

Que nous apprennent les réponses des élèves à deux tâches de recherche d'information en ligne ?

What can we learn from students' responses to two online information-seeking tasks?

Audrey Kumps – audrey.kumps@umons.ac.be – <https://orcid.org/0000-0001-6481-0318>

Gaëtan Temperman – Gaetan.temperman@umons.ac.be – <https://orcid.org/0000-0002-0200-350X>

Charles Glineur – Charles.glineur@umons.ac.be – <https://orcid.org/0000-0002-9945-0348>

Bruno De Lièvre – bruno.delievre@umons.ac.be – <https://orcid.org/0000-0001-8843-1582>

Service d'ingénierie Pédagogique et Numérique éducatif, Université de Mons - Belgique

Pour citer cet article : Krumps, A., Temperman, G., Glineur, C. et De Lièvre, B. (2024). Que nous apprennent les réponses des élèves à deux tâches de recherche d'information en ligne ? *Évaluer. Journal international de recherche en éducation et formation*, 10(3), 49-75. <https://doi.org/10.48782/e-jiref-10-3-49>

Résumé

L'objectif de cette recherche a une portée diagnostique quant à l'enseignement de la recherche d'information en ligne. Elle présente les principaux résultats d'une analyse qualitative d'un échantillon de 1138 erreurs écrites, formulées par 260 élèves de 8 à 13 ans, dans le cadre de deux tâches de recherche d'information en ligne. Dans un second temps, une comparaison des erreurs en fonction du type de tâche réalisée ainsi que de l'âge des sujets est proposée. Cette analyse de contenu est réalisée sur la base de la typologie des erreurs d'Astolfi ainsi que sur les recherches empiriques mettant en avant les difficultés des élèves lors de cette activité. Nos résultats indiquent que les erreurs des élèves portent de manière plus prédominante sur des difficultés propres au contenu. En termes de perspectives, cette analyse des erreurs permet de formuler des pistes d'aides didactiques selon les différents temps de l'apprentissage.

Mots-clés

Recherche en ligne, traitement de l'information, analyse de contenu, erreurs, lecture en ligne

Abstract

The objective of this research has a diagnostic scope regarding the teaching of online information retrieval. It presents the main results of a qualitative analysis of a sample of 1,138 written errors made by 260 students aged 8 to 13 during two online information retrieval tasks. Subsequently, a comparison of the errors based on the type of task performed and the age of the subjects is provided. This content analysis is conducted based on Astolfi's typology of errors as well as empirical research highlighting students' difficulties during this activity. Our results indicate that students' errors predominantly relate to content-specific challenges. In terms of prospects, this error analysis allows us to suggest didactic support strategies tailored to the different stages of learning.

Keywords

Online research, information processing, content analysis, errors; online reading

1 Introduction

En Belgique francophone, une des nouvelles missions de l'enseignant sera de développer les compétences informationnelles des apprenants, notamment en ce qui concerne la recherche d'informations sur le web. Cette orientation vers le renforcement de ces compétences s'explique par leur pertinence sociétale avérée (Bricteux *et al.*, 2021 ; Collard *et al.*, 2021). De plus, elle découle du constat que de nombreux élèves font face à des difficultés dans ce domaine (Walraven *et al.*, 2009 ; Mullis *et al.*, 2017 ; Bricteux *et al.*, 2021) et que l'enseignement de ces compétences demeure encore peu répandu dans le contexte scolaire (TALIS, 2018 ; Digital Wallonia, 2018).

Mais les enseignants connaissent-ils les stratégies utilisées (adéquates ou non) de leurs élèves afin d'adapter leurs interventions pédagogiques ?

Lors de cette recherche, nous nous sommes intéressés aux erreurs que les élèves commettent lorsqu'ils effectuent une recherche en ligne. Dans l'enseignement, l'erreur peut être perçue négativement. Pourtant « tout apprentissage est source potentielle d'erreur. Il n'y a pas d'apprentissage sans erreurs, parce que cela voudrait dire que celui qui apprend sait déjà » (Cuq et Gruca, 2002, p.389). L'erreur est donc non seulement inévitable, mais elle est nécessaire. Elle constitue, en effet, un indice et un moyen d'apprentissage qui devraient être pris en compte comme des moyens pertinents vers la reconstruction d'un apprentissage stable (Algubbi, 2016).

Cette recherche opte pour une approche qui part du postulat que rechercher la source de l'erreur peut être bénéfique tant pour l'élève que pour l'enseignant. En effet, porter un regard réflexif sur ses erreurs donne la possibilité à l'élève de révéler ses stratégies et ses connaissances mobilisées pour pouvoir les ajuster. Pour l'enseignant, l'exploitation de l'erreur est un instrument de régulation pédagogique. Elle permet de prendre conscience des démarches d'apprentissage des élèves, d'identifier leurs besoins et de les traduire en objectifs didactiques afin de surmonter ces obstacles (Vanssay et Lozac'h, 2012).

Cette étude se situe dans une approche psycho-didactique des évaluations (Vantourout et Goasdoué, 2014) où l'analyse des erreurs des élèves peut procurer des informations sur le niveau de développement d'une compétence et, aussi, avoir une portée diagnostique. Il s'agit d' « approches qualitatives dont la spécificité est de prendre comme point de départ les connaissances à évaluer et de s'intéresser aux processus de réponse des évalués » (Vantourout et Blanc, 2017, p.215).

L'objectif de cette étude est de faire émerger des recommandations pédagogiques à partir des possibles erreurs des élèves et de fournir aux enseignants des indices suffisamment solides pour pouvoir interpréter les réponses de leurs élèves. Par ailleurs, les résultats de cette étude seront partagés avec les enseignants au travers d'une formation cublée afin de les sensibiliser aux défis spécifiques rencontrés par les élèves et d'améliorer les stratégies pédagogiques en matière de recherche d'information en ligne.

2 Cadre théorique

Afin de pouvoir analyser les erreurs commises par les élèves, nous nous sommes appuyés sur la typologie d'erreurs d'Astolfi comme cadre de référence ainsi que sur la littérature existante afin de spécifier chaque type d'erreur au regard des difficultés rencontrées lors d'une recherche d'information en ligne.

2.1 Typologie d'erreurs d'Astolfi

Astolfi (1997) considère l'erreur comme la trace d'une activité intellectuelle authentique permettant à l'apprenant d'apprendre ; et donnant à l'enseignant de la matière pour mieux comprendre le raisonnement de ses élèves. Sa typologie décrit huit différentes natures d'erreurs susceptibles d'être commises par les apprenants :

- Les erreurs relevant de la compréhension des consignes de travail données oralement ou par écrit. Ces erreurs peuvent mettre en avant, par exemple, des décodages inadéquats de l'implicite de la consigne ou du vocabulaire. En effet, le vocabulaire utilisé est spécifique à chaque discipline et peut avoir une signification différente. On peut également observer ce type d'erreurs lorsque les élèves distinguent difficilement les moments où on attend d'eux une réponse brève ou une réponse construite et originale.
- Les erreurs résultant d'habitudes scolaires ou d'un décodage erroné des attentes comme, par exemple, quand un élève est focalisé sur l'exigence de donner une réponse à problème posé et qu'il ne perçoit pas la possibilité d'une non-réponse.
- Les erreurs témoignant des conceptions alternatives des élèves, comme les représentations préalables à l'apprentissage.
- Les erreurs liées aux opérations intellectuelles impliquées lorsque l'élève n'a pas les prérequis cognitifs nécessaires.
- Les erreurs portant sur les démarches adoptées où le raisonnement, les stratégies et les démarches empruntées ne sont pas en adéquation avec la tâche demandée.
- Les erreurs dues à une surcharge cognitive lorsque la tâche est trop complexe.
- Les erreurs ayant leur origine dans une autre discipline. Les erreurs sont liées au fait que les apprenants ne font pas le rapprochement entre des outils déjà utilisés dans une discipline et ceux qui sont requis pour une autre discipline. Ce type d'erreur est en lien avec la notion de transfert d'une discipline à l'autre.
- Les erreurs causées par la complexité propre du contenu. Deux situations particulières sont à mettre en évidence : une simplification telle qu'il n'y a plus de difficulté donc d'apprentissage nouveau, une complexité telle que l'apprenant se décourage.

2.2 Les étapes de la recherche d'information en ligne

Les travaux s'intéressant aux comportements des utilisateurs du Web aboutissent à la conclusion que plusieurs étapes peuvent être dégagées d'une recherche d'information (Marchionini, 1995 ; Eisenberg et Berkowitz, 1988 ; Brand-Gruwel *et al.*, 2009 ; Dumouchel, 2016). Pour spécifier ces étapes, nous nous sommes appuyés sur le modèle *Information Problem Solving while using Internet* [IPS-I] de Brand-Gruwel *et al.*, 2009. Le modèle de recherche d'information en ligne de Brand-Gruwel *et al.* (2009) se distingue des autres modèles en proposant une approche centrée sur les compétences. Cette approche est bien accueillie par les pédagogues et les didacticiens (Dinet, 2015), car elle permet de prendre en compte les processus cognitifs et les habiletés de contrôle et de régulation tout au long de la recherche (Kumps, 2024). En raison de cette approche pédagogique et de son appréciation positive par les enseignants, nous avons décidé de choisir ce modèle comme cadre de notre réflexion conceptuelle et méthodologique.

Dans ce modèle, la recherche d'information est considérée comme une activité de résolution de problèmes car elle repose sur l'atteinte d'un but dans le cadre de contraintes matérielles et

souvent temporelles, tout en passant par des sous-butts intermédiaires, en mobilisant un certain nombre d'aptitudes et de processus cognitifs et métacognitifs.

Le modèle IPS-I définit 5 composantes appelées "compétences constitutives" décomposées en sous-compétences : (1) La définition de la résolution de problèmes commence par la reconnaissance d'un besoin d'information. Elle comprend les tâches consistant à lire le problème, à formuler des questions, à activer les connaissances antérieures en mémoire, à clarifier les exigences et à déterminer les informations nécessaires. (2) La recherche d'informations, où l'utilisateur doit choisir un moteur de recherche, une stratégie de recherche, spécifier les termes de la requête et évaluer les résultats. (3) Le balayage de l'information : il s'agit de balayer l'information, de l'évaluer et de retenir les informations pertinentes. (4) Le traitement de l'information : lire l'information en détail, l'évaluer et la retenir pour en développer le contenu. (5) L'organisation de l'information : formuler le problème, structurer les informations pertinentes et décrire le produit, le réaliser et développer le contenu.

La lecture est considérée comme itérative et métacognitive. C'est pourquoi le modèle IPS-I prévoit également 4 activités de régulation au cours du processus de recherche : l'orientation, le suivi, le pilotage et l'évaluation. Plus précisément, la régulation interviendrait pour : (1) gérer les réalisations de la tâche en fonction des conditions matérielles et temporelles, (2) gérer et ajuster ses comportements de recherche et d'analyse de l'information, (3) évaluer la crédibilité, la récence, la pertinence des informations retrouvées, (4) évaluer le produit issu de sa recherche après avoir traité l'information.

2.3 Difficultés des élèves lors d'une recherche en ligne

Sur le plan pédagogique, de nombreux auteurs (Kingsley *et al.*, 2015 ; Cho *et al.*, 2017 ; Perez *et al.*, 2018 ; Delhougne, 2021) s'accordent pour dire qu'il est primordial d'accompagner les élèves dans cette acquisition de compétences spécifiques que sont la recherche et le traitement de l'information en ligne. En effet, bien que motivés par l'utilisation d'Internet, les élèves ont bien du mal à repérer l'information pertinente.

Leurs difficultés peuvent se situer dans la définition du problème posé ainsi que lors de la formulation des requêtes au moteur de recherche. Ils montrent des difficultés à formuler des questions de recherche et à activer leurs connaissances antérieures (Walraven *et al.*, 2008). Il arrive également que lorsqu'ils prennent appui sur les connaissances préalables, ils ne confirment pas celles-ci en réalisant une recherche (Schillings et André, 2020). Dans la plus grande partie des cas, ils commencent immédiatement la tâche sans explorer le sujet ou planifier la recherche (Fleury, 2016). C'est pourquoi, à défaut de ne pas avoir assez réfléchi à leur besoin d'information avant de s'engager dans la lecture, les élèves perdent assez vite le fil de leur recherche (Rouet *et al.*, 2011).

L'utilisation d'un moteur de recherche peut également poser problème aux apprenants. Ceux-ci trient la quantité d'informations présentes sur Internet pour trouver les documents qui paraissent intéressants pour leurs besoins. Les ressources proposées peuvent être des pages web, des articles de forum, des images, des vidéos, des fichiers, etc. Contrairement à un texte continu, ces différentes informations ne sont pas destinées à être lues en profondeur dans l'objectif de former une représentation mentale chez le lecteur (Rouet *et al.*, 2011). Le lecteur doit donc choisir parmi ce grand nombre d'alternatives pour lesquelles seules des informations éparpillées à savoir un titre, un extrait de la page Web respective et son URL sont fournies. Et, sur la base de ces informations, le lecteur doit porter des jugements prédictifs

sur la pertinence et la fiabilité des documents disponibles. Ce travail qualitatif les amène à traiter l'information disponible de manière sélective en cherchant des catégories qui répondent à leur objectif de recherche (Rouet *et al.*, 2011). Les élèves du primaire et même du secondaire rencontrent parfois des difficultés à accomplir ces tâches de manière efficace. Dans de telles situations de décisions incertaines, les individus se fient souvent à des indices heuristiques pour expliquer leur choix. Ainsi, les apprenants distinguent de manière erronée les éléments pertinents et non pertinents contenus dans les liens (Macedo-Rouet *et al.*, 2019).

Une fois le lien sélectionné, le lecteur arrive sur une page web contenant différentes informations susceptibles de répondre à leur besoin de départ. À cette étape, les élèves omettent souvent d'examiner d'un œil critique quelles sont les sources ainsi que la pertinence de ces informations en ligne (Kumps *et al.*, 2022). Lors du traitement, les jeunes internautes lisent soit le premier paragraphe, soit toute la page (Vanderschantz *et al.*, 2014). Une partie des élèves pensent que, comme lors d'une lecture traditionnelle, tout doit être lu. (Schillings et André, 2020). Mais, ils passent peu de temps sur le contenu du lien (Fives et Dinsmore, 2017). Souvent, ils ne lisent pas les pages en profondeur et, par conséquent, ne peuvent donc pas déterminer si elles répondent (ou pas) à leur besoin. S'ils identifient la réponse à leur question dans un document, la recherche est terminée et, si ce n'est pas le cas, ils ajustent les termes de la recherche en modifiant leur requête ou en choisissant un des autres liens (Boudée et Tricot, 2010). De plus, en l'absence de confrontation de ses données personnelles et des données du texte tout au long de la lecture, certains élèves ne lisent pas le texte pour répondre aux questions posées et se fient uniquement à leurs connaissances (Renaud, 2020).

Enfin, les apprenants éprouvent des difficultés à synthétiser l'information récoltée pendant leur navigation sur le web. Une navigation efficace à travers les pages exige que les lecteurs traitent non seulement les relations sémantiques entre les pages, comme en témoigne le rôle des compétences en lecture sur la navigation, mais aussi les relations spatiales entre les pages et entre les sections d'une page particulière (Baddeley, 2012). Cette structure hypertextuelle tend à augmenter les exigences cognitives de la prise de décision et du traitement visuel, et cette charge cognitive supplémentaire nuit à la compréhension de la lecture (DeStefano et LeFevre, 2007). De plus, la multitude de documents présents sur le web demande en parallèle au lecteur de se créer une représentation des relations (complémentaires ou contradictoires) entre les informations disponibles sur les documents et de la complexité des connaissances abordées (Britt et Rouet, 2012). Aussi, la transcription de la réponse identifiée dans un document numérisé vers un format papier peut également mettre les élèves en difficulté. En effet, lors du traitement visuel d'informations présentées sous différents formats, une surcharge peut apparaître. Ainsi, l'effet d'attention partagée (Mayer et Morena, 2003) peut également nuire à la réussite de la tâche de recherche lorsque l'élève doit gérer plusieurs formats.

3 Méthodologie

3.1 Échantillon

La méthode d'échantillonnage utilisée est le recrutement par convenance. Il est composé de 96 élèves de 8-9 ans (43 filles) ; de 88 élèves de 10-11 ans (42 filles) et de 76 élèves de 12-13 ans (30 filles), soit 260 sujets répartis sur treize écoles. Ces trois groupes se composent d'élèves non redoublants, sans trouble d'apprentissage, issus d'établissements scolaires situés en Belgique francophone et reposent sur la participation volontaire des enseignants à la recherche. Aucun enseignant n'a déclaré organiser dans sa classe des activités spécifiques

concernant la recherche d'information en ligne avec ses élèves. Ceux-ci ont été observés en milieu d'année scolaire.

3.2 Questions de recherche

Cette étude se veut compréhensive des erreurs des élèves afin que les enseignants puissent les appréhender. Pour répondre à cet objectif, nous avons rédigé 3 questions de recherche.

- a. Quelles sont les erreurs commises par les élèves lorsqu'ils recherchent de l'information en ligne ?
- b. Les deux tâches de recherche proposées différencient-elles les erreurs commises ?
- c. L'âge des apprenants différencie-t-il les erreurs commises ?

3.3 Prise de données

Les réponses analysées sont issues des feuilles-réponses des élèves lors de la réalisation de deux tâches de recherche en ligne.

La tâche 1 (Tâche de localisation) est considérée comme de complexité faible puisqu'elle ne demande pas de transformation de l'information pour répondre à la question proposée (Kuiper, *et al.*, 2005 ; Walraven *et al.*, 2009 ; Kim, 2009). Les élèves retrouvent directement l'information demandée sur le lien consulté (retrouver et prélever des informations). Pour cette tâche, les élèves doivent retrouver 6 informations sur le personnage Georges Rémi : (1) son pseudonyme, (2) sa profession, (3) sa date de naissance, (4 et 5) deux œuvres connues, (6) sa nationalité (Figure 1).

Qui est Georges Rémi ?

Pseudonyme : _____

Date de naissance : _____

Oeuvres connues : _____

Nationalité : _____

Profession : _____

Figure 1 : Tâche de recherche 1

La tâche 2 (Tâche d'interprétation) est dite de complexité modérée étant donné qu'elle requiert un résumé de l'information (Kuiper, *et al.*, 2005 ; Walraven *et al.*, 2009 ; Kim, 2009). Tout ce qui est lu n'est pas forcément important pour répondre au besoin d'information. Néanmoins, la confrontation de plusieurs sources enrichit la réponse (retrouver, prélever des informations, les confronter à d'autres sources et les résumer). Pour cette tâche, les élèves doivent décrire 6 étapes pour intervenir lorsque quelqu'un saigne du nez (Figure 2).

Ton établissement scolaire décide de mettre en place un tout nouveau projet « Apprendre à porter secours ». Ta classe choisit de travailler sur le saignement de nez et de rédiger une fiche sur les bons gestes à adopter.

Recherche sur Internet les informations dont tu as besoin pour réaliser cette fiche.

Le saignement de nez

Comment intervenir ?

Figure 2 : Tâche de recherche 2

Ainsi, après avoir produit une requête sur un moteur de recherche pour répondre à la tâche de recherche posée, l'apprenant est amené à sélectionner une ou plusieurs références dans une liste de résultats et à traiter l'information pour répondre à chacune des trois questions. L'ensemble de la procédure de recherche est réalisé via une tablette. Pendant l'utilisation du dispositif, l'expérimentateur a veillé à ne pas rester à côté de l'élève afin de lui donner la possibilité d'aller vers les liens qu'il souhaitait, même s'ils n'étaient pas en rapport avec les tâches de recherche proposées.

Les tâches étaient présentées sur une feuille A4, en laissant suffisamment d'espace entre elles pour que les élèves puissent écrire leur(s) réponse(s). Cette feuille de travail a pour finalité de clairement définir chaque problème informationnel et de le rendre visible afin d'empêcher les élèves de changer d'objectif de recherche, et cela leur permet de rester concentrés sur la tâche à accomplir (De Vries *et al.*, 2005).

Le temps imparti pour réaliser l'activité était libre. Les élèves ont avancé à leur propre rythme. La non-définition d'un temps à consacrer pour la tâche vise à éviter que les élèves se sentent pressés ce qui peut occasionner, par exemple, de la négligence dans l'évaluation des sources ou des erreurs dans la sélection des informations (Dinet et Tricot, 2008 ; Walraven *et al.*, 2009).

Tous ont commencé par la tâche 1 et ont été autorisés à changer quand ils le souhaitaient. Cette consigne a permis à tous les élèves de suivre le même protocole expérimental (commencer par la tâche de complexité la plus faible).

Les apprenants étaient placés de manière à ne pas voir les écrans des autres étudiants. Nous avons choisi cet outil en raison de sa facilité de transport et de la possibilité d'enregistrer les différentes manipulations effectuées sur l'écran. L'expérimentateur a répondu à toutes les questions techniques liées à l'utilisation de la tablette, afin d'aider les élèves moins familiers avec cet outil et de garantir que chacun disposait des mêmes conditions pour l'expérience, sans que cela ne constitue un biais dans leur recherche.

Enfin, pour garantir une pratique de recherche éthique qui protège la sécurité, la vie privée et la confidentialité de nos participants, nous avons anonymisé les noms des participants à chaque activité en les remplaçant par un numéro.

3.4 Méthode d'analyse

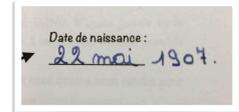
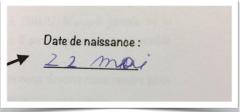
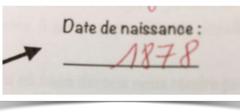
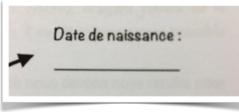
Il existe différentes méthodes d'analyse de données qualitatives (Langley, 1997), mais aucune n'est considérée plus efficace qu'une autre. Cependant, certaines méthodes sont plus appropriées que d'autres compte tenu du contexte dans lequel le chercheur travaille (Trudel et Gilbert, 1999). Dans notre cas, l'analyse de contenu semble pouvoir être pertinemment appliquée.

On peut considérer que l'analyse de contenu d'un document consiste, selon Mucchielli (2006, p.24) à « rechercher les informations qui s'y trouvent, dégager le sens ou les sens de ce qui est présenté, formuler et classer tout ce que « contient » ce document ». Bardin (1977, p.43) la définit, quant à lui, comme « un ensemble de techniques d'analyse des communications visant, par des procédures systématiques et objectives de description du contenu des énoncés, à obtenir des indicateurs (quantitatifs ou non) permettant l'inférence de connaissances relatives aux conditions de production/réception (variables inférées) de ces énoncés ». Notre objectif étant de pouvoir quantifier les erreurs commises pour établir les plus dominantes, nous optons pour la méthodologie développée par Bardin qui combine l'analyse de contenu avec d'autres méthodes comme la fréquence d'apparition.

Pour structurer notre analyse de contenu déductive, nous l'avons articulée autour de trois phases chronologiques : la préanalyse, l'exploitation du matériel ainsi que le traitement des résultats, l'inférence et l'interprétation (Bardin, 1977).

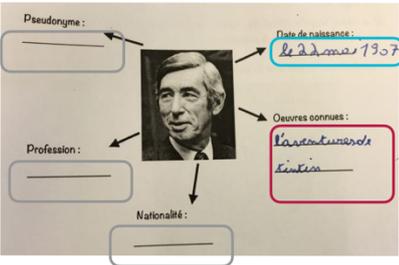
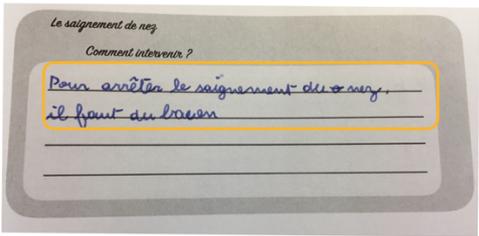
La préanalyse est une étape préliminaire intuitive d'organisation pour opérationnaliser et systématiser les idées de départ afin d'aboutir à un plan d'analyse. Nous avons choisi de soumettre à l'analyse les feuilles de réponses où les élèves ont résolu les tâches de recherche d'information en ligne. Nous avons réalisé une première lecture de ces documents en annotant librement le texte (lecture flottante). Celle-ci nous a permis de faire émerger 4 types de réponses (tableau 1) : des réponses correctes, des réponses incomplètes, des réponses erronées ou encore des non-réponses.

Tableau 1 : Lecture flottante : 4 types de réponses possibles

Réponse correcte	Réponse incomplète	Réponse erronée	Non-réponse
			

Cette lecture flottante a délimité notre champ d'investigation. Nous avons, de ce fait, choisi d'analyser les réponses erronées et les réponses incomplètes. Les réponses correctes ainsi que les non-réponses ont donc été écartées de l'analyse. En effet, les réponses correctes ne nous informent pas sur les erreurs des élèves et les non-réponses ne pourront pas être codées. Le tableau 2 présente deux exemples (un pour chacune des tâches) afin de visualiser les réponses écartées et celles qui seront analysées.

Tableau 2 : Distinction entre réponses analysées et réponses écartées

Tâche 1	Tâche 2
<p>Non réponse Non utilisée pour l'analyse</p>  <p>Réponse correcte Non utilisée pour l'analyse</p> <p>Réponse incomplète A analyser</p>	 <p>Réponse erronée A analyser</p>

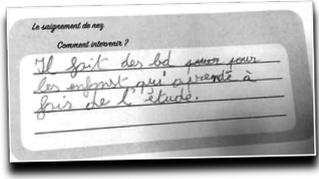
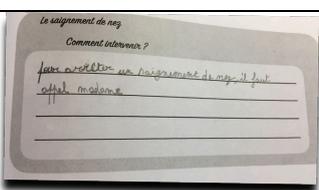
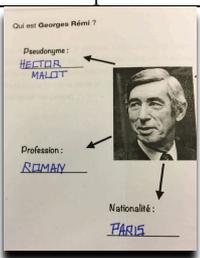
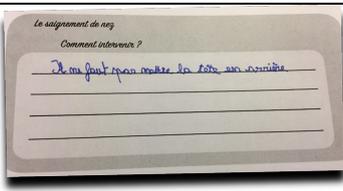
Lors de la deuxième étape de la méthode de Bardin, considérée comme centrale dans l'analyse de contenu, nous avons créé une grille de catégorisation personnelle. Le découpage du contenu peut être effectué en fonction d'unités de codage, entendues comme le « segment de contenu à considérer comme unité de base en vue de la catégorisation et du comptage fréquentiel » (Bardin, 2001, p. 135). Le chercheur peut se baser sur un critère sémantique (basé sur le sens) ou sur un critère syntaxique (mot, phrase, paragraphe ou message) pour déterminer son unité de codage. Dans le cadre de cette étude, c'est le critère syntaxique qui a été choisi. Ce choix repose sur le fait que les mots formant les réponses erronées ou incomplètes peuvent être directement associés à une catégorie.

Notre grille de catégorisation s'appuie, premièrement, sur la typologie des erreurs réalisée par Astolfi (1997) pour définir les catégories principales. Elle a été choisie pour sa vision compréhensive de la nature de l'erreur. De ce fait, ce modèle théorique a été opérationnalisé en une grille de codage reprenant ses 8 types d'erreurs en catégories principales. Ensuite, afin de définir des indicateurs, ces catégories principales ont été subdivisées en catégories secondaires en lien avec la littérature sur les stratégies et sur les difficultés des élèves lors d'une recherche d'information en ligne. Une attention particulière a donc été portée à la définition opérationnelle des catégories et des sous-catégories de codage afin de renforcer l'objectivité du codeur.

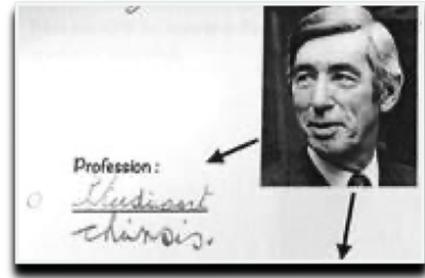
Le tableau 3 reprend les 8 catégories principales, leurs sous-catégories, les indicateurs se rapportant à chaque catégorie et sous-catégorie, ainsi que des exemples de réponses analysées dans le cadre de cette recherche.

Deux des 8 catégories d'erreurs d'Astolfi n'ont pas été utilisées dans notre analyse : « Mauvais décodage des attentes » et « Autre discipline impliquée ». En effet, nous n'avons pas, pour ces deux catégories, identifier de critère directement observable sur les feuilles-réponses des élèves.

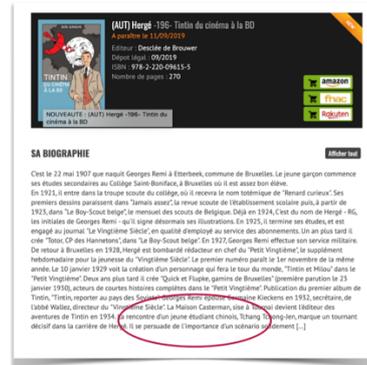
Tableau 3 : Grille de catégories

Typologie d'Astolfi	Sous-catégorie		Exemples
	Difficultés des élèves	Indicateurs	
Compréhension de la consigne	Difficulté de décodage de la consigne	de la réponse donnée ne répond pas à l'attendu	 <p>L'élève n'a pas compris la question.</p>
Mauvais décodage des attentes	Lecture de toute la page	<i>La catégorie « Décodage erroné des attentes » n'est pas observable à partir des documents analysés.</i>	
Conceptions alternatives	Utilisation des connaissances antérieures sans vérifier	La réponse donnée ne se trouve pas dans les documents proposés à l'élève	 <p>L'élève répond sans consulter de documents.</p>
Opérations intellectuelles	Repérage des informations pertinentes	La réponse contient des informations contenues dans des sites non pertinents	 <p>L'élève prend de l'information sur un site non pertinent pour répondre à la question.</p>
	Erreurs de retranscription de la réponse trouvée	La réponse est erronée, car l'élève a mal recopié l'information présente sur le document	 <p>L'élève ne recopie pas correctement l'information trouvée.</p>
Démarches adoptées	Réponse prise dans une partie des informations présentes	La réponse est incomplète, car elle ne reprend que les réponses d'une partie des documents à disposition de l'élève.	 <p>L'élève n'a retranscrit que partiellement le document.</p>

Surcharge cognitive Difficulté de compréhension des informations lues La réponse comprend une erreur due à la mauvaise compréhension du document

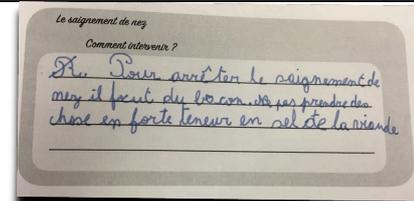


L'élève a mal interprété une partie des informations du document.



Autre discipline impliquée Méconnaissance du moteur de recherche Difficultés propres au contenu

La catégorie « Méconnaissance du moteur de recherche » n'est pas observable à partir des documents analysés La réponse reprend des informations contenues dans des documents non fiables



L'élève consulte un lien non fiable.

Lecture d'un seul lien La réponse n'est pas complète car l'élève a transcrit des informations ne se trouvant que sur un seul lien.

Vient ensuite l'application de la grille de catégories. Il s'agit de classer les éléments constitutifs d'un ensemble par différenciation, puis de les regrouper par analogie d'après des critères définis, afin de fournir une représentation simplifiée des données brutes (Bardin, 1977). Chaque réponse a donc été codée manuellement en fonction des catégories et des sous-catégories présentées dans le tableau 3, mais également en fonction du type relevé lors de la préanalyse (phase 1) : réponse incomplète et réponse erronée. Les erreurs demandant une part de subjectivité lors du codage n'ont pas été utilisées pour l'analyse comme le montre la figure 3.

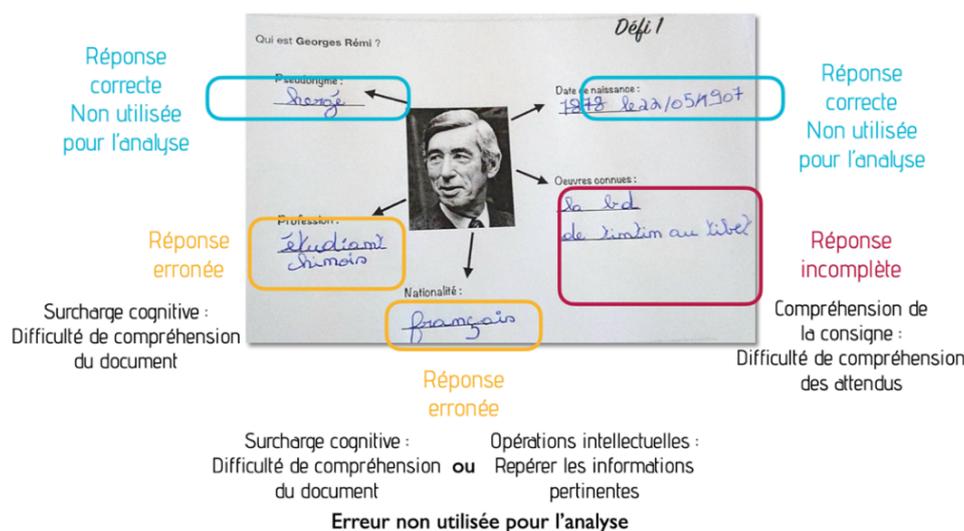


Figure 3 : Distinction entre réponses analysées et réponses écartées

Ainsi lors de la phase de traitement, interprétation et inférence, les données brutes sont traitées de manière à être significatives et valides. À partir de ces données, les fréquences exprimées en pourcentages (calculés par rapport à l'effectif total des erreurs) nous permettent d'établir des tableaux de résultats et mettent en relief les informations apportées par l'analyse (Bardin, 1977). Dans cette recherche, la fréquence d'apparition du critère dans les feuilles-réponses des élèves est l'indicateur qui a été considéré.

Pour plus de rigueur, nous avons soumis nos résultats à un test de validité via un contrôle de la fiabilité intercodeurs. Afin de contrôler la fiabilité de la grille de codage ainsi que du codage effectué, nous avons réalisé un double codage, tel qu'il est préconisé par Huberman et Miles (1991). Nous avons, tout d'abord, effectué notre codage des différentes erreurs relevées dans les productions des élèves. Après ce premier codage des données à partir de la grille des codes, nous avons fait coder séparément une partie des données (300 erreurs) par un chercheur extérieur à la recherche. La comparaison des codes attribués par les deux chercheurs a permis de définir la fiabilité du codage à partir de la formule suivante :

$$\text{Fiabilité} = \frac{\text{nombre d'accords}}{\text{nombre total d'accords} + \text{désaccords}}$$

Le double codage a donc été effectué sur 26,36% des données recueillies. Nous avons obtenu dans un premier temps une validité de 93,55%, ce qui constitue une fiabilité très satisfaisante. Après avoir discuté des désaccords, nous avons obtenu un rapport de 100 %.

4 Analyse des résultats

4.1 Préanalyse

A partir des données brutes, nous avons tout d'abord dissocié les réponses correctes des réponses incorrectes, incomplètes et des non-réponses, et ce pour les deux tâches de recherche. Le tableau 4 les distingue en fonction des deux tâches de recherche analysées.

Tableau 4 : Nombre d'erreurs

	Les réponses non analysées			Les réponses retenues pour l'analyse	
	N	Les réponses correctes	Les non-réponses	Les réponses erronées	Les réponses incomplètes
Tâche 1	1560	999	133	218	210
Tâche 2	863	36	105	326	396
Total	2423	1035	238	544	606
		1273 (52,53%)		1150 (47,47%)	

Sur les 2423 réponses des élèves aux deux tâches de recherche, nous avons écarté 1273 (52,53%) réponses (voir tableau 2). 1035 d'entre elles sont des réponses correctes et 238 sont des non-réponses. En examinant la répartition de ces réponses en fonction de l'âge des participants, aucune tendance descriptive marquée ne se dégage en faveur d'un groupe d'âge particulier. En effet, le pourcentage des réponses correctes se répartit de manière relativement équilibrée : 34,56 % pour les 12-13 ans, 33,87 % pour les 10-11 ans, et 31,57 % pour les 8-9 ans. De même, les non-réponses suivent un schéma similaire : 32,44 % pour les 12-13 ans, 33,12 % pour les 10-11 ans, et 34,44 % pour les 8-9 ans.

En fin de cette première étape de l'analyse, nous obtenons 1150 unités traitées, soit 47,47% des réponses des élèves.

4.2 Classification

Sur les 1150 erreurs observées, 12 unités n'ont pas pu être codées. En effet, l'erreur pouvait être associée à plusieurs catégories ou sous-catégories (Figure 4). Afin d'éviter toute subjectivité et extrapolation, nous avons décidé de les retirer de notre analyse. En conclusion, ce sont 1138 erreurs qui ont pu être codées grâce à notre analyse de contenu. Le tableau 5 reprend la répartition des erreurs en fonction de la tâche et du type d'erreur.

Tableau 5 : Nombre d'erreurs à analyser

	Les réponses retenues pour l'analyse	
	Les réponses erronées	Les réponses incomplètes
Tâche 1	213	210
Tâche 2	319	396
Total	532 (46,75%)	606 (53,25%)
	1138 (100%)	

Sur ces 1138 réponses à traiter, 532 (46,75%) sont des réponses erronées et 606 (53,25%) sont des réponses incomplètes. Le tableau 5 nous informe également que 423 (37,17%) d'entre elles proviennent de la tâche 1 et 715 (62,83%) de la tâche 2.

4.3 Traitement

Les données ont ensuite été traitées en fonction des catégories, des deux tâches proposées et des différents âges afin de pouvoir répondre à nos trois questions de recherche.

4.3.1 Quelles sont les erreurs commises par les élèves lorsqu'ils recherchent de l'information en ligne ?

Le tableau 6 reprend les fréquences d'apparition d'erreurs dans les différentes catégories principales et secondaires.

Tableau 6 : Fréquences d'apparition des erreurs en fonction des catégories et sous-catégories

Catégories Astolfi (1997)	Sous-catégories Difficultés des élèves	Fréquences d'apparition
Compréhension de la consigne	Difficulté de décodage/compréhension de la consigne	176 (15,47 %)
Mauvais décodage des attentes	Lecture de toute la page	/
Conceptions alternatives	Utilisation des connaissances antérieures sans les vérifier	70 (6,15 %)
Opérations intellectuelles	Repérage des informations pertinentes	42 (3,69 %)
	Erreur de retranscription de la réponse trouvée	26 (2,84 %)
Démarches adoptées	Réponse prise dans une partie des informations présentes dans un document	114 (10,02 %)
Surcharge cognitive	Difficulté de compréhension de l'information lue	139 (11,80 %)
Autre discipline impliquée	Méconnaissance du moteur de recherche	/
Difficultés propres au contenu	Repérage des informations fiables	342 (30,05 %)
	Lecture d'un seul lien	229 (20,12 %)
TOTAL		1138 (100,00 %)

Parmi les huit sous-catégories présentées dans le Tableau 6, deux montrent une prédominance significative (plus de 20 % d'apparition). Il s'agit de « Repérer les informations fiables » avec 30,05 % et de « Lecture d'un seul lien » avec 20,12 %, qui sont les sous-catégories les plus fréquemment observées.

Les trois sous-catégories suivantes, représentant entre 10 % et 20 % des apparitions, sont : « Difficulté de compréhension de la consigne » (15,47 %), « Difficulté de compréhension de l'information lue » (11,80 %), et « Prise d'informations dans une partie du document » (10,02 %).

Les trois dernières sous-catégories apparaissent dans moins de 7 % des réponses analysées : « Utilisation des connaissances antérieures sans les vérifier » (6,15 %), « Repérer les informations pertinentes » (3,69 %), et la sous-catégorie la moins représentée, « Erreur de recopiage de la réponse trouvée » (2,84 %).

4.3.2 Les deux tâches de recherche proposées différencient-elles les erreurs commises ?

Le tableau 7 reprend les fréquences d'apparition des types d'erreurs en fonction des deux tâches de recherche

Tableau 7 : Fréquences d'apparition des erreurs en fonction des deux tâches de recherche

Catégories Astolfi (1997)	Sous-catégories Difficultés des élèves	Fréquences d'apparition	
		Tâche 1 <i>Tâche de localisation</i>	Tâche 2 <i>Tâche d'interprétation</i>
Compréhension de la consigne	Difficulté de décodage de la consigne	166 (39,24 %)	10 (1,40 %)
Mauvais décodage des attentes	Lecture de toute la page	/	/
Conceptions alternatives	Utilisation des connaissances antérieures sans les vérifier	26 (6,15 %)	44 (6,15 %)
Opérations intellectuelles	Repérage des informations pertinentes	42 (9,93 %)	00 (0,00 %)
	Erreur de retranscription de la réponse trouvée	26 (6,15 %)	00 (0,00 %)
Démarches adoptées	Réponse prise dans une partie des informations présentes dans un document	21 (4,96 %)	93 (13,01 %)
Surcharge cognitive	Difficulté de compréhension de l'information lue	125 (29,55 %)	14 (1,59 %)
Autre discipline impliquée	Méconnaissance du moteur de recherche	/	/
Difficultés propres au contenu	Repérage des informations fiables	00 (0,00 %)	342 (47,83 %)
	Lecture d'un seul lien	17 (4,02 %)	212 (29,65 %)
TOTAL		423 (100,00 %)	715 (100,00 %)

Lors de notre analyse descriptive, nous constatons que cette comparaison met tout d'abord en avant un nombre total plus élevé d'erreurs pour la deuxième tâche (62,83 % des erreurs pour la tâche 2 ; N=715 vs 37,17 % pour la tâche 1 ; N=423). L'application d'un « test de Mann-Whitney » indique cependant que cette différence est non significative ($p=0.958$).

On remarque néanmoins des écarts de fréquence entre la répartition des sous-catégories d'erreurs au sein des deux tâches de recherche.

Pour la tâche 1, dite de complexité faible, on remarque que les élèves obtiennent des fréquences d'apparition plus élevées par rapport à celles de la tâche 2 pour les erreurs de compréhension de la consigne (39,24 % des erreurs commises pour la tâche 1 vs 1,40 % des erreurs de la tâche 2), de repérage des informations pertinentes (9,93 % vs 00, 00 %), de compréhension des informations lues (29,55 % vs 1,59 %) et de retranscription de l'information (6,15 % vs 0, 00 %).

En ce qui concerne la tâche 2, on retrouve des différences sur les erreurs où la réponse n'est issue que d'une partie du document (tâche 1 : 4,96 % vs tâche 2 : 13,01 %), le repérage des informations non fiables (0, 00 % vs 47,83 %) et l'élaboration de la réponse en ne consultant qu'un seul lien (4,02 % vs 29,65 %).

L'utilisation des connaissances antérieures montre, quant à elle, des fréquences d'apparition identique pour les deux tâches (6,15 % vs 6,15 %).

4.3.3 L'âge des apprenants différencie-t-il les erreurs commises ?

Si on compare les erreurs commises en fonction de l'âge des élèves (Tableau 8), on constate que le nombre d'erreurs est inversement proportionnel à l'âge des élèves. En effet, ce sont les 8-9 ans qui commettent le plus d'erreurs ($n=563$; 49,47%) et les élèves de 12-13 ans qui en réalisent le moins ($n=218$; 19,15%).

Tableau 8 : Fréquences d'apparition des erreurs en fonction des âges

Catégories	Sous-catégories	Groupe 1 8-9 ans	Groupe 2 10-11 ans	Groupe 3 12-13 ans
Compréhension de la consigne	Difficulté de décodage/compréhension de la consigne	103 (18,29 %)	53 (14,85 %)	20 (9,17 %)
Mauvais décodage des attentes	Lecture de toute la page	-	-	-
Conceptions alternatives	Utilisation des connaissances antérieures sans les vérifier	32 (5,68 %)	21 (5,88 %)	17 (7,79 %)
Opérations intellectuelles	Repérage des informations pertinentes	21 (3,73 %)	11 (3,08 %)	7 (3,21 %)
	Erreur de retranscription de la réponse trouvée	18 (3,19 %)	5 (1,40 %)	3 (1,38 %)
Démarches adoptées	Réponse prise dans une partie des informations présentes dans un document	54 (9,59 %)	37 (10,36 %)	23 (10,55 %)
Surcharge cognitive	Difficulté de compréhension de l'information lue	86 (15,28 %)	40 (11,20 %)	13 (5,96 %)
Autre discipline impliquée	Méconnaissance du moteur de recherche	-	-	-
Difficultés propres au contenu	Repérage des informations fiables	144 (25,58 %)	111 (31,09 %)	90 (41,28 %)
	Lecture d'un seul lien	105 (18,65 %)	79 (22,13 %)	45 (20,64 %)
Total		563 (100,00 %)	357 (100,00 %)	218 (100,00 %)

Les résultats du « test de Kruskal-Wallis » indiquent, cependant, une p-valeur de 0.087 suggérant qu'il n'y a pas de différence significative entre les groupes concernant le nombre d'erreurs commises. Cette conclusion est appuyée par les résultats du « post hoc test », qui a comparé les groupes deux à deux et n'a révélé aucune différence significative, quelle que soit la comparaison effectuée (Groupe 1 vs Groupe 2 : $p_{\text{tukey}} = 0.372$; Groupe 1 vs Groupe 3 : $p_{\text{tukey}} = 0.078$; Groupe 2 vs Groupe 3 : $p_{\text{tukey}} = 0.629$).

On remarque, néanmoins que, quel que soit l'âge, les erreurs dues aux difficultés propres au contenu sont celles qui obtiennent les moyennes les plus élevées : « Repérage des informations fiables » (Groupe 1 : 25,58 % ; Groupe 2 : 31,09 % ; Groupe 3 : 41,28 %) ; « Lecture d'un seul lien » (Groupe 1 : 18,65 % ; Groupe 2 : 22,13 % ; Groupe 3 : 20,64 %).

Viennent ensuite les erreurs dues au décodage/compréhension de la consigne pour les élèves de 8 à 11 ans (Groupe 1 : 18,29 % ; Groupe 2 : 14,85 %) et la prise d'information dans une partie du document (le début du document) pour les élèves plus âgés (Groupe 3 : 10,55%).

On observe également que les « Difficultés de compréhension de l'information lue » obtiennent un taux élevé d'erreurs pour les élèves plus jeunes (Groupe 1 : 15,28 % et Groupe 2 : 11,20 %).

5 Discussion et conclusions

D'un point de vue épistémologique, nous avons pris appui sur le postulat qui met en avant que l'analyse d'erreurs peut permettre une meilleure analyse des stratégies de l'élève et une meilleure compréhension de leur raisonnement. Notre description des erreurs des élèves lors d'une recherche d'information en ligne a donc comme objectif de proposer certaines pistes didactiques pour l'enseignement spécifique de cette compétence.

Astolfi (1997) a élaboré une typologie des erreurs dite utilisable pour différentes disciplines scolaires. Nous confirmons, tout d'abord, la pertinence de celle-ci pour les tâches de recherche d'informations en ligne. Cette pertinence va permettre de répondre à notre objectif : adapter les stratégies d'intervention des enseignants aux besoins de chaque apprenant. Ainsi, en partant des erreurs des élèves, nous définissons les obstacles à l'apprentissage qui peuvent être traduits en objectifs didactiques et en aides proposées à l'élève.

Lors de cette étude, 1138 erreurs au total ont pu être analysées au regard de cette typologie. Ce taux d'erreurs montre que, malgré que le numérique offre des facilités d'accès à l'information, ces nouvelles pratiques ne sont pas pour autant aisées pour les apprenants de 8 à 13 ans. Notre recherche confirme la nécessité d'un encadrement pédagogique auprès de ce public, comme l'affirme la FW-B dans son référentiel numérique ou encore les auteurs comme Kingley *et al.* (2015), Perez *et al.* (2018) ou encore Delhougne (2021). Pour que les lecteurs puissent consulter, de manière efficace, la vaste quantité d'informations disponibles sur le Web, ils doivent être formés à l'utilisation de stratégies adéquates de navigation, de recherche et d'évaluation (Cho *et al.*, 2017).

Cette analyse des erreurs permet également la détermination des types de compétences que visera chacune des aides didactiques proposées lors d'une étape spécifique de la recherche d'information. En effet, nous pouvons relier chacune des catégories d'Astolfi à des erreurs spécifiques d'élèves lorsqu'ils réalisent une recherche en ligne. Ainsi, si nous utilisons les étapes de la recherche d'information de Brand-Gruwel (2009), il est possible de spécifier, pour chacune de celles-ci, le type d'erreurs qu'il est possible de commettre. Le tableau 9 synthétise les erreurs des apprenants en fonction du type d'erreurs d'Astolfi et du moment à partir duquel elles se produisent.

Tableau 9 : Les erreurs des apprenants en fonction du type d'erreurs d'Astolfi et du moment auquel elles se produisent

Étapes IPS (Brand-Gruwel et al., 2009)	Typologie d'Astolfi (1997)	Erreurs rencontrées lors de notre analyse
Définir le problème	Compréhension de la consigne	Erreurs dues à la mauvaise compréhension du problème de recherche d'information - vocabulaire non compris, mauvaise compréhension du texte.
Balayer l'information	Opérations intellectuelles	Erreurs dues au repérage des informations pertinentes.
	Difficultés propres au contenu	Erreurs dues à la vérification de la fiabilité des informations.
Traiter l'information	Surcharge cognitive	Erreurs dues à la mauvaise compréhension du document lu - vocabulaire non compris, mauvaise compréhension du texte.
	Difficultés propres au contenu	Erreurs dues à la prise d'information d'un seul lien.
	Conceptions alternatives	Erreurs dues à l'utilisation des connaissances antérieures pour répondre au besoin d'informations sans les vérifier en faisant une recherche en ligne.
	Démarches adoptées	Erreurs dues à la prise d'information dans une seule partie du document.
Organisation de l'information	Opérations intellectuelles	Erreur de retranscription

Enfin, l'analyse d'erreurs constitue un outil privilégié. Elle différencie les erreurs selon leur nature (connaissances, stratégies cognitives générales ou spécifiques au domaine) et permet donc de différencier les interventions pédagogiques.

5.1 Agir avant de rechercher l'information en ligne pour prévenir certaines erreurs

Notre analyse laisse apparaître que les élèves commettent des erreurs en termes de compréhension de l'information recueillie (11,80 % des erreurs commises) et en termes de réponse élaborée à partir d'une lecture partielle du document (10,02 %).

On remarque que les difficultés de compréhension de l'information recueillie apparaissent, auprès de nos trois groupes d'âge, avec une diminution des erreurs de compréhension avec l'âge. Lorsqu'on compare les deux tâches proposées, la fréquence de ce type d'erreurs est plus élevée lors de la première tâche. Nous expliquons ce constat par la précision de la réponse attendue et le type de documents à lire pour répondre à cette première tâche (Tâche de localisation : Kim, 2009 ; Walraven *et al.*, 2009). Notre étude confirme donc que les

compétences en compréhension de texte impactent la performance d'une recherche en ligne (Kingsley *et al.*, 2015), quel que soit l'âge ou le type de tâche proposée (Tâche de localisation - Tâche d'interprétation).

Les erreurs dues à l'absence de lecture de l'ensemble du document (Lecture partielle du document) apparaissent de manière quasiment similaire quels que soient les groupes d'âge et se révèlent plus fréquentes lors de la tâche 2. Nous pouvons expliquer ce constat par le fait que les élèves s'appuient dès qu'ils le peuvent sur des solutions de résolution plus simples à mettre en œuvre, qui ne leur demandent que peu d'engagement sur le plan cognitif (Merrill et Twitchell, 1994). Les apprenants ont une tendance innée à réduire la charge cognitive autant que possible. Ainsi, dans le cadre d'une recherche en ligne, dès que les élèves ont l'impression d'avoir trouvé une information qui répond à leur besoin, ils interrompent leur lecture du document. Ils passent, de ce fait, très peu de temps sur le contenu d'un lien (Loffreda, 2011 ; Fives et Dinsmore, 2017). Nos résultats montrent la persistance de ce phénomène, quel que soit l'âge des apprenants, et son augmentation en lien direct avec le degré de difficulté de résolution de la tâche.

Les interventions pédagogiques concernant la recherche d'information en ligne doivent donc être mises en œuvre avec la même précision que celle exigée par la lecture sous un format papier. Ainsi, pour rechercher efficacement des informations en ligne, il faut, premièrement, former les élèves aux caractéristiques spécifiques des genres numériques : reconnaître et distinguer un site d'un billet de blog, un message d'un forum d'un article de wiki, etc. afin qu'ils puissent les discriminer pour en identifier l'objectif et les intentions des auteurs (Rouet, 2016 ; Renaud, 2020 ; FW-B, 2022).

De plus, l'élève doit pouvoir passer rapidement en revue certains blocs de sens pour éventuellement ne pas les considérer immédiatement et plutôt se concentrer sur d'autres parties du texte. Ainsi, l'élève doit intégrer le fait que la lecture d'une page Web peut s'avérer différente de celle d'un roman (Renaud, 2020). Il faut pouvoir doter les élèves de connaissances structurelles (le document est composé de...) et fonctionnelles (le sommaire sert à ...) des documents présents en ligne pour qu'ils puissent orienter leur lecture en fonction de leur objectif de recherche d'information (Rouet, 2016). Notons également que ces connaissances ont une importance cruciale sur les écrans où la capacité d'affichage est moindre que sur les pages d'un texte imprimé. Le lecteur doit donc s'appuyer sur ces connaissances (le sommaire, par exemple) pour identifier les éléments qui n'apparaissent pas à l'écran, mais vers lesquels il peut se diriger en utilisant la souris (Renaud, 2020).

Le scénario didactique mis en œuvre par l'enseignant peut donc ici être structuré autour d'activités débranchées qui permettront un traitement des documents numériques sortis de leur contexte en ligne afin de pouvoir les analyser de manière spécifique.

5.2 Agir au début de la recherche d'information en ligne pour avoir une réflexion sur le besoin informationnel

Notre analyse permet de mettre en évidence des erreurs en termes de compréhension de la consigne (15,47 % des erreurs commises), de l'utilisation des connaissances antérieures pour répondre sans vérification de celles-ci (6,15 %) et également de prise d'informations qui ne sont pas en lien avec le problème posé (3,69 %).

Deux de ces 3 types d'erreurs (Compréhension de la consigne et utilisation des connaissances antérieures) sont apparus lors des deux tâches de recherche proposées. Nous pouvons émettre l'hypothèse que ces deux erreurs peuvent être considérées comme récurrentes lors d'une recherche en ligne pour le public cible. La prise d'information non pertinente, quant à elle, n'est présente que pour la tâche qui ne demande aucune transformation de l'information recueillie (tâche 1 ; tâche de localisation). Nous pouvons expliquer ce constat par le fait que ce type de tâche demande des réponses précises. Ainsi, l'importance des mots-clés semble plus décisive lorsque la tâche exige une réponse spécifique afin d'obtenir un maximum de liens pertinents sur la page de résultats (Kim, 2009 ; Walraven *et al.*, 2009). Nous avons également constaté que ces trois types d'erreurs pouvaient être présentes chez nos trois groupes d'âge. En ce qui concerne la compréhension de la consigne, il est à noter que la proportion d'erreurs commises diminue avec l'âge. Ce constat peut être justifié par le caractère évolutif de la compétence réceptive qu'est la lecture (Dehaene, 2007).

Ces trois types d'erreurs apparaissent à des moments différents lors d'une recherche en ligne, mais peuvent être, du point de vue pédagogique, travaillées au début de l'activité. Des interventions, au début de l'apprentissage dans l'objectif de prévenir ces trois types d'erreurs, peuvent donc être envisagées. En effet, l'absence d'une réflexion suffisante sur le besoin d'information, avant de se lancer dans la recherche, peut engendrer chez l'élève la perte de celui-ci en mémoire de travail (Rouet *et al.*, 2011).

Au vu de nos résultats, nous pouvons anticiper ces erreurs dans une première activité de type « définition du problème ». Cette activité doit être orientée vers la lecture du problème posé, la formulation de questions autour du problème posé, l'activation des connaissances antérieures des élèves et la détermination des informations attendues (Brand-Gruwel *et al.*, 2009).

Lors de cette étape préparatoire à la recherche d'information, le niveau de lecture va influencer la performance de l'élève. Nos résultats montrent que plus l'élève est jeune plus un travail de compréhension sur son besoin initial d'information est nécessaire. Il semble crucial que l'enfant comprenne d'abord le vocabulaire qui compose la consigne. Celle-ci est le premier élément qui permet à l'enfant d'entrer dans l'activité (Astolfi, 1997). L'enfant doit d'abord comprendre la consigne, se l'approprier et savoir quel est l'objectif de cette tâche de recherche d'information afin de définir une intention ou un but de lecture et se représenter la tâche qu'il aura à mener.

Ainsi, après avoir fixé la compréhension des termes de la consigne, Mottet et Fatoux (2014) proposent la méthode des 3QPOC (ou Qui, Quoi, Quand, Pourquoi, Où, Comment) afin de définir le travail et de cerner l'objet de la recherche d'information. Au travers des questionnements « Quel est le thème de la recherche ? Quel est le type de production à réaliser ? Quel est le public cible ? Quel est l'objectif de la production ? Quelles sont les consignes à respecter ? Quelles sont les étapes de travail à franchir ? », l'élève peut ainsi clarifier la recherche qu'il souhaite effectuer avant de se lancer dans l'encodage des mots-clés dans le moteur de recherche. Avec les réponses aux différentes questions du 3QPOC, il est alors possible de structurer les informations, sous forme de cartes mentales par exemple. Cela permet d'avoir une vision d'ensemble et donc de faciliter les recherches. C'est ce que Rouet et Britt (2011) nomment la construction du modèle de la tâche. Cette première étape de lecture en ligne permet de définir ce que Renaud (2020) appelle « ce que je cherche ». Formuler des réponses précises à ces questions aidera les élèves à garder en mémoire leur intention de lecture. Ce qui leur permettra de ne pas s'éparpiller pendant le processus de

recherche, de ne pas se décourager en cours de route et d'identifier des sources d'information pertinentes tout en étant guidés par une lecture sélective et finalisée orientée vers un but (Rouet *et al.*, 2011). C'est également l'occasion de mobiliser les compétences encyclopédiques des élèves en faisant émerger leurs connaissances antérieures sur le thème afin de les confronter à celles des pairs et ensuite aux informations contenues dans les documents. Cette confrontation au texte et à l'intention de lecture permet de représenter et anticiper le type d'information que le texte peut contenir, cela incite à choisir de le prendre en compte ou non et d'en évaluer sa pertinence (Renaud, 2020).

5.3 Agir en fin de recherche d'information en ligne en mettant l'apprenant en position de correction active

Les erreurs des élèves peuvent également concerner la retranscription de l'information (2,84 %), la vérification de la fiabilité du document consulté (30,05 %) et la confrontation de la réponse trouvée avec d'autres documents (20,12 %).

À y regarder d'un peu plus près, les erreurs de transcription sont les erreurs les moins fréquentes qu'ont commises les sujets de notre échantillon. Elles apparaissent surtout chez les élèves les plus jeunes et seulement lors de la tâche 1. Ce type d'erreur concerne essentiellement des dates ou des noms. Par contre, le manque de vérification de la fiabilité ainsi que la non-confrontation des multiples sources représentent les erreurs les plus fréquemment apparues. Nos analyses montrent qu'elles sont présentes pour toutes les tranches d'âge et de manière plus spécifique pour la tâche 2 qui exige de résumer l'information.

Pour pallier ces deux types d'erreurs, du point de vue pédagogique, la littérature propose plusieurs solutions. Les auteurs parlent d'auto-évaluation et de lecture latérale. Une première option pédagogique est de proposer une grille d'évaluation (ou check-list) à l'élève pour lui permettre de se positionner au regard d'un certain nombre de critères afin de vérifier sa démarche (Mottet et Fatoux, 2014). Celle-ci peut être remplie en répondant à des questions sur l'identification de l'auteur, l'objectif du site, les références, le contenu, la présentation, la facilité de navigation, et d'autres critères importants. En utilisant cette grille, les élèves peuvent prendre une décision éclairée quant à l'utilité et à la fiabilité du site, ce qui leur permettra de mieux utiliser les ressources disponibles en ligne.

Une deuxième possibilité (Programme CTRL ; Pavlounis *et al.*, 2021) consiste à (1) vérifier la source de l'information, en se posant des questions telles que «Qui a publié cette information?» ou «Quelle est la réputation de cette source, de cette personne ou de ce groupe?» ; (2) vérifier l'affirmation, en se demandant si des sources fiables affirment la même chose ou si des vérificateurs de faits ont déjà établi que l'affirmation était fautive ; (3) retrouver la source originale de l'information, en se demandant si le contenu s'est transformé au fil des réécritures, des partages et des reformulations. L'approche de la lecture latérale préconisée par le programme CTRL-F implique que l'élève quitte momentanément le site qu'il consulte afin d'évaluer sa fiabilité en effectuant une recherche en parallèle. Cette recherche en parallèle peut consister à rechercher des informations sur l'auteur, à vérifier si les informations retenues sont également présentes ailleurs ou à retrouver l'information originelle (comme une image, par exemple). L'objectif est de trouver des informations contextuelles pertinentes pour évaluer les sources.

En ce qui concerne les erreurs de retranscription de l'information, Goigoux (2017) souligne l'impératif d'enseigner aux élèves la technique de recopie. Cela implique de démontrer la

segmentation afin de faciliter la mémorisation, d'expliciter les procédures de maintien de la mémoire, telles que l'autorépétition, et de sensibiliser aux points de vigilance orthographiques.

En conclusion, la catégorisation et la classification d'erreurs des élèves lors d'une recherche d'information en ligne se sont révélées intéressantes. Dans une logique descriptive, elles nous ont permis de constater la prépondérance de difficultés spécifiques au contenu (évaluation de la fiabilité et confrontation des documents). Mais également, que l'âge ainsi que le type de tâche à réaliser pouvaient moduler le type d'erreur commise. Les pistes didactiques proposées ci-dessus tiennent compte des sources des erreurs observées et ont l'objectif d'être concrètes afin d'être menées en contexte scolaire. De plus, ces pistes didactiques ont le souci d'une intégration spécifique des activités de recherche en ligne en complémentarité à celles de lecture dite traditionnelle sous format papier. Comme pour les études précédentes, ces différents constats pourront fournir des recommandations didactiques basées sur des expérimentations scientifiques ainsi que des exemples concrets pour construire notre contenu de formation à destination des enseignants de la FW-B.

6 Limites et perspectives

Lors de l'interprétation des résultats, deux limites doivent être considérées. Premièrement, il existe un biais potentiel lié à l'intérêt des situations étudiées. La deuxième tâche de recherche, qui consistait à résumer les étapes d'intervention lors d'un saignement de nez, pourrait avoir été perçue par les élèves comme plus utile et pertinente. Cette tâche, renvoyant à une situation concrète et potentiellement vécue, risque d'avoir suscité un engagement plus important de la part des élèves. Ce biais pourrait avoir influencé les résultats en rendant cette tâche plus attrayante, ce qui ne reflète pas nécessairement le comportement des élèves dans des contextes moins engageants.

En second lieu, les compétences rédactionnelles des élèves présentent une autre limite. Le principal défi pour les élèves pourrait ne pas être seulement la lecture et la compréhension des documents, mais aussi leur capacité à rédiger une réponse claire et cohérente à partir de ce qu'ils ont lu. Cette difficulté en production écrite peut fausser l'évaluation des compétences des élèves, menant potentiellement à une sous-estimation de leur véritable niveau de compréhension. Ces deux limites suggèrent des orientations pour des recherches futures, notamment en vue d'affiner l'analyse des compétences des élèves.

Enfin, une dimension qui mérite d'être explorée dans des recherches futures est celle des différences potentielles entre les filles et les garçons dans leur approche de la recherche d'information en ligne. Bien que cette étude n'ait pas spécifiquement abordé cette question, il serait pertinent de déterminer si des différences existent et, le cas échéant, comment elles influencent les performances des élèves. Explorer ces variations pourrait fournir des pistes pour adapter les stratégies d'enseignement en fonction des besoins spécifiques de chaque groupe, tout en tenant compte des dynamiques de genre dans l'apprentissage numérique.

7 Références bibliographiques

- Algubbi, S. (2016). L'erreur : un outil fondamental dans la classe de FLE. *Norsud*, 7,103-116.
- Astolfi, J.-P.(1997). *L'erreur, un outil pour enseigner*. Paris.
- Baddeley, A. (2012) Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63, 1-29.
- Bardin, L. (1977). *L'analyse de contenu*. Presses Universitaires de France.
- Boubée, N. et Tricot, A. (2010). *Qu'est-ce que rechercher de l'information?* Villeurbanne, Presses de l'ENSSIB.
- Brand-Gruwel S., Wopereis I. et Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, 53, 1207-1217.
- Bricteux, S., Crépin, F., Quittre, V. et Lafontaine, D. (2021). Résultats de PISA 2018 en Fédération Wallonie-Bruxelles. Les jeunes de 15 ans et la lecture. *Les Cahiers des Sciences de l'éducation*, 45.
- Britt, M. A. et Rouet, J. F. (2012). Learning with multiple documents: Component skills and their acquisition. Dans J. Kirby et M. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (p. 276–314). Cambridge University Press.
- Cho, B-Y., Woodward, L., Li, D. et Barlow, W. (2017). Examining adolescents' strategic processing during online reading with a question generating task, *American Educational Research Journal*, 54, 691–724.
- Collard, Y., Culot, M., Debrabandere, T., Guffens, B., Huys, S., Lugurgo, V., Orban de Xivry, A-C. et Verniers, P. (2021). #Génération 2020. *Les jeunes et l'info*. Conseil supérieur de l'Éducation aux Médias de la FWB.
- Cuq, J-P, et Gruca, I. (2002). *Cours de didactique du français langue étrangère et seconde*. Grenoble : Presses Universitaires de Grenoble.
- Dehaene, S. (2007). *Les neurones de la lecture*. Paris : Odile Jacob.
- Delhougne, C. (2021). *En littératie numérique, l'implémentation, en première année de l'enseignement secondaire, de séquences pédagogiques pour un apprentissage explicite de la navigation et de la recherche d'informations permet-elle de développer ces compétences liées à la lecture numérique ?* [Mémoire de master]. Université de Liège.
- DeStefano, D. et LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior*, 23(3), 1616–1641.
- De Vries, H., Mesters, I., van de Steeg, H. et Honing, C. (2005). The general public's information needs and perceptions regarding hereditary cancer: an application of the Integrated Change Model, *Patient Education and Counseling*, 56(2). <https://doi.org/10.1016/j.pec.2004.01.002>.
- Digital Wallonia. (2018). *Baromètre 2023 de maturité numérique des citoyens wallons : Équipement, usages et compétences numériques des citoyens wallons*. Digital Wallonia. <https://www.digitalwallonia.be/fr/publications/citoyens-2023/>.
- Dinet, J., De Cara, B., Therouanne, P., Chanquoy, L., Rouet, J-F., Tricot, A., Vivian, R. et Dumercy, L.(2010). L'utilisation des moteurs de recherche par les jeunes : Impact des connaissances du domaine et des connaissances procédurales sur les stratégies d'exploration visuelle. Communication présentée au 7ème Colloque Technologies de l'Information et de la Communication pour l'Enseignement, Nancy.
- Dinet, J. et Tricot, A. (2008). Recherche d'information dans les documents électroniques. Dans A. Chevalier (Eds.), *Ergonomie des documents électroniques* (pp. 35-69). Presses Universitaires de France. <https://doi.org/10.3917/puf.cheva.2008.02.0035>.

- Dumouchel, G. (2016). *Les compétences informationnelles des futurs enseignants du Québec sur le Web*. [Thèse de doctorat, Université de Montréal].
- Eisenberg, M. B. et Berkowitz, R. E. (1988). *Curriculum initiative: An agenda and strategy for library media programs*. Ablex Publishing Corporation.
- Fives H. et Dinsmore D. (2017). *The Model of Domain Learning: Understanding the Development of Expertise*. Routledge.
- Fleury, P. (2016). *Comment montrer aux élèves à faire de bonnes recherches sur le Web*. Chenelière éducation.
- Fournier, H. (2007). *Stratégies de recherche et de traitement de l'information dans des environnements informatiques et sentiment d'efficacité personnelle des futurs enseignants à l'égard de ces stratégies* (Thèse de doctorat non publiée). Université du Québec, Montréal.
- Kingsley, T. L. (2015). *Successfully Promoting 21st Century Online Research Skills: Interventions in 5th-Grade Classrooms*. ScholarWorks at WMU.
https://scholarworks.wmich.edu/reading_horizons/vol54/iss2/5/
- Kim, J. (2009). Describing and Predicting Information-Seeking Behavior on the Web. *Journal of the american society for information science and technology*, 60 (4), 679–693.
- Kuiper, E., Volman, M. et Terwel, J. (2005). The Web as an Information Resource in K-12 Education: Strategies for Supporting Students in Searching and Processing Information. *Review of Educational Research*, 75(3), 285-328.
- Kumps, A., Temperman, B. et De Lièvre, B. (2022). A Decision Tree Approach to Predict the Online Information Search Performance of 8-13 Year Old Students in French-Speaking Belgium. *Turkish Online Journal of Educational Technology*, 21(1), 1-17.
- Langley, A. (1997). Stratégies d'analyse de données processuelles. *Recherches qualitatives*, 17, 206-231.
- Loffreda, M. (2011). Pratiques de recherche d'information sur Internet dans des activités de documentation. Analyse des activités des apprenants dans le contexte scolaire. *EduTice*.
<https://edutice.hal.science/file/index/docid/685338/filename/a1110b.htm>
- Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge University Press.
- Mayer, R. E. et Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6
- Merrill D., Twitchell D. (1994). *Instructional Design Theory*. New Jersey : Educational Technology.
- Miles, M.B. et Huberman, A.M. (2003). Analyse des données qualitatives. De Boeck (2^e éd.).
- Mottet, M. et Fatoux, C. (2014). Former à la recherche d'infos : pourquoi et comment ? *Résonances – Mensuel de l'école valaisanne*, 22-23.
- Muchielli, A. (2006). Les processus intellectuels fondamentaux sous-jacents aux techniques et méthodes qualitatives. Conférence au Colloque international « recherche qualitative : Bilan et prospective ». Béziers, 27-29 juin 2006.
- Mullis, I., Martin, M., Foy, P. et Hooper, M. (2017). *Résultats internationaux d'ePIRLS 2016 en lecture d'information en ligne*. Boston College. <http://timssandpirls.bc.edu/pirls2016/international-results/ISBN-978-1-889938-50-9>.
- Pérez, A., Potocki, A., Stadler, M., Macedo-Rouet, M., Paul, J., Salmerón, L. et Rouet, J-F. (2018). Fostering teenagers' assessment of information reliability: Effects of a classroom intervention focused on critical source dimensions. *Learning and Instruction*, 58, 53-64.
<https://doi.org/10.1016/j.learninstruc.2018.04.006>
- Renaud, J. (2020). Quelles cibles didactiques viser dans l'enseignement de la lecture documentaire sur support numérique au cycle 3 ? *Repères*, 61.

- Rouet, J., Ros, C., Goumi, A., Macedo-Rouet, M. et Dinet, J. (2011). The Influence of Surface and Deep Cues on Primary and Secondary School Students' Assessment of Relevance in Web Menus. *Learning and Instruction*, 21, 205-219.
- Schillings, P. et André, M. (2020). *Analyse Clinique des obstacles rencontrés par des élèves de 4^e et 5^e années primaires face à un texte de lecture numérique*. Communication présentée au Didactif en 2020, Liège.
- Stordeur, J. (2014). *Comprendre, apprendre, mémoriser. Les neurosciences au service de la pédagogie*, De Boeck Education.
- Sung, Y.-T., Wu, M.-D., Chen, C.-K. et Chang, K.-E. (2015). Examining the online reading behavior and performance of fifth-graders: Evidence from eye-movement data. *Frontiers in Psychology*, 6, 1–15.
- Trudel, P. et Gilbert, W. (2000). Compléter la formation des chercheurs avec le logiciel NUD*IST. *Recherches qualitatives*, 20, 87-111.
- Vanderschantz, N., Hinze, A., et Cunningham, S.J (2014). Sometimes the internet reads the question wrong: Children's search strategies & difficulties. *Proceedings of the American Society for Information Science and Technology*, 51(1). <https://doi.org/10.1002/meet.2014.14505101053>
- Vantourout, M. et Goasdoué, R. (2014). Approches et validité psycho-didactiques des évaluations. *Éducation et formation*, e-302, 139-156.
- Vantourout, M. et Blanc, C. (2017). Une approche psycho-didactique pour évaluer la lecture au cours préparatoire ? *Repères*, 56, 215-231.
- Vanssay, S. et Lozac'h, A. (2012). L'erreur pour apprendre, *Cahiers pédagogiques*, 494.
- Walraven A., Brand-gruwel S. et Boshuizen H. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior*, 24(3), 623–648.
- Walraven, A., Brand-Gruwel, S. et Boshuizen, H. (2009). How students evaluate information and sources when searching the World Wide Web for information. *Computers & education*, 52(1), 234-246. <https://doi.org/10.1016/j.compedu.2008.08.003>