

[Page de garde]

HEC MONTRÉAL

**Cognitive Diversity in AI-Generated Text Reading: The Roles of Visual
Propensities and Reading Habits**

par
Anne-Sophie Gaudet

Sylvain Sénécal
HEC Montréal
Codirecteur de recherche

Constantinos Coursaris
HEC Montréal
Codirecteur de recherche

Sciences de la gestion
(Spécialisation Expérience Utilisateur)

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)*

Décembre 2025
© Anne-Sophie Gaudet, 2025

Résumé

Les modèles de langage (LLM) ont profondément transformé notre façon de chercher de l'information. Ils produisent des réponses rapides, structurées et cohérentes, devenant des outils incontournables pour l'accès à l'information. Or, les textes qu'ils génèrent diffèrent fondamentalement de ceux rédigés par des humains. Certaines particularités propres aux LLM sont susceptibles d'influencer la manière dont les lecteurs mobilisent leur attention visuelle et interprètent l'information. La présente étude examine comment les propensions cognitives individuelles façonnent l'attention visuelle et la charge cognitive perçue lors de la lecture de réponses générées par l'IA. Nous nous intéressons à deux propensions : la recherche visuelle orientée vers un objectif, soit la capacité à repérer efficacement l'information pertinente parmi des éléments distracteurs, et la discrimination visuelle, soit l'habileté à distinguer des éléments visuellement similaires. Bien que ces propensions n'aient pas significativement prédit les mesures d'attention visuelle, la recherche visuelle était associée à une charge cognitive perçue plus faible, tandis que la discrimination visuelle correspondait à une charge plus élevée. Les habitudes de lecture ont également révélé des tendances marginales suggérant un possible effet modérateur, atténuant l'impact des propensions visuelles sur la charge cognitive. Ce travail contribue à une compréhension plus fine du traitement cognitif dans la lecture médiée par l'IA et appuie la conception d'interfaces adaptatives tenant compte de la diversité des lecteurs et de leurs profils cognitifs.

Mots-clés: Texte généré par l'IA, attention visuelle, charge cognitive, oculométrie, recherche visuelle, discrimination visuelle, habitudes de lecture, différences individuelles.

Méthodes de recherche: Étude expérimentale, oculométrie, mesure psychométrique, plan intra-sujet, analyse quantitative.

Abstract

Large Language Models (LLMs) have fundamentally reshaped how people search for information. They generate quick and well-structured answers, which has positioned them as a primary gateway to information for many users. Yet, the texts they produce differ from those written by humans. Such distinctive characteristics may shape how readers allocate visual attention and interpret the information they receive. This study investigates how individual cognitive propensities influence visual attention and perceived cognitive load when reading AI-generated answers. We focus on two specific propensities: goal-directed visual search, the ability to efficiently locate relevant information amid distracting elements, and visual discrimination, the ability to differentiate visually similar details. Although these propensities did not significantly predict visual-attention measures, goal-directed visual search was associated with lower perceived cognitive load, whereas visual discrimination was linked to higher perceived load. Reading habits also showed marginal trends suggesting a potential moderating role, softening the influence of visual propensities on cognitive load. Overall, this work deepens our understanding of how individuals process information when reading AI-generated text and highlights the importance of designing adaptive interfaces that accommodate diverse readers and cognitive profiles.

Key words: AI-generated text, visual attention, cognitive load, eye-tracking, visual search, visual discrimination, reading habits, individual differences.

Research methods: Experimental study, eye-tracking, psychometric measurement, within-subjects design, quantitative analysis.

Table des matières

Résumé.....	iii
Abstract.....	iv
Liste des tableaux et des figures	vii
Liste des abréviations	viii
Avant-propos	ix
Déclaration concernant l'utilisation de l'IA.....	x
Chapitre 1	1
1. Introduction.....	1
1.1 Mise en contexte.....	1
1.2 Objectifs et question de recherche	3
1.3 Contributions	4
1.4 Structure du mémoire	5
1.5 Contribution de l'étudiant et responsabilités.....	5
Chapitre 2	8
1. Introduction.....	8
2. Literature review	11
2.1 Foundations of Reading and Comprehension	11
2.2 Cognitive Load Theory and Information Design	15
2.3 Individual Differences in Cognitive Processing.....	16
2.4 Summary	21
3. Methods.....	22
3.1 Procedure.....	22
3.2 Materials and Measures.....	23
3.3 Participants	29
3.4 Data Extraction and Analysis	29
4. Results.....	30
4.1 Visual Attention Focus	30
4.2 Cognitive Load	31

4.3 Moderating Effects of Reading Habits	31
5. Discussion	33
5.1 Summary of Key Findings	33
5.2 Interpretation by Hypothesis Family	33
6. Conclusion	38
6.1 Summary	38
6.2 Theoretical Contributions.....	38
6.3 Design Implications.....	39
6.4 Limitations and Research Avenues.....	40
7. Bibliographie	41
Chapitre 3	47
1. Conclusion	47
1.1 Synthèse des résultats.....	47
1.2 Contributions théoriques	49
1.3 Implications pratiques	49
1.4 Limites.....	50
1.5. Pistes de recherches futures.....	51

Liste des tableaux et des figures

Chapitre 1

Tableau 1.1 Tableau de contribution

Chapitre 2

Figure 2.1 Research Model

Figure 3.1 Area of Interest (AOI) Defined on the Result Page

Figure 3.2 Overview of the Large Grid Paradigm

Figure 3.3 Overview of the Visual Symbol Search Task

Table 3.4 Operationalization of Constructs, Measures, and Data Collection Instruments

Table 4.1 Validity Test Results

Liste des abréviations

(FR)

IA – Intelligence artificielle

CER – Comité d'éthique de la recherche

M. Sc. – Maîtrise ès sciences

(EN)

ACM – Association for Computing Machinery

AI – Artificial Intelligence

AOI – Area of Interest

CES – Cognitive Effort Scale

CLT – Cognitive Load Theory

LLM – Large Language Model

NASA-TLX – NASA Task Load Index

RAG – Retrieval-Augmented Generation

RHS – Reading Habits Scale

SD – Standard Deviation

VSS – Visual Symbol Search

UX – User Experience

Avant-propos

Ce travail a été effectué dans le cadre d'une maîtrise en sciences de la gestion (M. Sc.) avec spécialisation en expérience utilisateur (UX) à HEC Montréal. Le sujet de mémoire a été enregistré et approuvé par l'administration du programme, et une autorisation a été reçue pour rédiger ce mémoire par article. L'inclusion de l'article dans ce mémoire a été approuvée par tous les co-auteurs.

L'article principal (chapitre 2) a été rédigé en vue d'une soumission à la revue *Behaviour & Information Technology* (BIT). Au moment de la remise de ce mémoire, l'article n'a pas encore été soumis à la revue.

La recherche présentée dans ce mémoire a été approuvée par le Comité d'éthique de la recherche (CER) de HEC Montréal en avril 2025 sous le numéro de certificat 2025-6409. La collecte de données s'est déroulée entre mi-mai et mi-juin 2025 au laboratoire Tech3Lab de HEC Montréal. Conformément aux exigences éthiques, le projet a été mené dans le respect des normes de recherche avec participants humains, incluant l'obtention du consentement éclairé, la confidentialité des données et le droit de retrait.

Ce projet s'inscrit dans une collaboration entre le Tech3Lab et un partenaire média national, dans le cadre d'un projet plus large visant à examiner l'expérience utilisateur dans des contextes de recherche d'information médiée par l'intelligence artificielle. La contribution spécifique de l'étudiante au sein de ce projet est détaillée au chapitre 1 (section 1.5).

Déclaration concernant l'utilisation de l'IA

Dans le cadre de la préparation de ce travail, les auteurs ont utilisé des outils d'intelligence artificielle générative (ChatGPT et Claude) pour des tâches de soutien rédactionnel, notamment la révision linguistique, la reformulation stylistique et l'amélioration de la clarté du texte. Aucun contenu conceptuel, analytique ou empirique n'a été généré automatiquement : toutes les idées, interprétations, analyses, décisions méthodologiques et conclusions proviennent entièrement des auteurs. Après l'utilisation de ces outils, les auteurs ont procédé à une vérification systématique, ont corrigé ou modifié le texte au besoin, et assument l'entière responsabilité du contenu final.

Chapitre 1

1. Introduction

1.1 Mise en contexte

Les modèles de langage ont profondément transformé notre façon de chercher de l'information (Ai *et al.*, 2023). Ils produisent des réponses rapides, structurées et cohérentes, devenant ainsi des outils incontournables dans l'accès à l'information (Ai *et al.*, 2023). Toutefois, les textes qu'ils génèrent demeurent fondamentalement différents de ceux produits par des humains : ils résultent d'un processus statistique plutôt que d'une intention communicative réelle (Chen, Li and Ye, 2024). Produites sans planification hiérarchique ni organisation structurée du discours (Kim *et al.*, 2025), ces réponses peuvent également reproduire des biais présents dans les données d'entraînement, manquer de traçabilité quant à l'origine des informations fournies et contenir des erreurs factuelles ou des hallucinations, souvent sans signaux explicites d'incertitude (Kim *et al.*, 2024; Lin *et al.*, 2024). De telles particularités sont susceptibles de modifier la manière dont les lecteurs mobilisent leur attention visuelle, naviguent dans ces contenus et interprètent l'information présentée.

Concrètement, ces situations se manifestent lorsque des utilisateurs consultent des réponses générées par l'IA dans des contextes de recherche d'information, par exemple pour s'informer sur un enjeu d'actualité, comprendre une décision politique ou obtenir une synthèse explicative sur un sujet complexe. L'utilisateur est alors confronté à une réponse textuelle produite automatiquement, souvent présentée sous la forme d'un paragraphe dense et linéaire, sans hiérarchisation explicite de l'information ni intention éditoriale humaine. Il doit décider quels éléments lire attentivement, lesquels ignorer, et dans quelle mesure il peut s'appuyer sur le contenu présenté, ce qui sollicite à la fois son attention visuelle et ses ressources cognitives.

Dans ce mémoire, l'intelligence artificielle étudiée correspond à des modèles de langage (LLM) générant des réponses textuelles uniques en réponse à une requête, dans un contexte de lecture non conversationnel. Plus précisément, les stimuli utilisés s'apparentent à des sorties de systèmes de génération de texte augmenté par récupération (Retrieval-Augmented Generation), où le lecteur consulte une réponse générée sans possibilité d'interaction ou de reformulation. Ce contexte se distingue des usages dialogiques de type *chatbot* et se rapproche davantage d'une situation de lecture informationnelle, comparable à celle d'un article ou d'un encadré explicatif produit automatiquement.

Des études en oculométrie confirment que les lecteurs traitent les textes générés par IA différemment des textes humains : leurs fixations sont plus courtes et leurs mouvements oculaires plus rapides (Ilyas *et al.*, 2025). Ces textes ne sont pas nécessairement plus difficiles à lire, mais ils sollicitent des processus de lecture différents. En effet, leur cohérence ne découle pas d'une organisation délibérée du propos par un auteur; elle repose plutôt sur des schémas rédactionnels génériques, c'est-à-dire des formules récurrentes et des structures prévisibles, plutôt que sur des marqueurs discursifs explicites. Le lecteur s'appuie ainsi sur la prévisibilité du format plutôt que sur des indices textuels qui guideraient sa progression (Herbold *et al.*, 2023).

Or, la compréhension d'un texte ne dépend pas uniquement de son contenu, mais aussi des processus cognitifs mobilisés par le lecteur pour le traiter (Follmer and Sperling, 2018). Deux dimensions sont particulièrement pertinentes pour étudier cette dynamique. L'attention visuelle oriente la sélection de l'information — ce que le lecteur regarde, dans quel ordre et pendant combien de temps (Andreou *et al.*, 2024). La charge cognitive, quant à elle, reflète l'effort mental requis pour maintenir et intégrer l'information en mémoire de travail (Sweller, 1988). Face à des textes générés par IA, ces deux dimensions permettent d'observer comment le lecteur s'adapte à un contenu structuré différemment.

Mais pourquoi certains lecteurs naviguent-ils plus efficacement dans les textes que d'autres? Pourquoi certains ressentent-ils moins d'effort cognitif face au même contenu? Des différences dans les propensions cognitives individuelles pourraient expliquer ces

variations (McVay and Kane, 2012). Parmi celles-ci, deux capacités perceptuelles semblent particulièrement pertinentes dans le contexte de la lecture de textes générés par l'IA. La recherche visuelle orientée vers un objectif renvoie à la capacité de repérer rapidement une information pertinente au sein d'un ensemble visuel contenant des distracteurs (Wendt, Brand and Kollmeier, 2014). La discrimination visuelle correspond quant à elle à l'habileté à distinguer des éléments visuellement similaires et à comparer des détails subtils (Wendt, Brand and Kollmeier, 2014). Ces propensions pourraient jouer un rôle dans la façon dont les individus traitent visuellement et cognitivement les textes générés par IA.

Toutefois, ces propensions cognitives ne s'exercent pas en isolation. L'expérience individuelle de lecture pourrait également entrer en jeu. Les lecteurs qui lisent fréquemment, qui éprouvent du plaisir à lire et qui se sentent compétents pour comprendre différents types de textes ont développé des automatismes susceptibles de faciliter le traitement de l'information (Vaessen and Blomert, 2010). Par extension, ces différences pourraient également influencer la façon dont les utilisateurs traitent des contenus générés par IA (Ronconi, Altoè and Mason, 2025).

Dans un environnement où les réponses générées sont structurées mais dépourvues d'intention communicative humaine (Chen, Li and Ye, 2024), la compréhension repose en grande partie sur ce que le lecteur apporte lui-même. Examiner le rôle des propensions cognitives et de l'impact des habitudes de lecture dans le traitement de textes générés par IA devient essentiel pour concevoir des interfaces d'information adaptées à la diversité des lecteurs.

1.2 Objectifs et question de recherche

L'objectif de cette recherche est de mieux comprendre comment certaines propensions cognitives influencent la manière dont les individus traitent des textes générés par l'intelligence artificielle. Plus précisément, cette étude examine le rôle de deux capacités perceptives — la recherche visuelle orientée vers un objectif et la discrimination visuelle

— et analyse comment ces différences individuelles influencent l’attention visuelle et la charge cognitive lors de la lecture de réponses textuelles générées par IA. L’étude considère également les habitudes de lecture comme un facteur susceptible de moduler ces relations.

La question de recherche qui guide ce mémoire est la suivante :

Dans quelle mesure les propensions cognitives des lecteurs — la recherche visuelle orientée vers un objectif et la discrimination visuelle — influencent-elles l’attention visuelle et la charge cognitive perçue lors de la lecture de textes générés par l’intelligence artificielle?

1.3 Contributions

Ce mémoire apporte un éclairage empirique sur une question encore peu explorée : comment les différences individuelles influencent le traitement cognitif de textes générés par IA. En examinant conjointement des propensions cognitives (recherche visuelle et discrimination visuelle), des mesures comportementales (attention visuelle par oculométrie) et des indicateurs subjectifs (charge mentale et effort cognitif perçus), cette recherche documente des variations interindividuelles qui restaient jusqu’ici peu caractérisées dans la littérature sur la lecture de contenus algorithmiques.

Sur le plan méthodologique, ce mémoire propose un protocole, utile à d’autres chercheurs, intégrant des paradigmes standardisés de psychologie cognitive (tâches de recherche et de discrimination visuelles) comme mesures prédictives dans un contexte de lecture écologique. Cette articulation entre mesures expérimentales contrôlées et mesures en situation de lecture réelle offre une approche reproductible pour examiner le rôle de capacités perceptives spécifiques dans le traitement de textes générés par l’intelligence artificielle.

Enfin, d'un point de vue pratique, cette recherche a permis la formulation de recommandations pour la conception d'interfaces génératives. En documentant comment certaines propensions cognitives et habitudes de lecture influencent l'expérience utilisateur, les résultats suggèrent que les systèmes de génération de texte gagneraient à proposer des options d'affichage ou de structuration adaptables au profil du lecteur. Ces pistes ouvrent la voie à des interfaces plus inclusives, tenant compte de la diversité cognitive des utilisateurs.

1.4 Structure du mémoire

Le chapitre 1 introduit la problématique générale, présente le contexte théorique et empirique, précise les objectifs de recherche, les contributions et responsabilités de l'étudiant.

Le chapitre 2 regroupe l'article scientifique principal, rédigé en vue d'une soumission à la revue *Behaviour & Information Technology* (BIT). Cet article présente une étude expérimentale menée auprès de 27 participants ayant lu sept réponses générées par un système d'IA, tandis que leurs mouvements oculaires et leur charge cognitive perçue étaient enregistrés. Les résultats indiquent que, si les propensions visuelles ne prédisent pas significativement les mesures d'attention visuelle, la recherche visuelle orientée vers un objectif est associée à une charge cognitive perçue plus faible, tandis que la discrimination visuelle est liée à une charge cognitive plus élevée. Les habitudes de lecture semblent modérer marginalement ces relations.

Enfin, le chapitre 3 propose une conclusion générale du mémoire. Il synthétise les principaux résultats, discute des contributions théoriques et pratiques, met en lumière les limites de l'étude et ouvre des pistes pour de futures recherches.

1.5 Contribution de l'étudiant et responsabilités

Le tableau suivant présente les principales étapes du processus de recherche ainsi que la contribution associée à chacune d’elles.

Tableau 1.1 Tableau de contribution

Étapes du processus	Contribution
Question de recherche	Identification des lacunes dans la littérature et définition du problème de recherche [80%] – Problématique initiale proposée par le partenaire – Contexte affiné par l’équipe du laboratoire Approfondissement théorique et reformulation du problème effectués par l’étudiante
Revue de littérature	Recherche et synthèse approfondie d’articles scientifiques pour établir l’état actuel des connaissances [100%] - Collecte, analyse critique et intégration des sources effectuées entièrement par l’étudiante
Design expérimental	Soumission de l’étude au Comité d’éthique de la recherche (CER) de HEC Montréal [70%] – Rédaction et préparation du document d’approbation en collaboration avec l’équipe du laboratoire Développement du protocole expérimental et du stimulus [80%] – Conception du protocole par l’étudiant à partir d’un modèle du laboratoire – Adaptation des consignes et création des stimuli en collaboration avec le partenaire
Collecte de données	Recrutement des participants pour la collecte [50%] – Recrutement effectué en partie par l’étudiante, complété par le panel du Tech3Lab Configuration du laboratoire [50%] – Programmation du timeline pour l’oculométrie – Conception du questionnaire sur Qualtrics

	– Montage et calibration des instruments de collecte avec le soutien du personnel du Tech3Lab Prétests et supervision des opérations de collecte de données [100%] – Réalisation de tests préliminaires – Optimisation de la procédure et supervision des sessions utilisateurs
Analyse statistique	Analyse des données psychophysiologiques et comportementales [70%] – Extraction et synchronisation des données effectuées avec l'aide du statisticien du Tech3Lab – Analyses statistiques réalisées par l'étudiante dans SAS à partir d'un modèle fourni par le laboratoire – Interprétation et présentation des résultats rédigées par l'étudiante
Rédaction	Rédaction du mémoire et de l'article [95%] – Rédaction complète par l'étudiante – Révisions, commentaires et ajustements apportés par les codirecteurs du laboratoire

Chapitre 2

Cognitive Diversity in AI-Generated Text Reading: The Roles of Visual Propensities and Reading Habits

Large Language Models (LLMs) have fundamentally reshaped how people search for information. They generate quick and well-structured answers, which has positioned them as a primary gateway to information for many users. Yet, the texts they produce differ from those written by humans. Such distinctive characteristics may shape how readers allocate visual attention and interpret the information they receive. This study investigates how individual cognitive propensities influence visual attention and perceived cognitive load when reading AI-generated answers. We focus on two specific propensities: goal-directed visual search, the ability to efficiently locate relevant information amid distracting elements, and visual discrimination, the ability to differentiate visually similar details. Although these propensities did not significantly predict visual-attention measures, goal-directed visual search was associated with lower perceived cognitive load, whereas visual discrimination was linked to higher perceived load. Reading habits also showed marginal trends suggesting a potential moderating role, softening the influence of visual propensities on cognitive load. Overall, this work deepens our understanding of how individuals process information when reading AI-generated text and highlights the importance of designing adaptive interfaces that accommodate diverse readers and cognitive profiles.

Key words: AI-generated text, visual attention, cognitive load, eye-tracking, visual search, visual discrimination, reading habits, individual differences.

1. Introduction

By August 2025, generative AI adoption had risen sharply: 54.6% of U.S. adults aged 18-64 reported using generative AI for work or non-work purposes, up from 44.6% the previous year (Bick, Blandin and Deming, 2025). Large language models are increasingly

embedded in search, productivity, and learning tools, functioning as primary information intermediaries (Annapureddy, Fornaroli and Gatica-Perez, 2025). Yet these systems produce text that differs fundamentally from human-authored content: responses arise from statistical patterning rather than communicative intent (Chen, Li and Ye, 2024), are not guided by deliberate discourse planning (Kim *et al.*, 2025), and may present inaccuracies without uncertainty cues (Kim *et al.*, 2024). These characteristics introduce several limitations, including the reproduction of training-data biases, limited source traceability, and a heightened risk of factual errors or hallucinations (Lin *et al.*, 2024; Kim *et al.*, 2025).

Information search already requires evaluation, credibility judgment, and integration, processes that impose cognitive load and intensify with task complexity (Niwanputri, Tob-Ogu and Hosseinzadeh, 2025). Early evidence shows that LLM-generated answers do not consistently ease this burden. Their structural and epistemic properties can fragment attention, hinder the integration of information, and ultimately increase mental effort rather than support comprehension (Ai *et al.*, 2023). This raises a central question: how do users process and make sense of AI-generated information under these constraints?

This variability in user experience points to an overlooked dimension: individual differences in visual processing. Comprehension does not depend solely on content itself, but also on the cognitive resources readers mobilize and adjust based on that content (Follmer and Sperling, 2018). Visual attention determines what readers look at, in what order, and for how long (Andreou *et al.*, 2024), while cognitive load reflects the mental effort required to process and understand information (Sweller, 1988). When reading AI-generated text, users rely on these resources. Yet current research offers little insight into how stable cognitive propensities shape this attentional deployment and cognitive effort.

Two specific individual propensities may be particularly relevant. Goal-oriented visual search, the ability to rapidly locate target information amid distractors, determines how efficiently users can identify relevant elements without exhaustive scanning (Wendt, Brand and Kollmeier, 2014). Visual discrimination, the capacity to distinguish visually similar elements and compare subtle details, affects how users compare and verify information

(Wendt, Brand and Kollmeier, 2014). These propensities may directly influence both where attention is allocated and how much cognitive effort is required when processing AI-generated responses.

Reading habits may further modulate these effects. Readers who read frequently, derive pleasure from reading, and feel confident in their comprehension have developed automaticity that can facilitate information processing (Vaessen and Blomert, 2010). By extension, these differences could also influence how users process AI-generated content (Ronconi, Altoè and Mason, 2025). Yet, whether reading habits moderate the relationship between visual propensities and cognitive load in AI-mediated contexts remains empirically unexamined.

This gap motivates the present study. We ask: *To what extent do individual propensities for goal-oriented visual search and visual discrimination influence visual attention and cognitive load when reading LLM-generated text, and do reading habits moderate these relationships?*

We conducted a controlled laboratory experiment with 27 participants who read seven AI-generated responses while their eye movements and subjective cognitive load were recorded. Gaze was analyzed within a predefined Area of Interest (AOI) covering the AI-generated answer, and post-task questionnaires captured perceived mental demand and cognitive effort. In the present study, the AI-generated answers were designed to simulate the type of output produced by a retrieval-augmented generation (RAG) system. RAG models retrieve relevant external information and incorporate it into the generation process to produce responses that blend retrieved evidence with algorithmic synthesis (Cui *et al.*, 2025). This hybrid structure reflects the kinds of AI-mediated reading environments users increasingly encounter, and thus provides an ecologically meaningful context for examining how cognitive propensities shape attention and effort during AI-assisted information processing.

We make three contributions. First, we show that goal-oriented visual search is associated with lower subjective cognitive load, specifically lower mental demand and self-reported effort, while visual discrimination predicts higher mental demand when reading LLM-generated text. These findings demonstrate that stable visual propensities shape the subjective experience of cognitive load during AI-mediated reading, a dimension largely overlooked in prior research.

Second, we find that reading habits marginally moderate these effects, with stronger habits buffering the association between visual propensities and cognitive load. This extends Cognitive Load Theory by showing that habitual reading can provide compensatory regulation when processing algorithmically structured text (Stanovich, 1986; Sweller, 1988).

Third, we link these findings to interface design, demonstrating that cognitive load depends on the match between user propensities and output structure. This supports adaptive presentation strategies, such as clearer segmentation for users with weaker visual search, or reduced density for those with high discrimination sensitivity, to make LLM responses more cognitively accessible across diverse users.

The article proceeds as follows. We synthesize research on visual attention, cognitive load, and individual differences in reading, then derive hypotheses. We describe the experimental setup, instruments, and analytic approach, followed by results on visual behavior, subjective load, and moderation. We close with theoretical implications and design recommendations.

2. Literature review

2.1 Foundations of Reading and Comprehension

This literature review is structured around the conceptual framework guiding the present study. The framework proposes that individual visual propensities influence how readers

process AI-generated text, both in terms of visual attention and perceived cognitive load. Reading habits are further considered as a factor that may shape or attenuate these relationships.

To support this framework, the review first examines foundational theories of reading and comprehension, with particular attention to the cognitive processes involved in allocating attention and integrating information during text reading. It then introduces Cognitive Load Theory to explain how characteristics of information presentation and task demands shape mental effort during reading. Finally, the review focuses on individual differences in cognitive and perceptual processing, emphasizing visual search, visual discrimination, and reading habits as key sources of interindividual variability relevant to AI-mediated reading contexts.

Together, these perspectives provide the theoretical grounding for the hypotheses developed in this study and clarify how each construct contributes to the overall research logic.

2.1.1 Cognitive Processes

Reading is a cognitively demanding activity that depends on the continuous coordination of attention, memory, and language comprehension (Rayner, 1998). As they move through a text, readers must focus on relevant information, retain what they have just processed, and integrate new material into an evolving mental representation, all while managing limited cognitive resources (Niwanputri, Tob-Ogu and Hosseinzadeh, 2025).

Comprehension emerges from the interplay between lower-level linguistic processes and higher-level reasoning. On the linguistic side, readers identify words, parse syntax, and link phrases into propositions. On the conceptual side, they infer intentions, build situation models, and connect ideas into coherent meaning (Kintsch, 1998; Rayner, 1998). When lexical access is slow or sentence structure is complex, more effort is required at the linguistic level, leaving fewer resources available for higher-order integration (Baker and Love, 2023).

One factor that strongly shapes how efficiently this process unfolds is text readability. Readability refers to the ease with which a text can be processed and understood based on its linguistic characteristics (Zhang and Lu, 2025). It has both a clarity dimension, concerning how transparently ideas are expressed, and a processing dimension, concerning how efficiently words and sentences can be decoded and linked together (Skalicky and Dascalu, 2019). A clear and well-structured text reduces unnecessary interpretive effort, allowing readers to devote more attention to meaning rather than to resolving ambiguity (Dubay, 2004).

Beyond the properties of the text itself, comprehension is also constrained by the reader's cognitive resources. Working memory plays a central role by supporting the temporary storage and integration of information across words and sentences (Kennison and Clifton, 1995; Krejtz *et al.*, 2018). Readers with greater working memory capacity are better able to maintain partial interpretations, track references, and connect new input with previously read material, which directly contributes to coherent understanding (Just and Carpenter, 1980).

Prior knowledge and reading fluency further shape how efficiently readers construct meaning. Familiarity with a topic allows new information to be mapped onto existing mental structures, reducing uncertainty and effort (Kintsch, 1998; McNamara *et al.*, 2002). Over time, repeated exposure fosters automaticity in lexical and syntactic processing, making word recognition and parsing faster and less demanding (LaBerge and Samuels, 1974). As these lower-level processes become more automatic, readers can allocate a greater share of their attention to higher-level reasoning—inferring, evaluating, and integrating ideas into a broader understanding (Acheson, Wells and MacDonald, 2008).

In essence, reading efficiency reflects a balance between task demands—the complexity and organization of the text—and the resources available to the reader, such as working memory, prior knowledge, and fluency (Hautala *et al.*, 2022). Skilled readers who are both practiced and knowledgeable process text more efficiently and with less mental effort because key aspects of comprehension become more automatic (Smith *et al.*, 2021). This leaves more capacity for deeper comprehension, such as evaluating claims or integrating

information across sources. These dynamics become especially salient when reading AI-generated content, where informational hierarchy is often implicit rather than explicitly signaled, requiring readers to mentally reconstruct the organization and determine relevance through visual scanning rather than relying on authorial guidance (Joty, Carenini and Ng, 2015; Kim *et al.*, 2025).

This intricate balance between information structure, cognitive resources, and processing effort lies at the heart of Cognitive Load Theory (Sweller, 1988; Mayer and Moreno, 2003). CLT predicts that learning and comprehension are hindered when the demands imposed by a material exceed the learner's limited working memory capacity. It also distinguishes between intrinsic load (task complexity), extraneous load (presentation and structure), and germane load (effort invested in making sense of the content) (Sweller, 1988). These predictions provide a foundation for understanding how the design and presentation of AI-generated text can either support or hinder comprehension.

2.1.2 Eye-Movement Behavior in Reading

Eye-tracking provides a non-invasive and ecologically valid method for studying how readers process written information in real time (Hyönä, 2010; Mézière *et al.*, 2023). By recording fixations (pauses where the eyes remain still) and saccades (rapid movements between fixations), this method offers fine-grained insight into attention allocation and processing effort during reading (Frazier and Rayner, 1982; Rayner, 2009). Unlike post-task measures, eye-tracking captures cognitive processing as it unfolds, making it an effective tool for examining online comprehension and mental workload (Hyönä, 2010; Krejtz *et al.*, 2018).

Eye-movement behavior during reading is guided by two key principles. The Eye–Mind Hypothesis (Just and Carpenter, 1980) posits that readers process the information they are currently fixating, linking gaze position with attention. The Immediacy Assumption (Just and Carpenter, 1980) proposes that comprehension occurs almost simultaneously with reading, primarily during fixations. Based on these principles, gaze metrics such as fixation duration, saccade length, and regressions serve as indicators of processing effort. Longer

fixations generally reflect increased processing difficulty, often due to lexical or syntactic complexity, while regressions (backward movements) typically indicate comprehension problems or the need for reanalysis (Rayner and Duffy, 1986; Rayner, 2009; Hyönä, 2010).

Eye-tracking measures are typically analyzed at different levels. Global measures (e.g., average fixation duration, total reading time on AOIs) capture overall reading efficiency and workload distribution across a text. Local measures focus on specific words or phrases to examine moment-to-moment processing difficulty (Mézière *et al.*, 2023). In this study, the analysis focuses on global and aggregated local measures, as participants engaged in continuous text reading rather than isolated word processing.

Taken together, eye-movement data provide a dynamic picture of how readers allocate attention and manage effort as they interact with text. Patterns of fixations, saccades, and regressions reveal where readers focus their attention and how much mental effort they invest in understanding. These observable fluctuations in gaze behavior offer a valuable window into cognitive load, a concept further explored in the next section through the lens of Cognitive Load Theory (Sweller, 1988; Mayer and Moreno, 2003).

2.2 Cognitive Load Theory and Information Design

2.2.1 Principles of Cognitive Load Theory

Cognitive Load Theory (CLT) provides a structured framework to explain how the presentation and organization of information affect comprehension efficiency (Mayer and Moreno, 2003; Kuperman, Matsuki and Van Dyke, 2018). According to this theory, the mental effort required during reading depends not only on the inherent complexity of the material but also on how clearly the content is presented (Sweller, 1988; Mayer and Moreno, 2003).

Cognitive Load Theory (CLT) distinguishes three forms of cognitive load. Intrinsic load refers to the inherent complexity of the material, how many elements must be understood and how interdependent they are (Sweller, 1988; Paas, Renkl and Sweller, 2003). Extraneous load arises from the way information is presented, such as unclear layouts,

unnecessary details, or redundant visual scanning caused by poor design (Paas, Renkl and Sweller, 2003). Finally, germane load represents the useful mental effort invested in organizing and integrating new information into long-term memory (Paas, Renkl and Sweller, 2003).

In CLT, not all effort is useful. Intrinsic effort is unavoidable because it comes from the material's inherent complexity, while germane effort is desirable because it supports constructing and refining meaning in memory. What can be reduced is the effort that comes from avoidable presentation issues. Clear and coherent structure, alignment of related information, and the removal of unnecessary redundancy help limit this avoidable load and free cognitive resources for actual comprehension (Sweller, 1988; Mayer and Moreno, 2003).

In digital and AI-mediated reading environments, these principles become even more relevant. Generated text often varies in organization, clarity, and reliability, requiring readers to search for relevance, verify claims, and manage uncertainty. Poor structure or verbosity increases extraneous load, while clear segmentation and explicit cues help conserve cognitive resources for comprehension and evaluation (Hyönä, 2010; Niwanputri, Tob-Ogu and Hosseinzadeh, 2025). From a cognitive standpoint, these effects underscore that design and presentation shape how information is processed and how much effort that processing demands.

The mental effort imposed by a given text is not experienced uniformly across readers. Some locate and prioritize relevant information efficiently; others spend more effort searching or resolving ambiguity. In that sense, cognitive load reflects not only properties of the text but also the reader's attentional control. These stable attentional propensities motivate the next section.

2.3 Individual Differences in Cognitive Processing

Readers do not all engage with text in the same way. Interindividual variability in reading reflects stable perceptual and attentional traits that shape how people locate relevant information, filter distraction, and manage cognitive effort (Rayner, 1998; Mézière *et al.*, 2023) . In other words, some readers can immediately focus on what matters, while others invest more effort simply navigating the text.

These differences are not random. They arise from enduring cognitive characteristics that guide visual search, selective attention, and moment-to-moment control of processing (Proulx, 2011; Clarke *et al.*, 2022). In demanding settings, such as AI-generated content, these traits influence how difficult the reading experience feels, and how much cognitive load it imposes.

Two visual propensities are especially relevant in this context: goal-oriented visual search and visual discrimination. Both have been widely studied as markers of attentional efficiency, and they rely on complementary mechanisms (Belardinelli, Herbort and Butz, 2015; Wolfe and Horowitz, 2017). Goal-oriented search supports rapid identification of task-relevant information. Visual discrimination reflects sensitivity to subtle differences and inconsistencies. Together, they capture two sides of efficient information processing: directing attention to what is important and evaluating the precision of what is seen.

2.3.1 Goal-Oriented Visual Search Propensity

Goal-oriented visual search refers to the ability to direct attention toward task-relevant stimuli. In classical visual search paradigms, efficient search depends on selective guidance mechanisms that prioritize likely targets (Treisman and Gelade, 1980; Wolfe, 1994). Neurocognitively, this top-down attentional control relies on prefrontal and parietal networks that coordinate attention according to task goals (Corbetta and Shulman, 2002).

Eye-tracking paradigms show that gaze allocation reflects task-relevant interpretation processes, with fixations shifting toward the target as soon as the linguistic information allows disambiguation (Wendt, Brand and Kollmeier, 2014; Belardinelli, Herbort and Butz,

2015). Individuals with stronger goal-oriented search ability can deploy attention more strategically, allowing faster detection of relevant information and reduced scanning effort.

Applied to reading, this propensity translates into more focused processing of meaningful textual segments and fewer unnecessary fixations on peripheral or redundant content. In AI-mediated information environments, efficiently filtering information reduces unnecessary navigation demands and frees cognitive resources for comprehension. AI-driven systems have been shown to lower extraneous cognitive load by simplifying information presentation and aligning content with user needs (Twabu, 2025).

Building on this attentional efficiency, we formulate the following hypothesis concerning visual attention:

H1. A higher goal-oriented visual search propensity is associated with more focused visual attention.

This efficiency in filtering and prioritizing information also informs our prediction regarding cognitive load:

H3. A higher goal-oriented visual search propensity is associated with lower cognitive load.

2.3.2 Goal-Oriented Visual Discrimination Propensity

Visual discrimination refers to the perceptual ability to detect and differentiate subtle variations among visual stimuli—such as small changes in form, texture, or spatial arrangement. This capacity reflects fine-grained perceptual sensitivity and supports detailed analysis of visual input. In reading contexts, individuals with high visual discrimination tend to examine text with greater precision, attending closely to lexical, syntactic, or formatting cues that others might overlook (Aron and Aron, 1997; Rayner, 1998).

This enhanced perceptual sensitivity can be advantageous when accuracy and error detection are required, for example when identifying contradictions or ambiguous phrasing (Hautus, Macmillan and Creelman, 2021). However, it may also lead to over-engagement with local details, increasing processing effort when the text is complex, ambiguous, or inconsistently structured. Cognitive research shows that discrimination sensitivity is closely linked to metacognitive evaluation—the ability to assess the reliability of one’s perceptual judgments—which recruits executive and working-memory resources in prefrontal networks (Maniscalco and Lau, 2015).

In the context of AI-generated outputs, individuals with stronger discrimination propensity may experience higher cognitive load. Their heightened sensitivity compels them to analyze and reconcile fine-grained discrepancies or inconsistencies within long or loosely structured responses. Eye-tracking research demonstrates that such detailed perceptual analysis increases fixation rates, re-inspection, and processing time when dealing with ambiguous or structurally complex material (Wendt, Brand and Kollmeier, 2014), thereby elevating perceived mental demand (Proulx, 2011).

Given this tendency to engage more extensively with local details, we derive the following hypothesis about visual attention:

H2. A higher visual discrimination propensity is associated with more distributed (less focused) visual attention.

The cognitive implications of this detail-oriented processing lead to the following prediction regarding cognitive load:

H4. A higher visual discrimination propensity is associated with higher cognitive load.

2.3.3 Reading Habits as a Moderating Factor

Visual propensities do not operate in isolation. Experience with reading also shapes how cognitive load is managed. Frequent readers develop greater automaticity in lower-level processes like word recognition and syntactic parsing, which frees cognitive resources for higher-level interpretation (LaBerge and Samuels, 1974). They also tend to engage in metacognitive regulation: monitoring understanding, slowing down when something is unclear, skimming when content is redundant, and rereading when needed (Guthrie and Wigfield, 2000). These strategies act as a form of cognitive resilience.

In AI-mediated reading, such experience may buffer effort in two ways. First, readers with strong self-regulation skills can compensate for weaker goal-oriented search by deploying deliberate strategies to locate relevant information. Second, when they possess high visual discrimination, experienced readers can avoid over-engaging with every local inconsistency by shifting flexibly between detailed inspection and global meaning, a form of strategic attentional control documented in skilled reading research (Rayner, 1998; McNamara *et al.*, 2002). Readers with less developed reading habits may be more vulnerable to increases in cognitive effort when processing texts. Because they rely more heavily on working-memory resources and have fewer automated strategies, they experience greater difficulty when text organization is insufficient to support comprehension (Peng *et al.*, 2018).

Because reading habits provide compensatory strategies and regulatory control, we expect them to moderate the role of visual search in shaping cognitive experience:

H5. The effect of goal-oriented visual search propensity on cognitive load is attenuated among users with stronger reading habits.

For similar reasons, we also anticipate a moderating effect on the relationship between visual discrimination and cognitive load:

H6. The effect of visual discrimination propensity on cognitive load is attenuated among users with stronger reading habits.

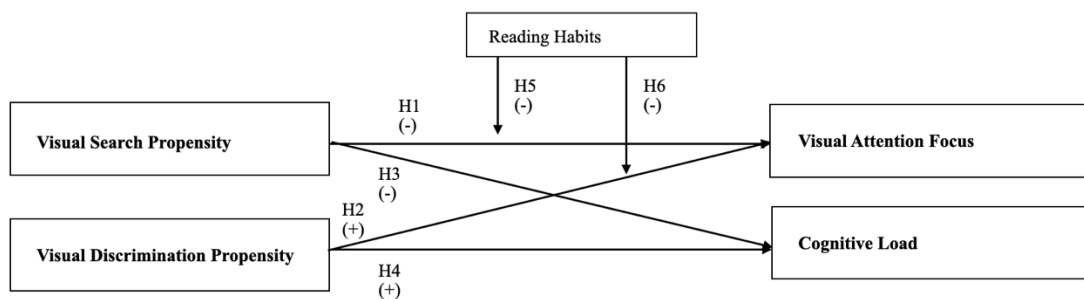
2.4 Summary

Taken together, these factors suggest that cognitive load during reading is shaped not only by the text itself, but also by the reader. Goal-oriented visual search supports efficient selection of relevant information. Visual discrimination supports fine-grained evaluation of what is being read but can increase perceived effort when the material is dense or inconsistent. Reading habits act as a stabilizing influence: they provide strategic control and help keep processing efficient, even under higher demand.

This framework motivates the present study. We test how these individual propensities relate to visual attention and cognitive load when readers interact with AI-generated text, and whether reading habits moderate these relationships.

Building on these theoretical foundations, we propose a conceptual framework that integrates the expected relationships between visual propensities, visual attention, cognitive load, and reading habits. Figure 2.1 presents the research model guiding the hypotheses tested in this study.

Figure 2.1 Research Model



3. Methods

3.1 Procedure

Before the main experimental tasks began, participants were welcomed and provided informed consent. A calibration procedure for the eye-tracker was then performed to ensure accurate gaze recording, followed by the administration of a short pre-experiment questionnaire on Qualtrics (Provo, UT). This questionnaire collected demographic information and participants' self-reported reading habits using the Reading Habits Scale (Acheson, Wells and MacDonald, 2008).

Participants first completed two standardized perceptual tasks assessing distinct visual propensities: a goal-directed visual search task and a visual discrimination task. The order of these tasks was randomized, and both are described in detail later in the Methods.

The main experimental phase then consisted of seven information-seeking tasks, during which participants used a Figma-based mock search interface to formulate a query and read an AI-generated response. For each task, participants were presented with a brief scenario describing an information need and were asked to type a corresponding query into the search bar. The scenarios served only to elicit a plausible query and to create varied reading contexts; they were not experimental manipulations.

Crucially, the system output was fully controlled. Regardless of the exact wording of the query, participants were shown the same pre-generated AI response for each task (i.e., a static “LLM output” screen). This design simulated a realistic search interaction while ensuring that all participants read identical textual stimuli.

Once the response was displayed, participants were instructed to read it as they normally would in an information-seeking context. Eye movements were recorded during reading, and subjective cognitive load was assessed after each task.

The seven tasks and their corresponding topic–response-type pairs were:

T1 – U.S. Politics (AI hallucination): the system produced an incorrect answer to a political query.

T2 – Entertainment (factual response): the system provided a concise and accurate factual answer.

T3 – Canadian Government (factual response): the system delivered a clear, factually correct response to a government-related question.

T4 – Canadian Politics (refusal to answer): the system declined to provide information for a politically sensitive query.

T5 – Consumer Product (no answer): the system returned an empty output.

T6 – Climate Change (short answer): the system generated a very brief and minimally developed response.

T7 – Climate Change (long answer): the system produced an overly long and highly detailed answer to the same climate-related question.

To avoid order effects, participants were assigned to one of three pre-randomized task sequences. None of the sequences began with inadequate responses (T1, T4, T5), and tasks T6 and T7 were never presented consecutively, as they covered the same topic.

3.2 Materials and Measures

Both oculometric (eye-tracking) and self-reported (questionnaire-based) measures were collected to capture complementary indicators of visual attention and cognitive load.

Table 3.4 summarizes the operationalization of the main constructs, associated measures, and data collection instruments.

Table 3.4 Operationalization of Constructs, Measures, and Data Collection Instruments

Construct	Measures	Data collection instrument
-----------	----------	----------------------------

Visual attention focus	• % total fixation duration in AOI (defined as the text block containing the LLM-generated answer) • % total number of fixations in AOI (same text block) • K-coefficient	Tobii Pro Lab (eye-tracking)
Cognitive load	• Perceived mental demand (NASA-TLX) • Perceived cognitive effort (CES)	Qualtrics (scales)
Reading habits	• Reading Habit Scale (RHS)	Qualtrics (scale)

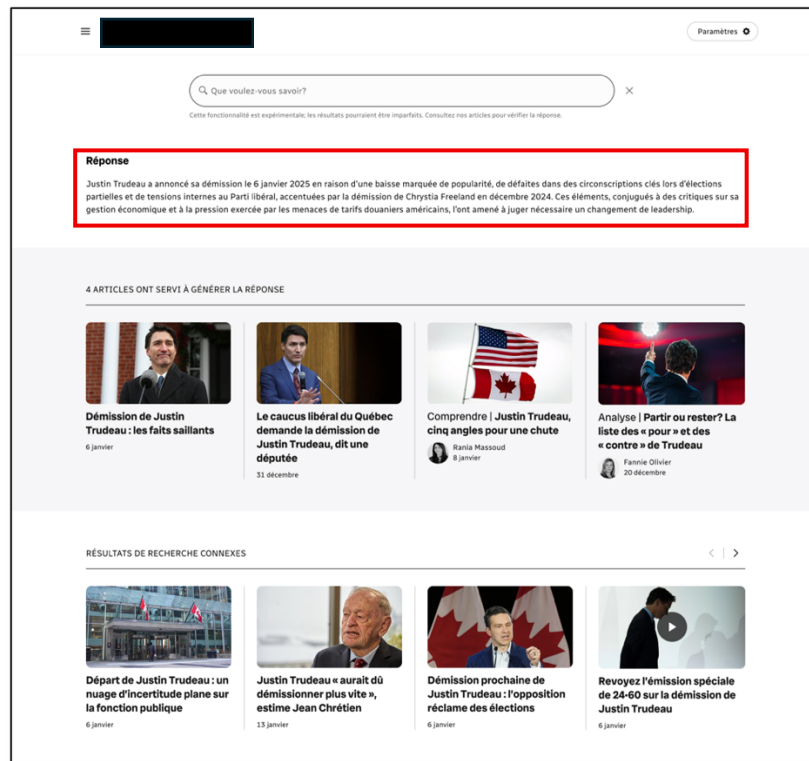
3.2.1 Eye-Tracking Setup and AOIs

Implicit measures were obtained from eye-tracking data. Visual attention was operationalized through two complementary indicators: (1) the proportion of fixations and total fixation duration within predefined AOIs, and (2) the K-coefficient of attentional dispersion. Because the study was conducted in a news information-seeking context, the prototype displayed AI-generated answers to questions about current events. News search typically involves evaluation, verification, and integration across sources, which increases cognitive demands and makes this context appropriate for examining visual attention.

The AOI was defined as the textual response block generated by the Retrieval-Augmented Generation (RAG) system, that is, the paragraph containing the model’s written answer to the user’s query. Supporting article snippets presented alongside the main response were excluded from the AOI to isolate attention directed toward the generated content itself rather than source-checking behaviour. This ensures that the AOI captures reading processes specific to the AI-generated answer.

An illustration of the results page is provided in Figure 3.1 to clarify the interface structure and the placement of the AOI. The AOI delineated by the red bounding box.

Figure 3.1 AOI Defined on the Results Page



The K-coefficient quantifies the spatial dispersion of fixations within the AOI and indicates whether gaze is distributed broadly or concentrated narrowly (Krejtz *et al.*, 2018). In this context, a lower K value reflects more ambient, broadly distributed scanning, whereas a higher K value reflects more focused, narrowly concentrated attention on specific parts of the text.

3.2.2 *Explicit Questionnaires*

Explicit measures included validated scales administered via Qualtrics (Provo, UT, United States). Prior to the experiment, participants reported their reading habits using an adapted version of the Reading Habits Scale (Acheson, Wells and MacDonald, 2008). The retained items were selected for their conceptual relevance and their statistical coherence with the construct of reading competence. Reliability analysis indicated acceptable internal consistency (Cronbach's $\alpha = .73$). The three selected items captured key aspects of habitual reading: frequency, engagement, and comprehension. Participants were asked how

frequently they read different types of documents each week, how much they enjoy reading in general, and how well they understand the content of a text when reading at their normal pace.

After each task, participants completed a brief post-task questionnaire. We used two subjective workload items.

First, perceived cognitive load was measured using the Mental Demand item from the NASA Task Load Index (Hart and Staveland, 1988): "How mentally demanding was the task?". Responses were given on a continuous scale from 0 (very low) to 100 (very high). The NASA-TLX was developed to assess mental workload along six dimensions, including mental demand, effort, and frustration, and has been shown to be sensitive to differences in task difficulty across a wide range of tasks (Rubio *et al.*, 2004). In cognitively driven, screen-based tasks, the Mental Demand item is commonly interpreted as an index of perceived cognitive load.

Second, we measured perceived effort using one item adapted from the Cognitive Effort Scale (Richter, Gendolla and Wright, 2016): "How much effort did you personally have to put forth to handle the system?" This item was rated on a 7-point Likert scale (1 = very low, 7 = very high) and is intended to capture the user's felt effort required to interact successfully with the interface.

All items were translated into French using a forward-backward translation procedure to preserve wording and meaning.

3.2.3 Cognitive Propensity Tasks

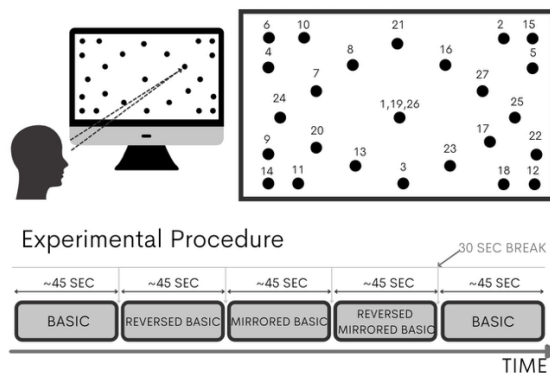
Two standardized visual tasks were used to assess individual cognitive propensities: a Large Grid paradigm (Son *et al.*, 2020) and a Visual Symbol Search task adapted from the Symbol Search subtest of the Wechsler Intelligence Scale (Drozdzick *et al.*, 2018)

(Wechsler, 2012; Drozdick *et al.*, 2018). An established computerized version of the Symbol Search task was used in this study (Kastrati *et al.*, 2021).

The Large Grid paradigm measures the efficiency of goal-directed visual search by requiring participants to fixate a sequence of dots appearing at different screen locations (Son *et al.*, 2020). Each trial displays one dot for 1.5 to 1.8 seconds at one of 25 predefined positions, covering the full screen and its center. In each block, 27 dots are presented, including three repetitions of the central dot. Five pseudo-randomized sequences are used to reduce predictability, producing uniform coverage of the screen across blocks.

For the present study, the stimulus sequence was presented in three consecutive blocks: a practice block followed by two experimental blocks with stable performance. Only Block 3 was retained for analysis, as it yielded the most reliable performance. Goal-oriented visual search propensity was computed as the mean gaze-shift latency between successive dots, a performance indicator derived from the sequential-fixation structure of the Large Grid paradigm (Son *et al.*, 2020). The figure below provides an overview of the Large Grid paradigm.

Figure 3.2 Overview of the Large Grid Paradigm



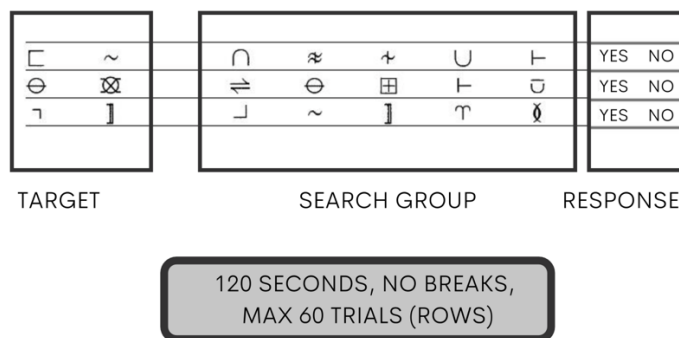
*Adapted from Kastrati and al. (2021)

The Visual Symbol Search (VSS) task is a computerized version of the Symbol Search subtest of the Wechsler Adult Intelligence Scale (Wechsler, 2012; Drozdick *et al.*, 2018). The task measures visual discrimination efficiency, that is, the ability to rapidly detect

whether specific target symbols appear among a set of similar distractors. In each row, participants are shown two target symbols, five search symbols, and two response options ("YES" / "NO"). They must indicate whether one of the targets appears in the search group. The task lasts 120 seconds, and participants complete as many rows as possible.

Performance was summarized using the number of correct responses completed within the allotted time, a standard metric for this paradigm. Higher scores reflect more efficient visual discrimination. As with the other visual task, only data from the third block were retained for analysis, as this block provides the most stable performance. The figure 3.3 below provides an overview of the VSS.

Figure 3.3 Overview of the Visual Symbol Search Task



*Adapted from Kastrati and al. (2021)

3.2.4 Laboratory Setup

The laboratory room was kept distraction-free, with constant lighting, sound isolation, and a neutral setting. Participants were seated at a workstation equipped with a 21-inch monitor, including an integrated Tobii Pro eye-tracker (Tobii AB, Sweden), a webcam positioned above the screen. The eye-tracker recorded fixations, visits, and gaze dispersion. The research team monitored the sessions from an adjacent control room through a one-way mirror, using three screens to (1) supervise the video feed and protocol, (2) take notes in

real time, and (3) track eye-tracking calibration and quality via Tobii Pro Lab (Tobii AB, Sweden).

3.3 Participants

Twenty-seven participants were recruited through our institution's participant recruitment panel. Participants were between 20 and 35 years old ($M = 25.39$, $SD = 4.24$), including 8 men and 19 women. This age range was chosen because young adults are typically very familiar with digital environments and technology-mediated search, having grown up in contexts where internet and digital tools are pervasive (Markoc, 2025). Inclusion criteria were: (1) being over 18 years old, (2) having advanced oral and written proficiency in French, and (3) being at least minimally familiar with a national public media, as the study was conducted in the context of a broader research project involving such an organization. Each 90-minute session took place at an academic research laboratory specializing in UX. Participants received CAD \$30 compensation. The study was approved by our institution's Research Ethics Board (certificate 2025-6409).

3.4 Data Extraction and Analysis

Analyses were conducted across all seven experimental tasks, which represented different AI-generated response behaviors. This design allowed for a comprehensive examination of how participants visually processed and evaluated varied AI responses.

Statistical analyses were performed in SAS 9.4 (SAS Institute Inc., Cary, NC, USA). Hypotheses were tested using mixed-effects models with random intercepts to account for repeated measures nested within participants. Continuous variables (e.g., fixation duration, K-coefficient, NASA-TLX scores) were analyzed with linear mixed models using the Kenward-Roger method. Count variables (e.g., fixation counts) were modeled with mixed-effects negative binomial regression. Ordinal variables (e.g., CES) were treated as continuous when distributional assumptions allowed. Dispositional variables (e.g., reading

habits) were added as moderators through interaction terms. As all hypotheses were directional, one-tailed tests were applied when reporting the p-values.

4. Results

4.1 Visual Attention Focus

H1 predicted that higher goal-oriented visual search propensity would be associated with more focused visual attention. A linear mixed-effects model was used to test this hypothesis across three indicators of visual attention: the proportion of total fixation duration within the AOI, the proportion of total fixations within the AOI, and the K-coefficient of attentional dispersion. Because the hypotheses were directional, p-values were adjusted using one-tailed tests. Visual search propensity did not significantly predict any of these measures. For fixation-duration proportion, the effect was essentially null ($\beta = 0.00$, $t(1084) = 0.00$, $p = 0.500$). The proportion of fixation counts showed a similar pattern ($\beta = 0.00$, $t(1) = 0.00$, $p = 0.500$). The K-coefficient likewise showed no detectable association ($\beta = -0.09$, $t(25.05) = -0.35$, $p = .36$). Overall, H1 was not statistically supported.

H2 posited that a higher propensity for visual discrimination would be associated with less focused visual attention, reflected in a more dispersed distribution of fixations. A linear mixed-effects model was used to test this hypothesis across the same three indicators of visual attention: the proportion of total fixation duration within the AOI, the proportion of total fixations within the AOI, and the K-coefficient of attentional dispersion. Because the hypotheses were directional, p-values were adjusted using one-tailed tests. Visual discrimination propensity did not significantly predict any of these measures. For fixation-duration proportion, the effect was essentially null ($\beta = 0.00$, $t(1084) = 0.00$, $p = 0.500$). The proportion of fixation counts showed a similarly null association ($\beta = 0.00$, $t(1) = 0.00$, $p = 0.500$). The K-coefficient showed a very small negative effect ($\beta = -0.00014$, $t(23.80) = -0.75$, $p = .23$), which likewise failed to reach significance. Overall, H2 was not statistically supported.

4.2 Cognitive Load

H3 predicted that a higher propensity for goal-oriented visual search would be associated with lower cognitive load. A linear mixed-effects model was used to test this hypothesis across two measures of perceived cognitive load: perceived mental demand (NASA-TLX) and perceived cognitive effort (CES). The results support this prediction. Goal-oriented visual search propensity was significantly associated with lower perceived mental demand ($\beta = -47.25$, $t(23.16) = -4.07$, $p < .001$). A similar pattern was observed for perceived cognitive effort, which also showed a significant negative association ($\beta = -1.66$, $t(25) = -1.83$, $p = .040$). Together, these findings indicate that individuals with higher goal-oriented visual search propensity experienced lower subjective cognitive load across both measures. Therefore, H3 is statistically supported.

H4 posited that a higher propensity for visual discrimination would be associated with greater cognitive load. A linear mixed-effects model was used to test this hypothesis across two measures of perceived cognitive load: perceived mental demand (NASA-TLX) and perceived cognitive effort (CES). The results partially support this prediction. Visual discrimination propensity was significantly associated with higher perceived mental demand ($\beta = 0.025$, $t(23.87) = 2.41$, $p = .012$), consistent with the directional hypothesis. In contrast, the association with perceived cognitive effort was positive but not significant ($\beta = 0.00024$, $t(25) = 0.32$, $p = .38$). Thus, while individuals with higher visual discrimination propensity reported greater perceived mental demand, this effect did not extend to cognitive effort. Therefore, H4 is partially supported.

4.3 Moderating Effects of Reading Habits

H5 posited that the effect of goal-oriented visual search propensity on cognitive load would be attenuated among users with stronger reading habits, implying a moderating role of

reading habits. A linear mixed-effects model was used to test this hypothesis across two measures of perceived cognitive load: mental demand (NASA-TLX) and cognitive effort (CES). Consistent with H3, goal-oriented visual search propensity was significantly associated with lower mental demand ($\beta = -147.68$, $t(20.48) = -2.37$, $p = .014$) and lower cognitive effort ($\beta = -9.53$, $t(23.08) = -1.87$, $p = .037$). The interaction between visual search propensity and reading habits did not reach statistical significance for mental demand ($\beta = 18.83$, $t(20.79) = 1.69$, $p = .104$). However, a marginal moderating effect was observed for cognitive effort ($\beta = 1.44$, $t(23) = 1.59$, $p = .063$), consistent with the directional hypothesis. These patterns suggest that stronger reading habits may attenuate the negative association between goal-oriented visual search propensity and perceived cognitive load, although this effect emerged only for cognitive effort. Overall, H5 receives marginal support.

H6 posited that the effect of visual discrimination propensity on cognitive load would be attenuated among users with stronger reading habits, meaning that reading habits would moderate the relationship between visual discrimination and perceived cognitive load. This hypothesis was tested using linear mixed-effects models across two complementary measures of cognitive load: mental demand (NASA-TLX) and cognitive effort (CES). The main effects of visual discrimination propensity were positive and reached or approached significance under the directional hypothesis for both outcomes: perceived mental demand ($\beta = 0.08$, $t(21.68) = 1.75$, $p = .047$) and perceived cognitive effort ($\beta = 0.006$, $t(23.15) = 1.72$, $p = .050$). For the moderation tests, the interaction between visual discrimination propensity and reading habits did not reach significance for perceived mental demand ($\beta = -0.011$, $t(22.68) = -1.30$, $p = .104$). However, the interaction for perceived cognitive effort was significant ($\beta = -0.001$, $t(23.15) = -1.73$, $p = .048$), indicating that stronger reading habits attenuated the positive association between visual discrimination and perceived effort. Taken together, these results provide partial support for H6: reading habits significantly moderated the effect of visual discrimination propensity on perceived cognitive effort, but not on perceived mental demand.

To summarize the outcomes of all tested hypotheses, Table 4.1 presents the validity results for each hypothesis, the constructs involved, and whether each prediction was supported by the data. When a hypothesis was tested using multiple indicators of the same construct, the overall conclusion reflects the combined pattern of results.

Table 4.1 Validity Test Results

Hypothesis	From	To	Status
H1	Visual search	Visual attention	Unsupported
H2	Visual discrimination	Visual attention	Unsupported
H3	Visual search	Cognitive load	Supported
H4	Visual discrimination	Cognitive load	Partially supported*
H5	Visual search x Reading habits	Cognitive load	Marginally supported
H6	Visual discrimination x Reading habits	Cognitive load	Partially supported*

** One of the two cognitive-load indicators showed a significant effect, while the other did not.*

5. Discussion

5.1 Summary of Key Findings

This study examined how individual visual propensities shape visual attention and cognitive load during the reading of AI-generated text. Although neither goal-oriented visual search nor visual discrimination significantly predicted gaze-based visual attention metrics, both propensities were meaningfully associated with subjective cognitive load. Goal-oriented visual search consistently predicted lower mental demand and cognitive effort, while visual discrimination showed a tendency to increase mental demand. Reading habits appeared to attenuate these effects.

5.2 Interpretation by Hypothesis Family

H1 and H2 predicted that individual visual propensities would shape how readers allocate attention when processing AI-generated answers. In our experiment, these answers simulated the type of output produced by a news-oriented RAG system, but visual attention

was operationalized strictly within a single AOI corresponding to the textual answer displayed on the screen. Although the broader interface also included supporting articles that participants could explore, all eye-tracking indicators were computed within this answer AOI, consistent with the study's objective of examining how users read the primary response. Across the three indicators of visual attention (proportion of fixation duration, proportion of fixations, and attentional dispersion measured with the K-coefficient), no significant effects of goal-oriented visual search or visual discrimination were observed. In practice, the effects were essentially null, indicating that individual visual propensities did not translate into measurable differences in how participants visually scanned the AI-generated paragraph itself. Several characteristics of the AI-mediated reading context likely contributed to these null findings.

First, these null effects should be interpreted in light of the experimental artefact and the resulting reading configuration. Because the AI response was displayed as a single, static paragraph within one predefined AOI, the answer functioned as a compact, self-contained summary: within this region, most words were potentially relevant to the task, leaving readers little incentive to adopt markedly different scanning strategies. As a result, gaze tended to remain within the paragraph, constraining between-participant variability in fixation allocation and dispersion (Orquin and Holmqvist, 2018). In other words, the short, dense, and linear structure of the generated answer likely limited how much visual attention could vary within the text region. Second, fixation duration within the answer is strongly driven by linguistic properties of the generated text (Rayner, 1998). Because all participants read the same outputs, and reading is a highly automatized skill, variance in fixation duration was likely dominated by sentence structure, word length, and local complexity rather than by stable visual tendencies (Rayner, 2009). This suggests that visual propensities may influence how readers evaluate or interpret the information provided by the RAG more than they affect low-level eye-movement timing during straightforward reading.

Third, the K-coefficient was computed over the entire response page but remained closely tied to the answer AOI, as participants quickly understood that this paragraph contained

the key information needed to complete the task. In a context where the answer is visually central and the supporting sources are peripheral, dispersion metrics become less diagnostic: readers rarely need to shift attention away from the text, and interindividual differences in spatial exploration are minimal.

Taken together, these findings indicate that within the AI-generated answer region, attentional trait differences are not readily captured by traditional eye-movement metrics. Instead, as shown by the cognitive load results, individual visual propensities seem to manifest more at a cognitive level (in perceived mental demand and cognitive effort) than at the level of overt gaze behavior. From a UX perspective, this suggests that when evaluating AI interfaces, eye-tracking within the answer AOI may be less sensitive to cognitive diversity than subjective or performance-based measures, especially when the answer is presented as a compact, linear block of text.

Hypotheses H3 and H4 examined how visual propensities shape the cognitive experience of reading AI-generated answers. Unlike the visual attention results, cognitive load revealed clear and meaningful differences between individuals. Goal-oriented visual search propensity was a consistent predictor of lower cognitive load, both for perceived mental demand and cognitive effort. In the context of an AI interface, this suggests that users with stronger goal-directed attentional control manage the answer more efficiently: they identify what matters within the generated paragraph, suppress irrelevant details, and maintain a stable attentional set while reading. This interpretation aligns with Cognitive Load Theory (Sweller, 1988; Paas, Renkl and Sweller, 2003) and with multimedia learning principles related to cognitive load (Mayer and Moreno, 2003). When interacting with AI-generated content, extraneous load arises largely from structural properties of the text and from uncertainty about how much of the answer must be retained or trusted. Readers skilled in goal-oriented visual search appear better equipped to manage these ambiguities. By filtering the AI-generated response strategically, they reduce unnecessary processing and devote working-memory resources to the core informational message.

From a UX perspective, these findings imply that AI systems may inherently advantage users with strong selective-attention skills: they navigate the generated text more easily, perceive the task as lighter, and extract meaning with less mental effort. For users with weaker goal-directed visual abilities, the same text may feel denser or more demanding, even without observable differences in eye-movement patterns.

In contrast, visual discrimination propensity, people's tendency to examine fine-grained or subtle visual differences, was associated with higher cognitive load. In an AI-reading context, this likely reflects a more analytical and detail-oriented processing style: readers high in discrimination may scrutinize the structure, coherence, or micro-level inconsistencies of the generated paragraph. Because AI-produced answers do not always follow humanlike discourse structure, this fine-grained evaluation can become taxing. Users may expend additional attention resolving ambiguities, comparing details, or verifying coherence within the text. This deep, detail-sensitive processing increases mental demand even if it does not translate into higher self-reported effort on the CES scale.

Within Cognitive Load Theory (Sweller, 1988; Paas, Renkl and Sweller, 2003), this pattern corresponds to a rise in intrinsic load (more semantic micro-analysis) and potentially extraneous load (extra attention devoted to low-level distinctions). This suggests that users high in visual discrimination may be more sensitive to subtle imperfections or structural irregularities in AI-generated text, making the experience feel cognitively heavier.

Hypotheses H5 and H6 examined whether reading habits moderated the influence of visual propensities on cognitive load while participants processed AI-generated news responses. Although the moderating effects were modest, marginal for goal-oriented visual search and significant only under a directional criterion for the interaction with visual discrimination on cognitive effort, the patterns across models consistently pointed toward an attenuating influence. Readers who reported stronger reading habits, meaning that they read frequently, enjoy reading, and perceive themselves as skilled readers, experienced less cognitive strain despite the cognitive tendencies that would otherwise predispose them to higher load.

This attenuation is particularly relevant within the context of a news-oriented RAG system. The generated answer that participants had to read was compact, linear, and often lacked explicit hierarchical cues, and users had no guidance regarding which parts of the paragraph were most important. Under such conditions, individuals with weaker goal-directed attentional control typically experienced higher cognitive effort, yet these demands were reduced among participants with strong reading habits. Likewise, individuals high in visual discrimination reported greater mental demand when processing the AI-generated answer, but this association was less pronounced among frequent and confident readers. These readers appeared better able to prioritize meaning over surface-level inspection, making the irregularities typical of generative outputs less mentally taxing.

A plausible explanation is that reading habits reflect a form of automatized cognitive infrastructure that reduces the workload associated with processing dense or ambiguously structured text. Extensive reading experience frees attentional resources that would otherwise be devoted to basic decoding or structural inference, allowing readers to focus more directly on meaning construction (LaBerge and Samuels, 1974). In an AI-news context, where sentences may not always flow smoothly and rhetorical cues can be hard to detect, this automatization seems to act as a compensatory mechanism. Readers with stronger reading habits depend less on effortful visual scanning and more on their well-practiced comprehension skills, which reduces the cognitive impact of their underlying visual propensities.

The motivational and affective dimensions of reading habits likely reinforce this effect. Readers who enjoy reading or feel confident in their abilities typically approach demanding textual material with greater persistence and a more adaptive mindset (Guthrie and Wigfield, 2000). When confronted with the uncertain structure of an AI-generated answer, such readers may be better equipped to tolerate ambiguity and deploy flexible comprehension strategies, which in turn reduces perceived effort. In this sense, reading habits act not only on cognitive mechanisms but also on metacognitive and affective regulation, shaping how visual tendencies translate into subjective load.

Altogether, the findings from H5 and H6 suggest that reading habits play a meaningful, if subtle, role in shaping users' cognitive experience when interacting with AI-generated news content. Rather than functioning as a simple performance variable, reading experience appears to modulate how attentional traits manifest during comprehension, absorbing part of the cognitive strain induced by the structure and quality of the generated text. This underscores the importance of considering individual reading backgrounds when evaluating UX in AI-mediated news environments, as well as the value of designing interfaces that support users who may not benefit from such protective cognitive routines.

6. Conclusion

6.1 Summary

Overall, the findings show that individual visual propensities meaningfully influence how readers experience cognitive load when processing AI-generated text. Goal-directed visual search reduces perceived demand, whereas fine-grained discrimination increases it. Although gaze measures did not capture these differences under the current task conditions, their directional alignment with subjective reports suggests that attentional trait differences manifest cognitively rather than behaviorally. Reading habits act as a protective factor, enabling more efficient processing of complex AI-generated content. These results advance theoretical understanding of cognitive diversity in AI-mediated reading and highlight the importance of designing adaptive AI interfaces that align information structure with users' cognitive capacities.

6.2 Theoretical Contributions

This study offers three main theoretical contributions. First, it extends Cognitive Load Theory by demonstrating that cognitive load is shaped not only by task structure but also by stable visual propensities (Sweller, 1988). Second, it highlights the dissociation between

subjective cognitive load and gaze behavior, reinforcing the theoretical distinction between cognitive and behavioral manifestations of attention (Rayner, 1998; Hyönä, 2010). Third, it identifies reading habits as a metacognitive buffer moderating trait-based differences, contributing to theories linking reading expertise, cognitive efficiency, and self-efficacy (Bandura, 1997; Guthrie and Wigfield, 2000).

6.3 Design Implications

The findings offer several implications for the design of AI systems that present textual information. First, the consistent association between goal-oriented visual search propensity and lower cognitive load suggests that interfaces should support efficient attentional guidance. When users can more easily identify relevant information, they expend fewer cognitive resources to interpret AI-generated content (Ledneva and Kovalev, 2021). Designers can support this process by incorporating subtle signaling mechanisms, such as emphasis cues, semantic grouping, or lightweight visual anchors, that facilitate selective attention without imposing rigid structural templates (Xie *et al.*, 2017).

Second, the positive association between visual discrimination propensity and perceived mental demand indicates that users who tend to focus on fine-grained details may become cognitively overloaded when reading AI-generated text. For these users, overly dense or visually uniform outputs may exacerbate unnecessary analytical effort. Interfaces can mitigate this by offering optional clarity-enhancing features, such as adjustable spacing, highlight-on-hover explanations, or simplified text modes that reduce surface-level complexity.

Third, the marginal moderating effects of reading habits point to the potential value of adaptive or personalization-oriented presentation strategies. Individuals with stronger reading habits appeared more resilient to cognitive strain, regardless of their visual propensity profiles. AI systems could therefore benefit from lightweight adaptivity, either implicit (e.g., detecting repeated rereading or scrolling patterns) or explicit (e.g., user-

selected “reading profiles”), to adjust presentation pace, density, or supportive cues based on user needs.

Overall, these findings underscore the value of designing generative interfaces that accommodate diverse cognitive profiles. Rather than assuming a one-size-fits-all presentation style, systems could provide adjustable supports that help users navigate AI-generated text more efficiently, reducing extraneous cognitive load and enhancing comprehension across a range of visual and reading propensities.

6.4 Limitations and Research Avenues

This study involved short, static texts displayed within a single AOI, which reduced natural variability in attention and limited sensitivity of the eye-tracking indicators. The heterogeneity of task scenarios might have limited the variability needed to reach statistical significance. Additionally, no physiological metrics such as pupil dilation, known to correlate with cognitive effort, were included. Future research should incorporate multi-AOI interfaces, interactive reading tasks, and richer physiological measures, and should further investigate the interplay between visual propensities, reading experience, and users’ verification strategies when interacting with AI-generated information.

7. Bibliographie

Acheson, D.J., Wells, J.B. and MacDonald, M.C. (2008) “New and updated tests of print exposure and reading abilities in college students,” *Behavior Research Methods*, 40(1), pp. 278–289. Available at: <https://doi.org/10.3758/BRM.40.1.278>.

Ai, Q. *et al.* (2023) “Information Retrieval Meets Large Language Models: A Strategic Report from Chinese IR Community.” arXiv. Available at: <https://doi.org/10.48550/arXiv.2307.09751>.

Andreou, G. *et al.* (2024) “Tracking Adults’ Eye Movements to Study Text Comprehension: A Review Article,” *Languages*, 9(12). Available at: <https://doi.org/10.3390/languages9120360>.

Annapureddy, R., Fornaroli, A. and Gatica-Perez, D. (2025) “Generative AI Literacy: Twelve Defining Competencies,” *Digit. Gov.: Res. Pract.*, 6(1), p. 13:1-13:21. Available at: <https://doi.org/10.1145/3685680>.

Aron, E.N. and Aron, A. (1997) “Sensory-processing sensitivity and its relation to introversion and emotionality,” *Journal of Personality and Social Psychology*, 73(2), pp. 345–368. Available at: <https://doi.org/10.1037/0022-3514.73.2.345>.

Baker, C. and Love, T. (2023) “The effect of time on lexical and syntactic processing in aphasia,” *Journal of Neurolinguistics*, 67, p. 101142. Available at: <https://doi.org/10.1016/j.jneuroling.2023.101142>.

Bandura, A. (1997) *Self-efficacy: The exercise of control*. New York, NY, US: W H Freeman/Times Books/ Henry Holt & Co (Self-efficacy: The exercise of control), pp. ix, 604.

Belardinelli, A., Herbort, O. and Butz, M.V. (2015) “Goal-oriented gaze strategies afforded by object interaction,” *Vision Research*, 106, pp. 47–57. Available at: <https://doi.org/10.1016/j.visres.2014.11.003>.

- Bick, A., Blandin, A. and Deming, D. (no date) *The State of Generative AI Adoption in 2025*. Available at: <https://www.stlouisfed.org/on-the-economy/2025/nov/state-generative-ai-adoption-2025> (Accessed: November 25, 2025).
- Chen, X., Li, J. and Ye, Y. (2024) “A feasibility study for the application of AI-generated conversations in pragmatic analysis,” *Journal of Pragmatics*, 223, pp. 14–30. Available at: <https://doi.org/10.1016/j.pragma.2024.01.003>.
- Clarke, A.D. *et al.* (2022) “Stable individual differences in strategies within, but not between, visual search tasks,” *Quarterly Journal of Experimental Psychology* (2006), 75(2), pp. 289–296. Available at: <https://doi.org/10.1177/1747021820929190>.
- Corbetta, M. and Shulman, G.L. (2002) “Control of goal-directed and stimulus-driven attention in the brain,” *Nature Reviews. Neuroscience*, 3(3), pp. 201–215. Available at: <https://doi.org/10.1038/nrn755>.
- Cui, H. *et al.* (2026) “An LLM-based cross-domain knowledge retrieval augmented generation method for bio-inspired solution design,” *Advanced Engineering Informatics*, 69, p. 104017. Available at: <https://doi.org/10.1016/j.aei.2025.104017>.
- Drozdzick, L.W. *et al.* (2018) “The Wechsler Adult Intelligence Scale—Fourth Edition and the Wechsler Memory Scale—Fourth Edition,” in *Contemporary intellectual assessment: Theories, tests, and issues, 4th ed.* New York, NY, US: The Guilford Press, pp. 486–511.
- Dubay, W. (2004) “The Principles of Readability,” *CA*, 92627949, pp. 631–3309.
- Frazier, L. and Rayner, K. (1982) “Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences,” *Cognitive Psychology*, 14(2), pp. 178–210. Available at: [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1).
- Guthrie, J.T. and Wigfield, A. (2000) “Engagement and motivation in reading,” in *Handbook of reading research, Vol. III*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 403–422.
- Hart, S.G. and Staveland, L.E. (1988) “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” in *Human mental workload*. Oxford, England: North-Holland (Advances in psychology, 52), pp. 139–183. Available at: [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9).
- Hautala, J. *et al.* (2022) “Task-oriented reading efficiency: interplay of general cognitive ability, task demands, strategies and reading fluency,” *Reading and Writing*, 35(8), pp. 1787–1813. Available at: <https://doi.org/10.1007/s11145-022-10265-7>.
- Hautus, M., Macmillan, N. and Creelman, C. (2021) *Detection Theory: A User’s Guide*. Available at: <https://doi.org/10.4324/9781003203636>.

- Herbold, S. *et al.* (2023) “A large-scale comparison of human-written versus ChatGPT-generated essays,” *Scientific Reports*, 13(1), p. 18617. Available at: <https://doi.org/10.1038/s41598-023-45644-9>.
- Hyönä, J. (2010) “The use of eye movements in the study of multimedia learning,” *Learning and Instruction*, 20(2), pp. 172–176. Available at: <https://doi.org/10.1016/j.learninstruc.2009.02.013>.
- Ilyas, C.M.A. *et al.* (2025) “Reading the Readers Mind through Eye Tracking: Can AI Generated Texts Match Human Authors?,” in *Proceedings of the 2025 Symposium on Eye Tracking Research and Applications. ETRA '25: 2025 Symposium on Eye Tracking Research and Applications*, Tokyo Japan: ACM, pp. 1–7. Available at: <https://doi.org/10.1145/3715669.3726846>.
- Jake Follmer, D. and Sperling, R.A. (2018) “Interactions between reader and text: Contributions of cognitive processes, strategy use, and text cohesion to comprehension of expository science text,” *Learning and Individual Differences*, 67, pp. 177–187. Available at: <https://doi.org/10.1016/j.lindif.2018.08.005>.
- Joty, S., Carenini, G. and Ng, R.T. (2015) “CODRA: A Novel Discriminative Framework for Rhetorical Analysis,” *Computational Linguistics*, 41(3), pp. 385–435. Available at: https://doi.org/10.1162/COLI_a_00226.
- Just, M.A. and Carpenter, P.A. (1980) “A theory of reading: From eye fixations to comprehension,” *Psychological Review*, 87(4), pp. 329–354. Available at: <https://doi.org/10.1037/0033-295X.87.4.329>.
- Kastrati, A. *et al.* (2021) “EEGEyeNet: a Simultaneous Electroencephalography and Eye-tracking Dataset and Benchmark for Eye Movement Prediction.” arXiv. Available at: <https://doi.org/10.48550/arXiv.2111.05100>.
- Kennison, S.M. and Clifton, C. (1995) “Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), pp. 68–81. Available at: <https://doi.org/10.1037/0278-7393.21.1.68>.
- Kim, S.S.Y. *et al.* (2024) “‘I’m Not Sure, But...’: Examining the Impact of Large Language Models’ Uncertainty Expression on User Reliance and Trust,” *The 2024 ACM Conference on Fairness Accountability and Transparency*, 3 June, pp. 822–835.
- Kim, Z.M. *et al.* (2025) “Align to Structure: Aligning Large Language Models with Structural Information.” arXiv. Available at: <https://doi.org/10.48550/arXiv.2504.03622>.
- Kintsch, W. (1998) *Comprehension: A paradigm for cognition*. New York, NY, US: Cambridge University Press (Comprehension: A paradigm for cognition), pp. xvi, 461.

- Krejtz, K. *et al.* (2018) “Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze,” *PLOS ONE*, 13(9), p. e0203629. Available at: <https://doi.org/10.1371/journal.pone.0203629>.
- Kuperman, V., Matsuki, K. and Van Dyke, J.A. (2018) “Contributions of reader- and text-level characteristics to eye-movement patterns during passage reading,” *Journal of experimental psychology. Learning, memory, and cognition*, 44(11), pp. 1687–1713. Available at: <https://doi.org/10.1037/xlm0000547>.
- LaBerge, D. and Samuels, S.J. (1974) “Toward a theory of automatic information processing in reading,” *Cognitive Psychology*, 6(2), pp. 293–323. Available at: [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2).
- Ledneva, T. and Kovalev, A. (2021) “Cognitive load measurement during navigation and information retrieval in digital text,” *Procedia Computer Science*, 192, pp. 2720–2730. Available at: <https://doi.org/10.1016/j.procs.2021.09.042>.
- Lin, Z. *et al.* (2024) “Towards trustworthy LLMs: a review on debiasing and dehallucinating in large language models,” *Artificial Intelligence Review*, 57(9), p. 243. Available at: <https://doi.org/10.1007/s10462-024-10896-y>.
- Maniscalco, B. and Lau, H. (2015) “Manipulation of working memory contents selectively impairs metacognitive sensitivity in a concurrent visual discrimination task,” *Neuroscience of Consciousness*, 2015(1). Available at: <https://doi.org/10.1093/nc/niv002>.
- Markoc, I. (2025) “The digitalization tendency of young adults: differences by living environment, gender and education,” *Online Information Review*, 49(1), pp. 136–163. Available at: <https://doi.org/10.1108/OIR-01-2023-0020>.
- Mayer, R.E. and Moreno, R. (2003) “Nine ways to reduce cognitive load in multimedia learning,” *Educational Psychologist*, 38(1), pp. 43–52. Available at: https://doi.org/10.1207/S15326985EP3801_6.
- McNamara, D. *et al.* (2002) “Are Good Texts Always Better? Interactions of Text Coherence, Background Knowledge, and Levels of Understanding in Learning From Text,” *Cognition and Instruction*, 14. Available at: https://doi.org/10.1207/s1532690xc1401_1.
- McVay, J.C. and Kane, M.J. (2012) “Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention,” *Journal of Experimental Psychology: General*, 141(2), pp. 302–320. Available at: <https://doi.org/10.1037/a0025250>.
- Mézière, D.C. *et al.* (2023) “Using Eye-Tracking Measures to Predict Reading Comprehension,” *Reading Research Quarterly*, 58(3), pp. 425–449. Available at: <https://doi.org/10.1002/rrq.498>.

- Niwanputri, G.S., Tob-Ogu, A. and Hosseinzadeh, M. (2025) “Untangling Cognitive Processes in Academic Information Searching: Investigating Tool Gaps, Cognitive Load, and User Satisfaction,” in *Proceedings of the 2025 International ACM SIGIR Conference on Innovative Concepts and Theories in Information Retrieval (ICTIR)*. ICTIR '25: *International ACM SIGIR Conference on Innovative Concepts and Theories in Information Retrieval*, Padua Italy: ACM, pp. 449–458. Available at: <https://doi.org/10.1145/3731120.3744615>.
- Orquin, J.L. and Holmqvist, K. (2018) “Threats to the validity of eye-movement research in psychology,” *Behavior Research Methods*, 50(4), pp. 1645–1656. Available at: <https://doi.org/10.3758/s13428-017-0998-z>.
- Paas, F., Renkl, A. and Sweller, J. (2003) “Cognitive load theory and instructional design: Recent developments,” *Educational Psychologist*, 38(1), pp. 1–4. Available at: https://doi.org/10.1207/S15326985EP3801_1.
- Peng, P. *et al.* (2018) “A meta-analysis on the relation between reading and working memory,” *Psychological Bulletin*, 144(1), pp. 48–76. Available at: <https://doi.org/10.1037/bul0000124>.
- Proulx, M.J. (2011) “Individual Differences and Metacognitive Knowledge of Visual Search Strategy,” *PLoS ONE*. Edited by S. Suzuki, 6(10), p. e27043. Available at: <https://doi.org/10.1371/journal.pone.0027043>.
- Rayner, K. (1998) “Eye movements in reading and information processing: 20 years of research,” *Psychological Bulletin*, 124(3), pp. 372–422. Available at: <https://doi.org/10.1037/0033-2909.124.3.372>.
- Rayner, K. (2009) “Eye movements and attention in reading, scene perception, and visual search,” *Quarterly Journal of Experimental Psychology* (2006), 62(8), pp. 1457–1506. Available at: <https://doi.org/10.1080/17470210902816461>.
- Rayner, K. and Duffy, S.A. (1986) “Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity,” *Memory & Cognition*, 14(3), pp. 191–201. Available at: <https://doi.org/10.3758/BF03197692>.
- Richter, M., Gendolla, G.H.E. and Wright, R.A. (2016) “Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don’t Know,” in A.J. Elliot (ed.) *Advances in Motivation Science*. Elsevier, pp. 149–186. Available at: <https://doi.org/10.1016/bs.adms.2016.02.001>.
- Ronconi, A., Altoè, G. and Mason, L. (2025) “Ink and pixels: Impact of highlighting and reading self-efficacy on adolescents’ cognitive load, epistemic emotions, and text comprehension,” *Learning and Individual Differences*, 117, p. 102604. Available at: <https://doi.org/10.1016/j.lindif.2024.102604>.

Rubio, S. *et al.* (2004) “Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods,” *Applied Psychology*, 53(1), pp. 61–86. Available at: <https://doi.org/10.1111/j.1464-0597.2004.00161.x>.

Skalicky, S. and Dascalu, M. (2019) “Moving beyond classic readability formulas: new methods and new models,” *Journal of Research in Reading*, 42. Available at: <https://doi.org/10.1111/1467-9817.12283>.

Smith, R. *et al.* (2021) “The Role of Background Knowledge in Reading Comprehension: A Critical Review,” *Reading Psychology*, 42(3), pp. 214–240. Available at: <https://doi.org/10.1080/02702711.2021.1888348>.

Son, J. *et al.* (2020) “Evaluating fMRI-Based Estimation of Eye Gaze During Naturalistic Viewing,” *Cerebral Cortex (New York, NY)*, 30(3), pp. 1171–1184. Available at: <https://doi.org/10.1093/cercor/bhz157>.

Stanovich, K.E. (1986) “Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy,” *Reading Research Quarterly*, 21(4), pp. 360–407. Available at: <https://doi.org/10.1598/RRQ.21.4.1>.

Sweller, J. (1988) “Cognitive Load During Problem Solving: Effects on Learning,” *Cognitive Science*, 12(2), pp. 257–285. Available at: https://doi.org/10.1207/s15516709cog1202_4.

Treisman, A.M. and Gelade, G. (1980) “A feature-integration theory of attention,” *Cognitive Psychology*, 12(1), pp. 97–136. Available at: [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5).

Twabu, K. (2025) “Enhancing the cognitive load theory and multimedia learning framework with AI insight,” *Discover Education*, 4. Available at: <https://doi.org/10.1007/s44217-025-00592-6>.

Vaessen, A. and Blomert, L. (2010) “Long-term cognitive dynamics of fluent reading development,” *Journal of Experimental Child Psychology*, 105(3), pp. 213–231. Available at: <https://doi.org/10.1016/j.jecp.2009.11.005>.

Wechsler, D. (2012) “Wechsler Adult Intelligence Scale--Fourth Edition.” Available at: <https://doi.org/10.1037/t15169-000>.

Wendt, D., Brand, T. and Kollmeier, B. (2014) “An Eye-Tracking Paradigm for Analyzing the Processing Time of Sentences with Different Linguistic Complexities,” *PLOS ONE*, 9(6), p. e100186. Available at: <https://doi.org/10.1371/journal.pone.0100186>.

Wolfe, J.M. (1994) “Guided Search 2.0 A revised model of visual search,” *Psychonomic Bulletin & Review*, 1(2), pp. 202–238. Available at: <https://doi.org/10.3758/BF03200774>.

Wolfe, J.M. and Horowitz, T.S. (2017) “Five factors that guide attention in visual search,” *Nature Human Behaviour*, 1(3), p. 0058. Available at: <https://doi.org/10.1038/s41562-017-0058>.

Xie, H. *et al.* (2017) “The more total cognitive load is reduced by cues, the better retention and transfer of multimedia learning: A meta-analysis and two meta-regression analyses,” *PLOS ONE*. Edited by M.A. Motes, 12(8), p. e0183884. Available at: <https://doi.org/10.1371/journal.pone.0183884>.

Zhang, X. and Lu, X. (2025) “Aligning linguistic complexity with the difficulty of English texts for L2 learners based on CEFR levels,” *Studies in Second Language Acquisition*, pp. 1–28. Available at: <https://doi.org/10.1017/S0272263125101125>.

Chapitre 3

1. Conclusion

1.1 Synthèse des résultats

Ce mémoire visait à examiner comment deux propensions cognitives — la recherche visuelle orientée vers un objectif et la discrimination visuelle — influencent l'attention visuelle et la charge cognitive lors de la lecture de réponses générées par l'intelligence artificielle, tout en évaluant le rôle modérateur des habitudes de lecture. Une expérimentation a été menée au Tech3Lab auprès de 27 participants, chacun effectuant 7 tâches de lecture de réponses textuelles produites par un prototype de modèle de langage. Cette étude a permis d'observer les comportements visuels et la charge cognitive perçue dans un environnement contrôlé.

Les résultats révèlent que les deux propensions cognitives exercent des effets opposés sur la charge cognitive. Les participants ayant une meilleure capacité de recherche visuelle rapportent une charge mentale plus faible, ce qui suggère qu'ils parviennent à cibler efficacement les éléments pertinents du texte sans effort excessif. À l'inverse, ceux qui obtiennent des scores élevés en discrimination visuelle rapportent une charge mentale plus

élevée, indiquant un traitement plus analytique et plus coûteux cognitivement. Les mesures oculométriques, bien que non significatives, vont dans le même sens : les lecteurs à forte recherche visuelle présentent une attention plus concentrée sur la zone textuelle, tandis que ceux à forte discrimination la parcourent plus largement.

Les habitudes de lecture ne montrent pas d'effet modérateur significatif, mais plusieurs tendances cohérentes émergent. Pour la recherche visuelle, l'interaction avec les habitudes de lecture approche le seuil de signification : les participants qui lisent fréquemment, qui aiment lire et qui se sentent compétents rapportent généralement une charge mentale plus faible, indépendamment de leur propension à la recherche visuelle. Autrement dit, de bonnes habitudes de lecture semblent réduire globalement la charge mentale. Pour la discrimination visuelle, une tendance similaire apparaît : l'association positive entre discrimination et charge mentale semble légèrement atténuée chez les lecteurs expérimentés, bien que cet effet reste non significatif. Ces observations suggèrent que les habitudes de lecture pourraient jouer un rôle de protection cognitive, aidant certains lecteurs à gérer plus efficacement la charge mentale face à des textes denses ou moins clairement structurés.

Dans l'ensemble, ces résultats révèlent que la lecture de contenus générés par l'IA sollicite des processus cognitifs comparables à ceux de la lecture traditionnelle, tout en accordant un rôle particulièrement déterminant aux différences individuelles dans la gestion de la charge cognitive. Les propensions visuelles influencent peu le comportement oculomoteur observable, mais modulent nettement la charge cognitive perçue : une recherche visuelle efficace l'allège, tandis qu'une sensibilité accrue aux détails tend à l'alourdir. Les habitudes de lecture, bien qu'elles ne présentent que des effets modérateurs marginaux, semblent conférer une forme de résilience cognitive en atténuant ces variations. Ainsi, la compréhension de textes générés par l'IA ne dépend pas uniquement du contenu présenté, mais également des ressources perceptives et de l'expérience de lecture que chaque lecteur mobilise pour interpréter l'information. Ces résultats soulignent l'importance de prendre en compte les différences individuelles dans l'étude, l'évaluation et la conception d'interfaces génératives.

1.2 Contributions théoriques

Ce travail contribue à la compréhension du traitement cognitif des textes générés par l'IA en démontrant le rôle déterminant des différences individuelles dans l'expérience de lecture. Alors que les recherches sur la lecture s'appuient principalement sur des textes humains, structurés par une intention éditoriale explicite, cette étude révèle que, dans un contexte génératif, la charge cognitive dépend davantage des capacités perceptuelles et de l'expérience du lecteur que de ses comportements oculaires observables.

Les résultats indiquent que la lecture de réponses générées par l'IA ne mobilise pas de processus entièrement nouveaux, mais exige une gestion plus active des ressources cognitives. Les lecteurs s'appuient fortement sur leurs propensions visuelles — notamment la recherche visuelle orientée vers un objectif et la discrimination visuelle — pour réguler la charge cognitive nécessaire à la compréhension. Certaines différences individuelles s'avèrent ainsi particulièrement déterminantes dans ce contexte.

En identifiant ces propensions perceptuelles comme prédicteurs de la charge cognitive, cette étude ouvre une voie empirique permettant d'intégrer des caractéristiques individuelles issues de la psychologie cognitive dans l'analyse de la lecture de contenus algorithmiques. Elle démontre que ces propensions ne devraient pas être considérées uniquement comme des variables de contrôle, mais bien comme des facteurs explicatifs centraux de l'expérience utilisateur face aux systèmes génératifs.

1.3 Implications pratiques

Les résultats de cette recherche ouvrent des perspectives pour concevoir des interfaces génératives plus sensibles aux différences individuelles. La charge cognitive variant selon les propensions perceptuelles des lecteurs, les systèmes de génération de texte gagneraient

à intégrer des mécanismes qui favorisent l'efficacité attentionnelle et réduisent la charge cognitive.

L'association entre une recherche visuelle efficace et une charge cognitive réduite suggère que les interfaces pourraient faciliter le repérage de l'information pertinente, notamment par des aides subtiles à l'orientation ou la mise en évidence d'éléments clés. À l'inverse, la discrimination visuelle étant liée à une charge mentale accrue, des options d'ajustement de la densité visuelle, du rythme de présentation ou de la clarté des éléments pourraient soutenir les utilisateurs portés à analyser finement les détails perceptifs.

Plus largement, ces observations plaident en faveur d'interfaces adaptables, permettant à chacun de moduler la présentation de l'information selon ses préférences et capacités cognitives. Alors que les modèles de langage s'imposent comme outils quotidiens d'accès à l'information, concevoir des systèmes qui reconnaissent et accommodent la diversité cognitive devient essentiel pour améliorer l'expérience de lecture et prévenir la surcharge mentale.

1.4 Limites

Cette étude comporte certaines limites qui invitent à la prudence dans l'interprétation des résultats. La taille de l'échantillon ($n=27$) et sa relative homogénéité limitent la généralisation des conclusions à des populations plus diversifiées.

L'hétérogénéité du matériel textuel constitue également une limite. Les réponses variaient en longueur, en complexité et en contenu, ce qui a pu interagir avec les propensions cognitives mesurées. De plus, certaines variables non contrôlées, comme la vitesse de lecture, la familiarité avec les sujets abordés, ou la charge cognitive de base, peuvent avoir influencé les résultats.

Sur le plan statistique, plusieurs effets cohérents avec les hypothèses n'atteignent pas le seuil de significativité. Ces tendances, bien que prometteuses, reflètent probablement une

puissance statistique limitée et la variabilité inhérente aux mesures oculométriques. Elles suggèrent néanmoins des pistes à explorer avec des échantillons plus robustes.

Enfin, le dispositif expérimental, soit un prototype statique, ne reflète qu'imparfaitement l'usage réel des systèmes génératifs, où les utilisateurs peuvent reformuler leurs requêtes, naviguer entre plusieurs réponses ou consulter d'autres sources en parallèle.

1.5. Pistes de recherches futures

Au-delà de ces limites méthodologiques, plusieurs directions de recherche pourraient enrichir la compréhension du traitement cognitif de contenus générés par IA.

Cette étude s'est concentrée sur des requêtes personnelles et informatives, mais il serait pertinent d'examiner comment les propensions individuelles et la charge cognitive évoluent dans d'autres contextes d'usage. En milieu professionnel, où les utilisateurs consultent des IA pour rédiger des rapports ou analyser des données, les objectifs de lecture et les critères d'évaluation diffèrent considérablement. De même, dans des contextes éducatifs — où l'apprentissage assisté ou la résolution de problèmes sont en jeu — ou créatifs — comme le brainstorming ou la rédaction collaborative — les stratégies cognitives mobilisées pourraient être différentes. Comparer ces contextes permettrait de mieux comprendre comment l'intention de lecture module le traitement de l'information générée.

Les modèles génératifs produisent également des contenus dans des formats variés : texte long, listes à puces, tableaux, code, ou combinaisons multimodales intégrant texte et images. Comprendre comment ces différents formats influencent l'attention visuelle et la charge cognitive permettrait d'adapter la présentation de l'information au type de tâche et au profil de l'utilisateur. Certains lecteurs pourraient traiter plus efficacement des réponses structurées en listes, tandis que d'autres préféreraient des paragraphes narratifs. Explorer ces préférences et leur interaction avec les propensions cognitives offrirait des pistes concrètes pour personnaliser les interfaces.

Une autre piste prometteuse serait d'examiner l'effet de la littératie en IA sur le traitement de contenus générés. La littératie en IA — comprendre les limites des modèles, savoir évaluer la fiabilité d'une réponse, développer des stratégies de vérification — pourrait significativement influencer la charge cognitive et les *patterns* d'attention. Des études interventionnelles pourraient tester si former les utilisateurs à interroger et évaluer les systèmes génératifs améliore leur efficacité cognitive et réduit l'effort mental requis. Cette piste aurait des implications directes pour l'éducation et la formation professionnelle dans un monde où l'IA devient un outil quotidien.

Au-delà des propensions visuelles et des habitudes de lecture, d'autres caractéristiques individuelles mériteraient également d'être explorées. La tolérance à l'ambiguïté — soit la capacité à gérer l'incertitude d'une réponse potentiellement inexacte — pourrait moduler l'anxiété et la charge cognitive ressenties face à des contenus dont la fiabilité doit être constamment évaluée. Les styles cognitifs (traitement analytique versus holistique) pourraient influencer la manière dont les lecteurs naviguent et intègrent l'information. La métacognition, c'est-à-dire la capacité à évaluer sa propre compréhension et à ajuster ses stratégies en conséquence, représente un autre facteur potentiellement déterminant. Enfin, les attitudes envers l'IA — confiance, scepticisme, ou familiarité technologique — pourraient façonner non seulement l'expérience subjective, mais aussi les comportements de lecture eux-mêmes. Intégrer ces dimensions offrirait une compréhension plus complète de la diversité des expériences utilisateurs.

Des travaux longitudinaux permettraient également d'examiner comment l'expérience répétée avec les systèmes génératifs transforme les stratégies de lecture. Les utilisateurs développent-ils des heuristiques pour évaluer rapidement la qualité d'une réponse? La charge cognitive diminue-t-elle avec la familiarisation, ou certains lecteurs restent-ils durablement plus vulnérables à la surcharge mentale? Ces questions sont cruciales pour comprendre l'adaptation cognitive à long terme dans des environnements informationnels de plus en plus médiés par l'IA.

Enfin, alors que cette étude examine la lecture de réponses isolées, la plupart des interactions avec les modèles de langage sont conversationnelles et itératives. Étudier comment les utilisateurs ajustent leurs requêtes, gèrent des échanges multi-tours, ou intègrent des informations provenant de plusieurs réponses successives permettrait de saisir la complexité réelle du traitement cognitif dans ces environnements dynamiques.

Ces pistes convergent vers un même objectif : mieux comprendre comment les humains s'adaptent cognitivement à des environnements informationnels où la production de contenu est déléguée à des systèmes algorithmiques, afin de concevoir des interfaces qui soutiennent véritablement la diversité des profils et des besoins.