6(3), 243–255.
- Sanchez, R. P., Lorch, E. P., & Lorch, R. F. (2001). Effects of Headings on Text Processing Strategies. *Contemporary Educational Psychology*, 26(3), 418-428.
<https://doi.org/10.1006/ceps.2000.1056>
- Santos-Espino, J. M., Afonso-Suárez, M. D., & Guerra-Artal, C. (2016). Speakers and Boards: A Survey of Instructional Video Styles in MOOCs. *Technical Communication*, 63(2), 101–115.
- Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, 1(1), 60.
- Scheiter, K. (2014). The learner control principle. *The Cambridge Handbook of Multimedia Learning*, 487–512.

- Scheiter, K., & Gerjets, P. (2007). Learner Control in Hypermedia Environments. *Educational Psychology Review*, 19(3), 285-307.
<https://doi.org/10.1007/s10648-007-9046-3>
- Schnotz, W. (2014). Integrated model of text and picture comprehension. Dans R. E. Mayer (Éd.), *The Cambridge handbook of multimedia learning* (2^e éd., pp. 72-103). (S.I.) : Cambridge University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and instruction*, 13(2), 141–156.
- Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, 14(3), 293-305.
<https://doi.org/10.1016/j.learninstruc.2004.06.005>
- Seel, N. M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning*, 1(1), 59–85.
- Sharit, J., Hernández, M. A., Czaja, S. J., & Pirolli, P. (2008). Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. *ACM Transactions on Computer-Human Interaction*, 15(1), 1-25.
<https://doi.org/10.1145/1352782.1352785>
- Sharma, P., & Hannafin, M. J. (2007). Scaffolding in technology-enhanced learning environments. *Interactive learning environments*, 15(1), 27–46.
- Simmons, P. E., & Lunetta, V. N. (1993). Problem-solving behaviors during a genetics computer simulation: Beyond the expert/novice dichotomy. *Journal of Research in Science Teaching*, 30(2), 153–173.

- Singh, A.-M., Marcus, N., & Ayres, P. (2012). The Transient Information Effect: Investigating the Impact of Segmentation on Spoken and Written text: The transient information effect. *Applied Cognitive Psychology, 26*(6), 848-853. <https://doi.org/10.1002/acp.2885>
- Slone, D. J. (2002). The influence of mental models and goals on search patterns during web interaction. *Journal of the American society for information science and technology, 53*(13), 1152–1169.
- Sorden, S. D. (2005). A cognitive approach to instructional design for multimedia learning. *Informing Science: International Journal of an Emerging Transdiscipline, 8*, 263–279.
- Spanjers, I. A. E., van Gog, T., Wouters, P., & Van Merriënboer, J. J. G. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education, 59*(2), 274-280. <https://doi.org/10.1016/j.compedu.2011.12.024>
- Spyridakis, J. H., & Standal, T. C. (1987). Signals in expository prose: Effects on reading comprehension. *Reading Research Quarterly, 28*5–298.
- Staggers, N., & Norcio, A. F. (1993). Mental models: concepts for human-computer interaction research. *International Journal of Man-machine studies, 38*(4), 587–605.
- Storey, M.-A., Fracchia, F. D., & Müller, H. A. (1999). Cognitive design elements to support the construction of a mental model during software exploration. *Journal of Systems and Software, 44*(3), 171–185.

- Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with computers*, 10(3), 321–351.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123-138.
<https://doi.org/10.1007/s10648-010-9128-5>
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, 10(3), 251–296.
- Trafimow, D., Sheeran, P., Conner, M., & Finlay, K. A. (2002). Evidence that perceived behavioural control is a multidimensional construct: Perceived control and perceived difficulty. *British Journal of Social Psychology*, 41(1), 101–121.
- Tsai, M.-J. (2009). Online Information Searching Strategy Inventory (OISSI): A quick version and a complete version. *Computers & Education*, 53(2), 473-483.
<https://doi.org/10.1016/j.compedu.2009.03.006>
- van der Meij, H., & van der Meij, J. (2013). Eight guidelines for the design of instructional videos for software training. *Technical communication*, 60(3), 205–228.
- van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & education*, 78, 150–159.
- van der Zee, T., Admiraal, W., Paas, F. G., Saab, N., & Giesbers, B. (2017). Effects of subtitles, complexity, and language proficiency on learning from online education videos. *Journal of Media Psychology*.
- van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. Dans R. E. Mayer (Éd.), *The Cambridge Handbook of Multimedia Learning* (2^e éd., pp.

263-278). Cambridge : Cambridge University Press.

<https://doi.org/10.1017/CBO9781139547369.014>

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20(2), 95-99.

<https://doi.org/10.1016/j.learninstruc.2009.02.009>

Vörös, Z., Rouet, J.-F., & Pléh, C. (2011). Effect of high-level content organizers on hypertext learning. *Computers in Human Behavior*, 27(5), 2047-2055.

<https://doi.org/10.1016/j.chb.2011.04.005>

Waern, Y. (1981). On the relation between comprehension and memory of a complex text. *Scandinavian Journal of Educational Research*, 25(1), 21-37.

<https://doi.org/10.1080/0031383810250103>

Weiler, A. (2005). Information-seeking behavior in generation Y students: Motivation, critical thinking, and learning theory. *The Journal of Academic Librarianship*, 31(1), 46–53.

Wilson, T. D. (2000). Human information behavior. *Informing science*, 3(2), 49–56.

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, 22(6), 449

-457. <https://doi.org/10.1016/j.learninstruc.2012.05.004>

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17(2), 89–100.

Woodall, W. G., Davis, D. K., & Sahin, H. (1983). From the boob tube to the black box: Television news comprehension from an information processing perspective.

Journal of Broadcasting, 27(1), 1-23.

<https://doi.org/10.1080/08838158309386469>

Wopereis, I., Brand-Gruwel, S., & Vermetten, Y. (2008). The effect of embedded instruction on solving information problems. *Computers in Human Behavior*, 24(3), 738-752. <https://doi.org/10.1016/j.chb.2007.01.024>

Wouters, P., Tabbers, H. K., & Paas, F. G. (2007). Interactivity in video-based models. *Educational Psychology Review*, 19(3), 327-342. <https://doi.org/10.1007/s10648-007-9045-4>

Yadav, K., Shrivastava, K., Mohana Prasad, S., Arsikere, H., Patil, S., Kumar, R., & Deshmukh, O. (2015). Content-driven multi-modal techniques for non-linear video navigation (pp. 333-344). ACM Press. <https://doi.org/10.1145/2678025.2701408>

Yujian, L., & Bo, L. (2007). A normalized Levenshtein distance metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 1091-1095. <https://doi.org/10.1109/TPAMI.2007.1078>

Yussen, S. R., Stright, A. D., & Payne, B. (1993). Where is it? Searching for information in a college textbook. *Contemporary Educational Psychology*, 18(2), 240-257.

Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology: Applied*, 9(2), 88-100. <https://doi.org/10.1037/1076-898X.9.2.88>

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness.

Information & Management, 43(1), 15-27.

<https://doi.org/10.1016/j.im.2005.01.004>

Zhang, X., & Chignell, M. (2001). Assessment of the effects of user characteristics on mental models of information retrieval systems. *Journal of the Association for Information Science and Technology*, 52(6), 445–459.

Annexes

Annexe 1. Questionnaire de RI (études 1, 2 et 3)

1. Quelles sont les deux formes de l'eau les plus courantes dans l'univers ? (étude 1) / Quand l'hydrogène apparaît-il ? (études 2 et 3)
2. L'eau marque la différence de formation entre deux types de planètes : lesquelles ?
3. Quel est le pourcentage d'eau dans la masse de notre corps ?
4. Sur quel satellite de Jupiter existerait-il un océan d'eau liquide ?
5. À quelle distance du Soleil se situe la ligne des glaces ?
6. Quelle est la forme du spectre de l'eau ? (étude 1) / Quel état de l'eau est recherché sur une planète pour savoir si elle est habitable ? (études 2 et 3)
7. Quelle est la configuration d'une molécule d'eau ?
8. Pour quelle raison Vénus est-elle plus chaude que la Terre ?
9. Grâce à quelle technique peut-on « voir » l'eau, la détecter ?

Annexe 2. Questionnaire de localisation (étude 2)

1. Quand l'hydrogène apparaît-il ?
2. L'eau marque la différence de formation entre deux types de planètes : lesquelles ?
3. Quel est le pourcentage d'eau dans la masse de notre corps ?
4. Sur quel satellite de Jupiter existerait-il un océan d'eau liquide ?
5. L'eau est-elle un bon solvant ?
6. Quelle est la forme du spectre de l'eau ?
7. L'atmosphère sur Mars était-elle plus ou moins dense dans le passé ?
8. Qu'est-ce que la ligne des glaces ?

Annexe 3. Questionnaire d'apprentissage (études 4 et 5)

Questions de mémorisation :

1. La dyslexie disparaît avec l'âge / On reste dyslexique toute sa vie / Je ne sais pas ?

2. Quel professionnel évalue les fonctions complexes exécutives, attentionnelles... ?
3. Quelles sont les 4 théories sur l'origine de la dyslexie développées dans cette vidéo ?
4. À partir de quand peut-on avoir des inquiétudes sur une potentielle dyslexie ?
5. Selon les chercheurs, combien d'enfants en âge scolaire en France souffriraient de dyslexie de développement ?
6. Qu'est-ce qu'un centre de référence ?
7. Qu'est-ce que la migration neuronale ?
8. Quel est le facteur clé de l'automatisation de la lecture ?
9. Quel est le professionnel qui est la base essentielle de la rééducation des troubles écrits ?
10. Selon la théorie phonologique, quelle est l'origine de la dyslexie ?
11. Quelle observation soutient l'idée d'une transmission héréditaire de la dyslexie ?
12. Comment appelle-t-on le protocole mis en place dans l'école quand des difficultés sévères sont repérées par un professionnel ?
13. Les enfants dyslexiques forment un groupe homogène : Vrai / Faux / Je ne sais pas ?
14. Selon la théorie visuelle, à quoi est due la difficulté à identifier les mots ?
15. La rééducation orthographique va réparer les anomalies de la migration neuronale : Vrai / Faux / Je ne sais pas ?

Questions de compréhension :

1. Que peut-on dire si un enfant n'améliore pas ses capacités de lecture après une prise en charge adaptée à la dyslexie ?
2. En quoi l'adaptation pour les examens proposée dans la vidéo peut aider un dyslexique ?
3. Certaines langues sont dites "transparentes" et sont moins handicapantes que d'autres pour les dyslexiques : c'est le cas de l'espagnol par rapport au français. Selon vous, pourquoi ?
4. La dyslexie touche 3 fois plus de garçons que de filles. Quelle théorie cela soutient-il ?

5. Certains dyslexiques traitent chaque mot comme un nouveau mot en rassemblant les lettres une à une. D'autres ne lisent les mots qu'en les adressant directement au lexique et ont beaucoup de mal à lire de nouveaux mots en les rassemblant. A quelles théories chacun de ces comportements correspond-il ?

Annexe 4. Questionnaire de RI (étude 5)

1. Quelles sont les conditions au niveau de l'enfant pour pouvoir parler de dyslexie de développement selon les chercheurs ?
2. Citez un professionnel qui accompagne les difficultés de graphisme et d'écriture au centre de référence :
3. On parle de contribution génétique et environnementale. Donnez des exemples de contribution environnementale :
4. Dans le cadre de la théorie phonologique, quelles sont les trois aires cérébrales impliquées dans la lecture ?

Annexe 5. Items de difficulté perçue (études 1, 2, 3, 4 et 5)

1. Pour moi, effectuer cette recherche d'information était difficile (études 1, 2 et 3) / Pour moi, apprendre cette vidéo était difficile (études 4 et 5)
2. Effectuer cette recherche d'information a été dur (études 1, 2 et 3) / Apprendre cette vidéo a été dur (études 4 et 5)
3. Il a été facile de réaliser cette recherche d'information – item inversé (études 1, 2 et 3) / Il a été facile d'apprendre cette vidéo – item inversé (études 4 et 5)

Annexe 6. Items de contrôle perçu (études 1 et 4)

1. J'ai eu un contrôle total sur la recherche d'information à effectuer (étude 1) / J'ai eu un contrôle total sur mon apprentissage de cette vidéo (étude 4)
2. J'ai trouvé que je contrôlais le comportement à effectuer (étude 1) / J'ai trouvé que je contrôlais mon apprentissage de la vidéo (étude 4)
3. La recherche d'information s'est réalisée avec un manque de contrôle –item inversé (étude 1) / L'apprentissage de cette vidéo s'est réalisé avec un manque de contrôle – item inversé (étude 4)

Annexe 7. Consigne de l'étude 1

Nous allons vous demander de rechercher des informations au sein d'une vidéo. Il s'agira de questions dont la réponse est dite explicitement au cours de la vidéo.

Pour chaque proposition, vous devrez indiquer **la réponse** à la question, ainsi que **le moment** dans la vidéo où vous l'avez repérée. Indiquez approximativement la minute et les secondes.

Par exemple, si la question était :

Questions	Réponse	Moment dans la vidéo
Quel est le premier élément qui se forme avec le Big Bang ?		

Il faudrait répondre :

Questions	Réponse	Moment dans la vidéo
Quel est le premier élément qui se forme avec le Big Bang ?	L'hydrogène	1.32 (correspondant à 1 minute et 32 secondes)

Même si vous connaissez déjà la réponse, nous vous demandons de la chercher **dans la vidéo**, et de noter l'instant où vous l'avez trouvée.

Les questions apparaîtront au fur et à mesure, sans possibilité de revenir en arrière. Pour chaque question, vous avez **5 minutes** pour trouver la réponse avant de passer automatiquement à la question suivante. Vous devrez écrire la réponse sur la feuille fournie.

Si vous trouvez la réponse avant la fin des 5 minutes, vous pouvez passer à la question suivante en cliquant sur le bouton en bas à droite :



Annexe 8. Consigne de l'étude 2 – étape 1 (tâche de RI)

Nous allons vous demander de rechercher des informations au sein d'une vidéo. Il s'agira de questions dont la réponse est dite explicitement au cours de la vidéo.

Pour chaque proposition, vous devrez indiquer **la réponse** à la question, ainsi que **le moment** dans la vidéo où vous l'avez repérée. Indiquez approximativement la minute et les secondes.

Par exemple, si la question était :

Questions	Réponse	Moment dans la vidéo
L'eau se présente sous forme de glace sur quelles planètes / satellites ?		

Il faudrait répondre :

Questions	Réponse	Moment dans la vidéo
L'eau se présente sous forme de glace sur quelles planètes / satellites ?	La Lune et Mercure	4.54 (correspondant à 4 minutes et 54 secondes)

Même si vous connaissez déjà la réponse, nous vous demandons de la chercher **dans la vidéo**, et de noter l'instant où vous l'avez trouvée.

Les questions apparaîtront au fur et à mesure, sans possibilité de revenir en arrière. Pour chaque question, vous avez **5 minutes** pour trouver la réponse avant de passer automatiquement à la question suivante.

Si vous trouvez la réponse avant la fin des 5 minutes, vous pouvez passer à la question suivante.

Annexe 9. Consigne de l'étude 2 – étape 2 (tâche de localisation)

Vous allez désormais répondre à des questions sans la vidéo.

Il s'agira de :

Donner la réponse à la question **selon vous** ;

Indiquer sur la barre de progression à **quel endroit** se trouverait l'information recherchée dans la vidéo.

En cliquant sur la barre de progression, vous placerez un intervalle de 20 secondes dans lequel est supposée se trouver la réponse selon vous.

Par exemple :

Les questions apparaîtront au fur et à mesure, sans possibilité de revenir en arrière. Pour chaque question, vous avez **5 minutes** pour trouver la réponse avant de passer automatiquement à la question suivante.

Si vous trouvez la réponse avant la fin des 5 minutes, vous pouvez passer à la question suivante.

Il s'agira pour moitié de questions déjà posées, et pour moitié de nouvelles.

Annexe 10. Consigne de l'étude 3

Nous allons vous demander de rechercher des informations au sein d'une vidéo. Il s'agira de questions dont la réponse est dite explicitement au cours de la vidéo.

Pour chaque proposition, vous devrez indiquer **la réponse** à la question, ainsi que **le moment** dans la vidéo où vous l'avez repérée. Indiquez approximativement la minute et les secondes.

Par exemple, si la question était :

Questions	Réponse	Moment dans la vidéo
L'eau se présente sous forme de glace sur quelles planètes / satellites ?		

Il faudrait répondre :

Questions	Réponse	Moment dans la vidéo
L'eau se présente sous forme de glace sur quelles planètes / satellites ?	La Lune et Mercure	4.54 (correspondant à 4 minutes et 54 secondes)

Même si vous connaissez déjà la réponse, nous vous demandons de la chercher **dans la vidéo**, et de noter l'instant où vous l'avez trouvée.

Les questions apparaîtront au fur et à mesure, sans possibilité de revenir en arrière. Pour chaque question, vous avez **5 minutes** pour trouver la réponse avant de passer automatiquement à la question suivante.

Si vous trouvez la réponse avant la fin des 5 minutes, vous pouvez passer à la question suivante.

Annexe 11. Consigne de l'étude 4

Nous allons vous demander d'apprendre le maximum d'informations dans une vidéo. Pour cela, vous pouvez naviguer **comme vous voulez** au sein de la vidéo, grâce à la barre de navigation.

Lorsque vous pensez avoir fini votre apprentissage, vous pouvez passer au **questionnaire de connaissances** en cliquant sur le bouton « Passer au questionnaire ».

Vous avez maximum **30 minutes** pour apprendre le maximum d'informations. Si vous n'êtes pas encore passé(e) au questionnaire avant ces 30 minutes, nous vous demanderons de bien vouloir le faire.

Annexe 12. Consigne de l'étude 5 – participants sans RI

Nous allons vous demander d'apprendre le maximum d'informations dans une vidéo. Pour cela, vous pouvez naviguer **comme vous voulez** au sein de la vidéo, grâce à la barre de navigation.

Lorsque vous pensez avoir fini votre apprentissage, vous pouvez passer au **questionnaire de connaissances** en cliquant sur le bouton « Passer au questionnaire d'apprentissage ».

Vous avez maximum **30 minutes** pour apprendre le maximum d'informations. Si vous n'êtes pas encore passé(e) au questionnaire avant ces 30 minutes, la vidéo s'arrêtera automatiquement.

La prise de notes n'est pas possible durant l'apprentissage.

Annexe 13. Consigne de l'étude 5 – participants avec RI

Nous allons vous demander de rechercher des informations au sein d'une vidéo. Il s'agira de questions dont la réponse est dite explicitement au cours de la vidéo.

Pour chaque proposition, vous devrez indiquer **la réponse** à la question, ainsi que **le moment** dans la vidéo où vous l'avez repérée. Indiquez approximativement la minute et les secondes.

Par exemple, si la question était :

Questions	Réponse	Moment dans la vidéo
Quelle est la bonne méthode d'apprentissage de la lecture dans le cerveau ?		

Il faudrait répondre :

Questions	Réponse	Moment dans la vidéo
Quelle est la bonne méthode d'apprentissage de la lecture dans le cerveau ?	Apprentissage des correspondances entre les graphèmes et les phonèmes	13.12 (correspondant à 13 minutes et 12 secondes)

Même si vous connaissez déjà la réponse, nous vous demandons de la chercher **dans la vidéo**, et de noter l'instant où vous l'avez trouvée.

Les questions apparaîtront au fur et à mesure, sans possibilité de revenir en arrière. Pour chaque question, vous avez **5 minutes** pour trouver la réponse avant de passer automatiquement à la question suivante. Si vous trouvez la réponse avant la fin des 5 minutes, vous pouvez passer à la question suivante.

A la fin de la dernière question, vous pourrez cliquer sur le bouton « Passer à la phase d'apprentissage ».

Nous vous demanderons alors d'apprendre le maximum d'informations dans la même vidéo. Pour cela, vous pourrez naviguer **comme vous voudrez** au sein de la vidéo, grâce à la barre de navigation.

Lorsque vous pensez avoir fini votre apprentissage, vous pouvez passer au **questionnaire de connaissances** en cliquant sur le bouton « Passer au questionnaire d'apprentissage ».

Vous avez maximum **30 minutes** pour apprendre le maximum d'informations. Si vous n'êtes pas encore passé(e) au questionnaire avant ces 30 minutes, la vidéo s'arrêtera automatiquement.

La prise de notes n'est pas possible durant l'apprentissage.

Index des figures

Figure 1. Exemples de style talking-head (1) sur la plateforme EdX, de style Khan (2) sur la plateforme Khan Academy et d'image du professeur intégrée à l'image présentée (3) sur la plateforme Coursera	22
Figure 2. Modèle de la théorie cognitive de l'apprentissage multimédia de Mayer (2014a)	27
Figure 3. Schématisation du modèle de Sharit et al. (2008)	43
Figure 4. Segmentation et structuration au sein d'un environnement vidéo (étude 1)	76
Figure 5. Environnements vidéo dans les conditions sans étayage (1), micro-étayage (2), macro-étayage (3) et double niveau d'étayage (4) de l'étude 1	80
Figure 6. Procédure de l'étude 1	85
Figure 7. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 1	89
Figure 8. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 1	91
Figure 9. Taux d'erreur moyens des participants en RI sur le premier clic selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 1	92
Figure 10. Environnements vidéo pour la tâche de RI dans les conditions avec (1) ou sans (2) étayage de l'étude 2	104
Figure 11. Environnement vidéo pour l'étape de localisation de l'étude 2....	105
Figure 12. Procédure de l'étude 2	108
Figure 13. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 2	110
Figure 14. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 2	112
Figure 15. Taux d'erreur moyens des participants en RI sur le premier clic selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 2	113

Figure 16. Durée totale d'exposition (en secondes) en fonction du score de localisation (marge d'erreur) pour l'étude 2	115
Figure 17. Possibilités d'interaction et d'intégration des deux niveaux d'étayage dans l'environnement vidéo (étude 3).....	128
Figure 18. Environnements vidéo dans les conditions expérimentales sans étayage (1), avec étayage non-interactif (2), avec étayage interactif (3), avec pop-ups (4) et avec pop-ups et présence visuelle (5) de l'étude 3	130
Figure 19. Procédure de l'étude 3	133
Figure 20. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 3	136
Figure 21. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 3	138
Figure 22. Taux d'erreur moyens des participants sur le premier clic en RI selon la condition expérimentale et le rang de la question (de 1 à 9) pour l'étude 3	139
Figure 23. Segmentation et structuration au sein d'un environnement vidéo (étude 4)	151
Figure 24. Environnements vidéo dans les conditions sans étayage (1), micro-étayage (2), macro-étayage (3) et double niveau d'étayage (4) de l'étude 4	152
Figure 25. Procédure de l'étude 4	156
Figure 26. Environnements vidéo dans les conditions contrôle (1), avec étayage (2), avec RI (3) ou avec étayage et RI (4) de l'étude 5	172
Figure 27. Procédure de l'étude 5	176
Figure 28. Scores de réussite moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 4) pour l'étude 5	178
Figure 29. Temps de réponse moyens des participants en RI selon la condition expérimentale et le rang de la question (de 1 à 4) pour l'étude 5	179
Figure 30. Taux d'erreur moyens des participants en RI sur le premier clic selon la condition expérimentale et le rang de la question (de 1 à 4) pour l'étude 5	180

Index des tableaux

Tableau 1. Caractéristiques du modèle mental et implications dans les performances en apprentissage	30
Tableau 2. Caractéristiques du modèle mental et implications dans les performances en RI	46
Tableau 3. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 1).....	86
Tableau 4. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 1)	90
Tableau 5. Données descriptives pour les variables de difficulté perçue, contrôle perçu et nombre de chapitres rappelés en fonction de la condition expérimentale (étude 1)	93
Tableau 6. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 2).....	109
Tableau 7. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 2)	111
Tableau 8. Données descriptives pour la variable de difficulté perçue en fonction de la condition expérimentale (étude 2).....	113
Tableau 9. Données descriptives pour le score (marge d'erreur) de localisation en fonction de la condition expérimentale (étude 2)	114
Tableau 10. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 3).....	135
Tableau 11. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 3)	137
Tableau 12. Données descriptives pour la variable de difficulté perçue en fonction de la condition expérimentale (étude 3)	140
Tableau 13. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 4).....	157

Tableau 14. Données descriptives pour les variables d'apprentissage (temps d'apprentissage et scores de mémorisation et de compréhension) en fonction de la condition expérimentale (étude 4)	158
Tableau 15. Nombre de révisions effectuées, temps total et temps moyen par révision en fonction de la condition expérimentale (étude 4).....	159
Tableau 16. Données descriptives pour les variables de difficulté perçue, contrôle perçu et distance de Levenshtein sur les chapitres rappelés en fonction de la condition expérimentale (étude 4)	161
Tableau 17. Données descriptives pour les variables d'intérêt préalable et de compétence perçue en fonction de la condition expérimentale (étude 5).....	177
Tableau 18. Données descriptives pour les échecs en RI en fonction de la condition expérimentale (étude 5)	179
Tableau 19. Données descriptives pour les variables d'apprentissage (temps d'apprentissage et scores de mémorisation et de compréhension) en fonction de la condition expérimentale (étude 5)	182
Tableau 20. Nombre de révisions effectuées, temps total et temps moyen par révision en fonction de la condition expérimentale (étude 5).....	183
Tableau 21. Données descriptives pour la variable difficulté perçue en fonction de la condition expérimentale (étude 5).....	184

ÉTAYAGE DES ACTIVITES DE RECHERCHE D'INFORMATION ET D'APPRENTISSAGE EN ENVIRONNEMENT VIDEO : APPORTS DE LA SEGMENTATION ET DE LA STRUCTURATION

Résumé

Les vidéos sont de plus en plus utilisées dans un contexte pédagogique, mais les formats de présentation sur les plateformes spécialisées (e.g., MOOC) sont variés et peu analysés. Il paraît cependant nécessaire de s'intéresser aux processus à l'œuvre lors des tâches d'apprentissage et de recherche d'information (RI) pour adapter au mieux les environnements vidéo aux besoins et ressources cognitifs des individus. Au cours de ces deux tâches (RI et apprentissage), la littérature scientifique fait apparaître l'importance de la qualité du modèle mental de l'individu. Pour favoriser la construction d'un modèle mental pertinent par l'individu, deux types d'étayage sont envisagés au cours de cette thèse : la segmentation de la barre de navigation et la structuration du contenu de la vidéo par une table des matières. Les trois premières études ont permis de montrer que la présence conjointe des deux niveaux d'étayage a favorisé la performance en RI (étude 1), mais qu'elle n'a pas permis la construction d'un modèle mental efficace du contenu de la vidéo (étude 2). Au cours de l'étude 3, différents formats de présentation ont été comparés : le format *pop-up* a permis de diminuer la difficulté perçue de la tâche de RI. Les deux dernières études se sont centrées sur les effets de l'étayage lors d'une tâche d'apprentissage. La présence d'étayage n'a pas amélioré la qualité de l'apprentissage (étude 4). Néanmoins, lorsque l'apprenant est rendu actif, grâce à une tâche de RI par exemple, la mémorisation du contenu de la vidéo a été améliorée (étude 5). Les résultats sont discutés en termes de rôle organisateur de l'étayage permettant une récupération efficace de l'information. L'étayage organiserait l'information entrante dans l'environnement vidéo ou la mémoire de l'individu selon la tâche.

Mots clés

Apprentissage, Recherche d'information, Vidéo, Étayage, Segmentation, Structuration

SCAFFOLDING INFORMATION SEEKING AND LEARNING IN VIDEO-BASED ENVIRONNEMENTS: CONTRIBUTIONS OF SEGMENTATION AND STRUCTURATION

Abstract

Videos are more and more used in pedagogical contexts, but presentation formats on specific platforms (e.g., MOOC) are varied and not so much analyzed. However, it seems necessary to focus on processes involved during learning and information seeking (IS) tasks to adapt video-based environments to individuals' needs and cognitive resources. During learning and IS, scientific literature highlights the importance of individuals' mental models. To encourage the construction of a relevant mental model, two types of scaffolding are considered in this thesis: the segmentation of the timeline and the structuration of the video content using a table of contents. The first three studies showed that the conjoint use of the two levels of scaffolding enhanced IS performance (study 1), but it did not promote the construction of a relevant mental model of the video content (study 2). During study 3, several presentation format were compared: the pop-up format reduced perceived difficulty of the IS task. The last two studies focused on the effects of scaffolding during a learning task. The presence of scaffolding did not enhanced learning performance (study 4). Nevertheless, when the learner was made active, thanks to an IS task for example, memorization of the video content was improved (study 5). Results are discussed in terms of organizing role of scaffolding enabling effective recovery of information. Scaffolding is here thought to organize incoming information in the video-based environment or in individual's memory according to the task.

Keywords

Learning, Information seeking, Video, Scaffolding, Segmentation, Structuration

