UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Towards Gaze-Informed AI Disclosure Interfaces: Eye-Tracking Attentional and Cognitive Load While Reading AI-Assisted News

Prajod, Pooja; Cools, Hannes; Röggla, Thomas; Cesar, Pablo; Ali, Abdallah El

Publication date
2026

Document Version
Final published version

License
CC BY

Link to publication

Citation for published version (APA):
Prajod, P., Cools, H., Röggla, T., Cesar, P., & Ali, A. E. (2026). *Towards Gaze-Informed AI Disclosure Interfaces: Eye-Tracking Attentional and Cognitive Load While Reading AI-Assisted News*. ArXiv. https://arxiv.org/abs/2605.14999v1

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:14 Oct 2026

arXiv:2605.14999v1 [cs.HC] 14 May 2026

Towards Gaze-Informed AI Disclosure Interfaces: Eye-Tracking Attentional and Cognitive Load While Reading AI-Assisted News

Pooja Prajod
Pooja.Prajod@cwi.nl
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands

Hannes Cools
University of Amsterdam
Amsterdam, The Netherlands

Thomas Röggla
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands

Pablo Cesar*
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands

Abdallah El Ali†
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands

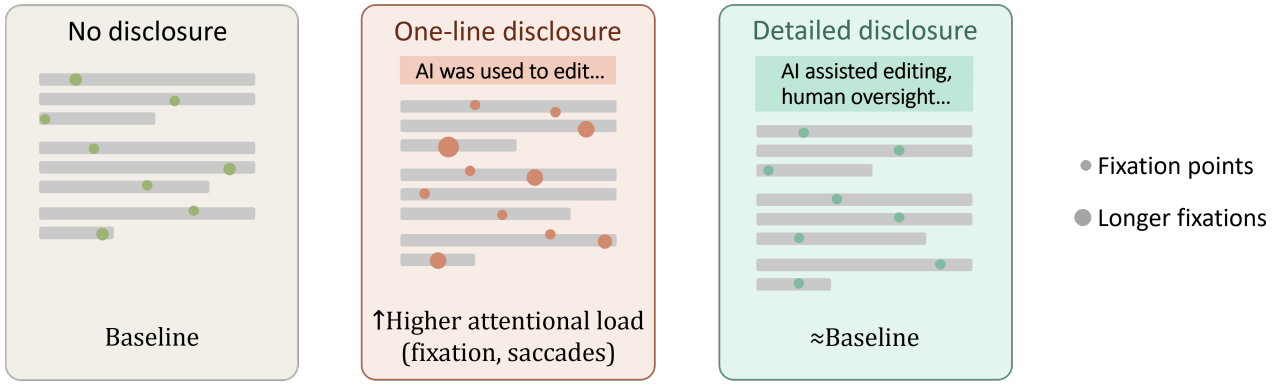

Figure 1: Illustration of key findings: one-line AI-use disclosures increase readers' attentional load (larger, more frequent fixation points) compared to no disclosure and detailed disclosure conditions.

Abstract

As generative AI becomes increasingly integrated into journalism, designing effective AI-use disclosures that inform readers without imposing unnecessary burden is a key challenge. While prior research has primarily focused on trust and credibility, the impact of disclosures on readers' attentional and cognitive load remains underexplored. To address this gap, we conducted a $3 \times 2 \times 2$ mixed factorial study manipulating the level of AI-use disclosure detail (none, one-line, detailed), news type (politics, lifestyle), and role of AI (editing, partial content generation), measuring load via NASA-TLX and eye-tracking. Our results reveal a significant attentional cost: one-line disclosures resulted in significantly higher fixation durations and saccade counts, particularly for AI-edited content. Detailed disclosures did not impose additional burden. Drawing on Information-Gap Theory, we argue that brief labels may trigger increased visual scrutiny by alerting readers to AI use without providing enough information. NASA-TLX scores and pupil diameter showed no significant differences across conditions, suggesting that AI-use disclosures do not impose cognitive burden regardless of the detail level. Interview insights contextualize these findings and reveal a strong preference for detailed or "detail-on-demand" designs. Our findings inform the design of gaze-informed adaptive disclosure interfaces that dynamically adjust transparency levels based on readers' attentional patterns and news context.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **User studies**; **Laboratory experiments**.

Keywords

AI-use Disclosures, Eye-tracking, Gaze Behavior, Attention, Cognitive Load, AI Journalism, Mixed-Methods, Physiological Computing

ACM Reference Format:
Pooja Prajod, Hannes Cools, Thomas Röggla, Pablo Cesar, and Abdallah El Ali. 2026. Towards Gaze-Informed AI Disclosure Interfaces: Eye-Tracking Attentional and Cognitive Load While Reading AI-Assisted News. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 10 pages. https://doi.org/XXXXXXX.XXXXXXX

*Also with TU Delft.
†Also with Utrecht University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference acronym 'XX, Woodstock, NY
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

As generative AI becomes increasingly integrated into news production, there are growing calls for transparency regarding its use [16, 58]. While disclosure statements are the primary mechanism for communicating AI involvement to readers, a key open challenge [50] is *designing disclosures that inform readers about AI use without imposing attentional and cognitive load.* Recent advancements in physiological computing suggest that real-time gaze sensing can enable interfaces to detect and adapt to such use states [13, 29, 45], which is a promising path towards user-centric, dynamic news interfaces. Given this potential, we investigate whether gaze signals can reveal when a disclosure design increases attentional and cognitive load during news reading.

Current AI-use disclosures in news are largely self-regulated and non-standardized [7, 58], ranging from minimal one-line labels to detailed transparency statements [2, 39, 44, 54]. AI's role in content production also varies, from editorial assistance to content generation [19, 30, 55]. To date, research has focused almost exclusively on trust and credibility [2, 32, 37, 39, 41, 44, 54], and the resulting reader engagement, such as sharing and subscription. However, the attentional and cognitive costs of AI disclosures remain underexplored. This gap is significant for news interface design as higher cognitive load can reduce information retention and increase susceptibility to misinformation [3, 15]. Understanding these effects requires real-time sensing methods, as subjective measures alone may not capture implicit processing differences [13, 33].

We bridge this gap by exploring two research questions central to the design of adaptive news interfaces:

- **Does the level of detail in AI disclosures influence readers' attentional and cognitive load?**
- **Does attentional and cognitive load differ depending on AI's role (editing vs. partial content generation)?**

We address these through a 3×2×2 mixed factorial study by manipulating AI disclosure detail (none, one-line, detailed), news type (politics, lifestyle), and AI role (editing, partial content generation). We measured load via NASA-TLX [22] and eye-tracking (pupil diameter, fixation duration, saccade count). NASA-TLX and pupil diameter capture cognitive load [29], while fixation duration and saccades reflect attentional load (visual search and spatial attention distribution) [49].

Our results reveal an attentional load cost for one-line disclosures, i.e., one-line disclosures increase attentional load, particularly for AI-edited content, whereas detailed disclosures do not impose additional burden despite containing more information. Drawing on Information-Gap Theory [35], we argue that brief labels may alert readers to AI use without providing enough information to resolve the reader's uncertainty, thereby triggering increased visual scrutiny as a form of information-seeking behavior. On the other hand, NASA-TLX scores and pupil diameter showed no significant differences across conditions, indicating that cognitive load remains unchanged regardless of disclosure detail.

Crucially, eye-tracking captured the attentional load differences, which highlights the value of real-time gaze sensing for designing and evaluating disclosure designs. Our findings inform the design of gaze-informed adaptive disclosure interfaces that dynamically adjust the detail level based on readers' attentional patterns and news context.

2 Related Work

2.1 Eye-Tracking Measures of Attentional and Cognitive Load

Eye-tracking measures are increasingly popular for assessing cognitive and attentional load in HCI [5, 28, 29]. In particular, pupil diameter is governed by the autonomic nervous system and is a well-established indicator of cognitive load, with larger pupil sizes reflecting higher cognitive effort [49]. Fixations, or focused gaze on areas of interest, and saccades, which are rapid eye movements between fixations, can be voluntary or involuntary. They are also associated with cognitive processing, but are more directly reflective of how attention is distributed across an interface. Longer fixation durations indicate greater processing difficulty at the point of gaze [14, 49], while higher saccade counts reflect increased visual search and re-reading behavior [49]. Although subjective measures such as NASA-TLX capture perceived task load, non-intrusive measures like eye-tracking are particularly valuable as they capture real-time, implicit responses that subjective measures may miss [23, 33, 46]. Moreover, real-time gaze measures have been proposed for adaptive interfaces [1, 4, 5, 31, 38, 40, 43, 53] and hence, eye-tracking is a crucial modality to consider for adaptive AI disclosures.

2.2 Attentional and Cognitive Load in News Reading

Recent works have used eye-tracking to investigate cognitive load in news reading contexts. In the misinformation domain, readers show longer fixation durations and increased pupil diameter when reading fake news compared to real news [20, 21, 48, 51]. For instance, an eye-tracking study [48] demonstrated the varying cognitive load and gaze patterns when reading true and false COVID-19 news headlines.

Recent works have also studied attentional and cognitive load in the context of AI-generated news content using eye-tracking and physiological measures [26, 56, 57]. These studies generally report higher load when reading AI-generated content compared to human-written content. In an eye-tracking study [26] comparing AI-generated and human-written texts, longer fixation durations were observed for AI-generated content, suggesting greater processing difficulty. This observation was not explained by differences in readability scores, leading the authors to attribute it to subtle stylistic differences. A multimodal study combining gaze and physiological measures explored how readers engage with human- and AI-generated real and fake news articles [56]. While readers' accuracy at distinguishing AI-generated from human-written articles was near chance level, machine learning models trained on physiological and gaze data were able to predict content source, indicating that implicit processing differences exist even when readers cannot explicitly identify them. Differences in psychophysiological responses to emotional AI-written news have been observed among young readers [57].

Notably, the above works focus on the load induced by fake news or AI-generated content itself, rather than by AI-use disclosures.

2.3 AI-Use Disclosures in News Reading

AI-use disclosures in news are currently self-regulated and non-standardized [7, 58]. Regulatory frameworks such as the EU AI Act propose transparency obligations for AI systems, but editorial processes and content generation typically fall outside their scope, leaving journalistic AI use largely unaddressed by explicit regulation [24, 58]. In practice, disclosure formats vary widely across news organizations, ranging from brief one-line labels to detailed transparency statements describing the specific production steps in which AI was involved [18, 27, 44]. Crucially, effective disclosure design requires balancing informativeness with usability [16, 30].

Prior research on AI-use disclosures in news has focused primarily on trust and credibility outcomes. Several studies have observed a "transparency dilemma" where disclosing AI involvement paradoxically reduces readers' trust [2, 39, 41, 54]. A recent study [2] found that labeling headlines as AI-generated reduced perceived accuracy regardless of the headline's veracity. The authors attributed this to AI-aversion caused by readers assuming fully AI-generated content. Similarly, studies reported that AI-generated news is believed less than human-written news [37]. Prior work has also observed differential effects based on news type, with political news being more sensitive to disclosure effects than lifestyle or entertainment news [39, 41]. A study [41] highlighted that AI-generated news outlets were trusted less, particularly for political news, with potential economic implications for subscription willingness. On the other hand, AI-role disclosures can be used to improve readers' short-term engagement with the article [19].

Recent work has begun exploring how disclosure design choices affect these outcomes. Studies investigating label designs for AI-generated content on social media demonstrated that more detailed labels can improve transparency perceptions [12, 18]. In contrast, detailed labels in a news context can lead to reduced trust, including on measures like willingness to subscribe [44]. Interestingly, different visual disclosure designs can lead to varying perceptions of AI contributions, even when the detail level is the same [30].

While these studies provide valuable insights into how disclosures affect trust and credibility, the attentional and cognitive load implications of AI-use disclosures remain largely unexamined. Understanding these implications is critical for designing disclosures that inform readers without imposing unnecessary processing burden, particularly given evidence that higher cognitive load can reduce information retention and increase misinformation sharing [3, 15]. One EEG study [34] found that AI-generated content labels increased attention and cognitive processing compared to unlabeled content. While this suggests that AI disclosure labels affect neural processing, the study used a single binary label (AI-generated vs. no label) and did not examine how the level of detail in disclosures or the role of AI in content production shape readers' attentional and cognitive load. We address this gap through an eye-tracking study that varies disclosure detail and AI role. Moreover, our findings contribute empirical evidence toward gaze-informed disclosure design that balances transparency with readers' cognitive and attentional load.

3 Methods

Prior work on AI disclosures has observed differential effects on trust and engagement based on the level of disclosure detail [25, 42, 44, 50], the type of news [18, 39, 41, 44], and the role of AI in content production [18, 19, 30]. This motivates investigating whether these factors similarly influence readers' attentional and cognitive load. We therefore conducted a lab study using a $3 \times 2 \times 2$ mixed factorial design to explore the impact of AI-use disclosures on readers' attentional and cognitive load. The study manipulated AI disclosure detail (between-subjects: no disclosure, one-line disclosure, detailed disclosure), news type (within-subjects: political, lifestyle), and AI role (within-subjects: editing, partial content generation).

3.1 News Articles and AI-use Disclosures

The news stimuli were based on six actual news articles (three from politics and three from lifestyle) from mainstream news organizations (e.g., BBC, CNN, NOS). Each article was manipulated using ChatGPT-4o to create two AI-assisted versions, resulting in 12 stimuli articles. AI-edited versions were generated by prompting the model to make minimal tweaks to the source article while preserving its structure: `Generate one short news article around 250-300 words based solely on the article. Make very small tweaks in the main text based on 'SOURCE'. You are allowed to change the title.` The resulting articles were minimally edited to ensure consistency with the source material. Partially AI-generated versions were created by prompting the model to generate a new 250-300 word article based on the source article: `Generate one short news article around 250-300 words, have a title, an introduction and specific text on 'TOPIC'. Make sure it is in the form of coherent article that is based on the following URL: 'SOURCE'.` The generated articles were used without additional human edits, except for deleting a paragraph when necessary to maintain the target word count.

We implemented two versions of AI disclosures: one-line and detailed, both worded to reflect AI's role (see Table 1). These statements followed existing AI transparency guidelines for journalism [6, 27, 55], and were phrased to avoid implying fully AI-generated content, as prior work indicates this can cause AI aversion [2]. Following insights from previous works [30, 36], disclosures were displayed at the top of the article as illustrated in Figure 1. The wording and presentation of the disclosures were pilot-tested with a small group (N=4) of participants.

To characterize the linguistic complexity of the stimuli, we computed two readability metrics for each article (Table 2). The Flesch Reading Ease score [17] estimates readability based on sentence length and syllable count, with higher scores indicating easier text. The Automated Readability Index (ARI) [47] estimates the grade level required to comprehend the text based on character and word counts, with lower scores indicating easier text. As shown in Table 2, AI-edited articles were consistently easier to read than partially AI-generated articles across both metrics. Lifestyle articles were also easier to read than political articles.

Table 1: AI disclosures statements used in the study.

Disclosures for AI in editing role
One-line: An AI tool was used to add an extra layer to the editing process of this article just before it was published.
Detailed: This article was produced with the assistance of an AI tool, which was used to support various stages of the editorial process, including content structuring, language refinement, and fact-check suggestions. All content was reviewed, edited, and approved by a human journalist before publication. This use of AI aligns with our commitment to transparency and responsible innovation in journalism. You can report errors at: webhomepage@news.com.
Disclosures for AI in partial content generation role
One-line: This article was largely generated with the help of an AI tool.
Detailed: This article was primarily generated using an AI tool, which was responsible for drafting substantial portions of the content, including initial story development, writing, and factual synthesis. Human editors reviewed the final version prior to publication. Our use of AI is guided by principles of transparency, accountability, and editorial oversight. Readers can report errors at: webhomepage@news.com.

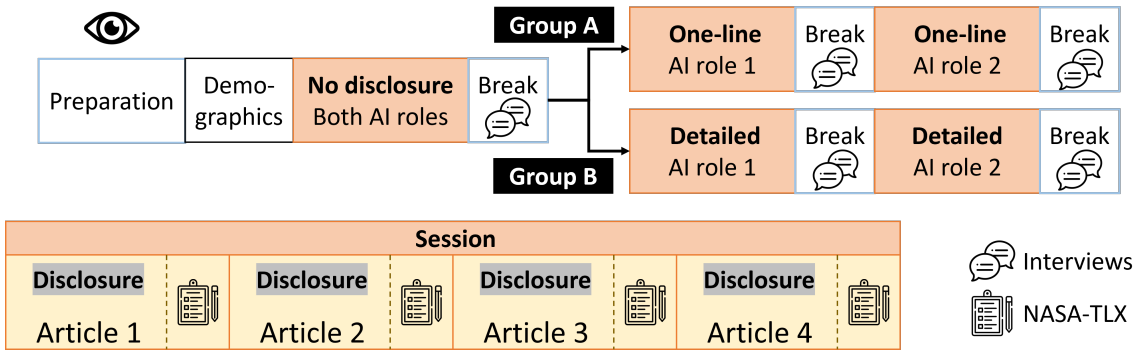

Figure 2: Overview of the experimental protocol. The eye tracker was calibrated during preparation and recorded gaze data throughout the experiment. AI role 1 and AI role 2 refer to editing and partial generation, counterbalanced across participants.

Table 2: Readability scores per article.

	AI-edited		Partially AI-generated	
	Flesch	ARI	Flesch	ARI
Political 1	27.2	17.87	23.2	17.22
Political 2	42.2	14.64	18.5	19.82
Political 3	28.3	14.22	10.0	17.79
Lifestyle 1	45.7	12.67	38.7	12.99
Lifestyle 2	47.3	13.26	29.8	16.11
Lifestyle 3	50.4	13.01	38.4	13.68
Mean	**40.18**	**14.28**	**26.43**	**16.27**

3.2 Measures

We used both subjective and objective measures to capture attentional and cognitive load in response to AI-use disclosures.

NASA-TLX: We used an adapted version of the NASA-TLX scale [22] for subjective assessment of perceived task load. NASA-TLX measures perceived load across six sub-scales, but we focused on mental demand, effort, and frustration, as our study did not involve physical exertion, time pressure, or competitive elements. Participants rated each sub-scale on a 10-point scale after reading each article, resulting in a combined score out of 30.

Eye-Tracking Measures: We recorded participants' gaze using a Tobii Pro Fusion eye tracker at a sampling rate of 120 Hz. Pupil diameter, total fixation duration, and saccade count were extracted using Tobii Pro Lab software. These metrics serve as real-time, implicit indicators of attentional and cognitive load [28, 29, 49]. As discussed in Section 2.1, pupil diameter reflects cognitive effort driven by the autonomic nervous system, while fixation duration and saccade count are more directly reflective of how attention is distributed across the news interface. Eye-tracking metrics were computed over the full reading block for each article, including both the disclosure statement and the article body. This means that the metrics capture the combined attentional and cognitive effects of reading the disclosure and subsequently processing the article, rather than isolating the disclosure reading alone.

Semi-Structured Interview: We also conducted semi-structured interviews after each session to capture participants' qualitative impressions of the articles and disclosures as well as broader perceptions of AI in journalism and disclosure preferences (details in Section 3.3). All interviews were recorded and transcribed.

3.3 Procedure

The experiment began with informed consent and a description of the news-reading interface. The eye tracker was then calibrated

using Tobii Pro Lab's five-point calibration procedure, where participants followed a series of dots on the screen. After calibration, participants accessed a locally hosted web interface where they read the news articles and completed the associated questionnaires. On the first page, they provided demographic information, including age range, gender, and news consumption habits.

As shown in Figure 2, the study was structured into three sessions (S1, S2, S3), each consisting of four articles (two political, two lifestyle). The articles for each session were assigned using a Latin square rotation, ensuring that across participants, each article appeared in all three disclosure conditions. The order of articles within each session was randomized but alternated between news types. Participants were randomly assigned to Group A or Group B. Each article sequence generated by the Latin square rotation was assigned to one participant from Group A and one from Group B, ensuring that paired participants read identical articles in identical order, with only the disclosure statements differing between them.

Both groups read articles without any AI disclosure in S1. The no-disclosure condition was always presented first so that participants would read those articles without prior knowledge of AI involvement in the study. In S2 and S3, Group A read articles with one-line disclosures while Group B read articles with detailed disclosures. Participants were instructed to read the disclosure before proceeding to the article. All articles in S2 were either AI-edited or partially AI-generated, with the alternative versions presented in S3, counterbalanced across participants.

After each article, participants completed the NASA-TLX subscales. A short semi-structured interview was conducted after each session to capture participants' immediate impressions of the articles and disclosures. After the final session, a longer interview (15–20 minutes) was conducted, which focused on broader perceptions of AI in journalism and disclosure preferences. The eye tracker recorded gaze data throughout the experiment. Each session lasted approximately 10 minutes, and the entire experiment lasted about one hour.

3.4 Participants

We recruited 40 participants from a research institute campus, including students, researchers, and non-scientific staff, and were compensated 10 Euros for their time. The study was approved by the institute's ethical committee.

During the post-session interviews, six participants (3 from each group) reported that they did not notice or read the disclosure statement in at least one session. Since these participants' responses could not be attributed to the disclosure condition, they were excluded from all further analyses, resulting in a final sample of 34 participants (21 male, 12 female, 1 non-binary). The majority of participants (n=22, 64.7%) were aged 25-34, with the remaining participants distributed across the 18-24 (n=2), 35-44 (n=4), 45-54 (n=2), 55-64 (n=2), and 65+ (n=2) age ranges. Most participants (n=19, 55.9%) reported consuming news multiple times a day, primarily via social media (38.2%) and online news websites (35.3%). Additionally, eye-tracking data were missing for four participants in some segments; these participants were excluded from eye-tracking analyses, resulting in 30 participants for eye-tracking measures.

Participants also reported high generative AI literacy (M=26.21, SD=3.19, out of 30), measured using six 5-point Likert scale items from Chan and Hu [11]. This relatively homogeneous and high AI literacy is consistent with the campus-based sample and is acknowledged as a limitation.

4 Results

We computed means and standard deviations for NASA-TLX scores and eye-tracking measures (pupil diameter, total fixation duration, saccade count). Statistical analyses using GLMs were conducted for measures where descriptive statistics indicated meaningful variation across conditions; measures with negligible variation (e.g., pupil diameter) were not tested as they would lack practical significance. Given the exploratory nature of the study, p-values were adjusted using the Benjamini-Hochberg procedure [8] to control for false discoveries due to multiple comparisons. Effect sizes (Cohen's d), z-values, and adjusted p-values are reported for pairwise comparisons.

4.1 NASA-TLX and Pupil Diameter

NASA-TLX scores (Figure 3a) were similar across disclosure conditions (no disclosure: 9.21; one-line: 8.91; detailed: 9.26). In all conditions, political articles elicited higher task load than lifestyle articles (e.g., no disclosure: political 10.37, lifestyle 8.04), consistent with expectations for hard versus soft news. Across all disclosure conditions, task load was similar for AI-edited (no disclosure: 8.94, one-line: 9.09, detailed: 9.08) and partially AI-generated (no disclosure: 9.47, one-line: 8.72, detailed: 9.44) articles. This suggests that participants did not perceive a difference in cognitive effort between reading AI-edited and partially AI-generated content, regardless of whether and how AI use was disclosed.

Similarly, pupil diameter (Figure 3b) was consistent across all disclosure conditions (no disclosure: 2.469; one-line: 2.477; detailed: 2.468). Unlike NASA-TLX, pupil diameter did not show notable differences between political and lifestyle articles (e.g., no disclosure: political 2.465, lifestyle 2.473). Disclosing AI involvement also did not affect pupil diameter (AI-edited: no disclosure 2.468, one-line 2.480, detailed 2.471; partially AI-generated: no disclosure 2.470, one-line 2.473, detailed 2.464).

►**Takeaway:** *Both subjective (NASA-TLX) and eye-tracking (pupil diameter) indicators of cognitive load showed no differences across disclosure conditions, suggesting that AI-use disclosures do not impose additional cognitive burden regardless of detail level.*

4.2 Fixation Duration and Saccade Count

Unlike cognitive load measures, fixation duration and saccade count showed promising differences across disclosure conditions.

Fixation durations (Figure 3c) were highest under one-line disclosures and lowest under detailed disclosures (no: 53.46, one-line: 58.70, detailed: 49.46). One-line disclosures led to significantly higher fixation durations compared to detailed disclosures overall ($z = 2.544$, $d = 0.34$, adjusted-$p = 0.02$). This difference was primarily driven by AI-edited articles (no: 54.40, one-line: 60.66, detailed: 47.97), where the difference between one-line and detailed conditions was significant ($z = 2.443$, $d = 0.43$, adjusted-$p = 0.02$). The effect was less pronounced for partially AI-generated articles (no:

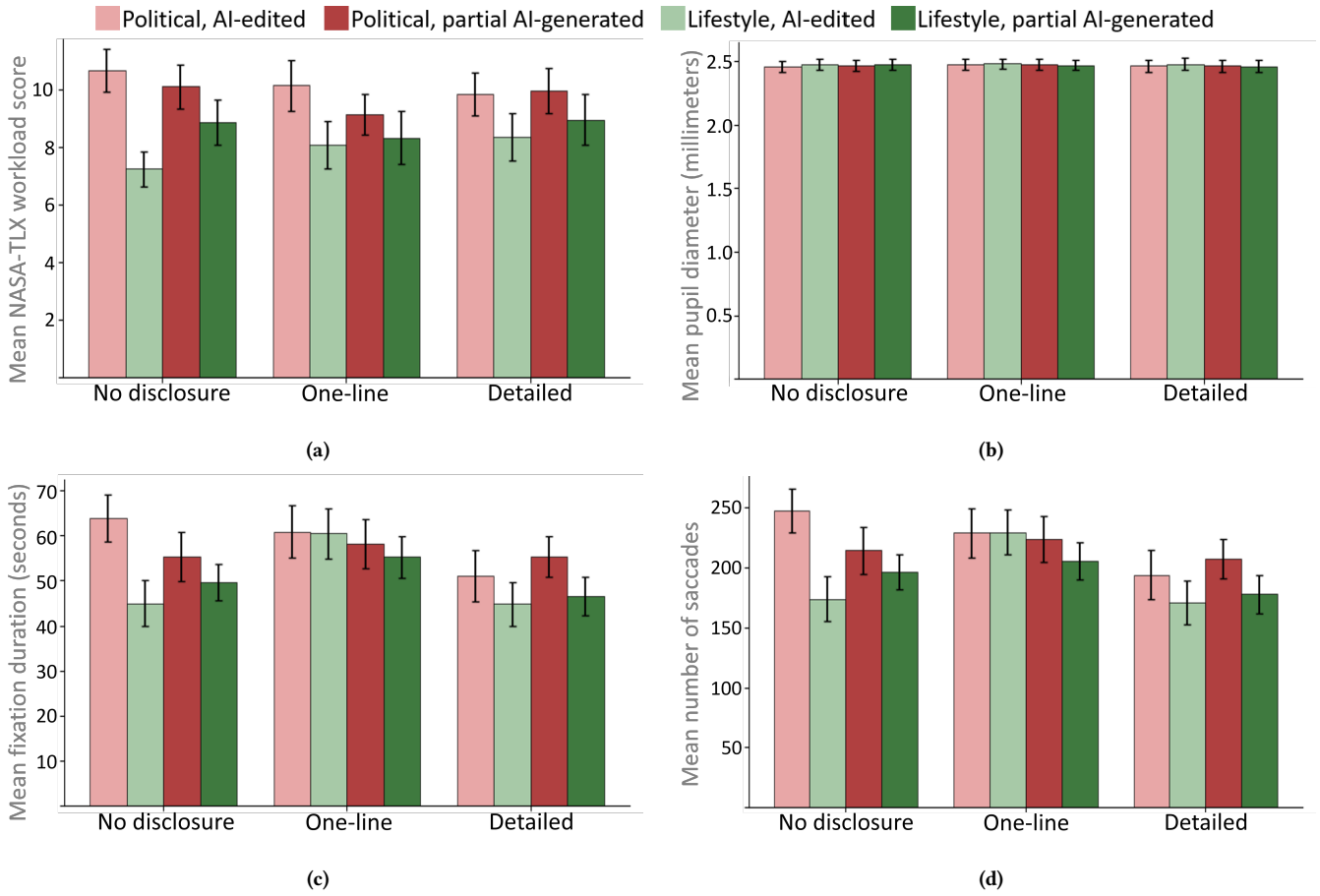

Figure 3: Grouped bar visualization of (a) mean NASA-TLX workload scores, (b) mean pupil diameter, (c) mean total fixation duration, and (d) mean saccade count across the three disclosures. Error bars represent the standard error of the mean.

52.52, one-line: 56.74, detailed: 50.95), where fixation durations were more similar across disclosure conditions. One-line disclosures also led to higher fixation durations in lifestyle articles (no: 47.32, one-line: 57.84, detailed: 45.75), with significant differences for one-line vs. no disclosure (z = 2.468, d = 0.41, adjusted-p = 0.02) and one-line vs. detailed (z = 2.120, d = 0.46, adjusted-p = 0.039). Although political articles showed higher fixation durations overall (e.g., no: political 59.60, lifestyle 47.32), the differences between disclosure conditions were less prominent for political articles compared to lifestyle articles. In other words, this difference was primarily driven by news type rather than disclosures.

Saccade counts (Figure 3d) followed an identical pattern (no: 207.63, one-line: 221.62, detailed: 187.33). One-line disclosures led to significantly higher saccade counts compared to detailed disclosures overall (z = 2.655, d = 0.35, adjusted-p = 0.02). In AI-edited articles (no: 210.33, one-line: 228.93, detailed: 182.37), the difference between one-line and detailed was significant (z = 2.518, d = 0.44, adjusted-p = 0.02). As with fixation duration, the effect was more pronounced for AI-edited articles than for partially AI-generated articles (no: 204.93, one-line: 214.30, detailed: 192.28). In lifestyle articles (no: 184.70, one-line: 217.10, detailed: 174.27), the difference was significant for one-line vs. detailed (z = 2.460, d = 0.46, adjusted-p = 0.02) and trending towards significance for one-line vs. no disclosure (z = 1.817, d = 0.35, adjusted-p = 0.069). Political articles again showed higher saccade counts overall (e.g., no disclosure: political 230.57, lifestyle 184.70), indicating greater attentional demands of hard news content rather than disclosure effects.

Given the brevity of the one-line disclosure, the observed differences in total fixation duration and saccade count cannot plausibly be attributed to repeated re-reading of the disclosure text alone and likely reflect altered processing of the article body. Furthermore, despite AI-edited articles being easier to read than partially AI-generated articles (Table 2), one-line disclosures led to significantly higher attentional load specifically for AI-edited articles, indicating that the observation stems from disclosures rather than text complexity.

►**Takeaway:** *One-line disclosures increased attentional load, as reflected in both fixation duration and saccade count, particularly for AI-edited and lifestyle articles. Detailed disclosures did not impose additional attentional burden.*

4.3 Interview Insights

The interviews covered participants' broader perceptions of AI in journalism, accountability, and disclosure designs. Interviews were analyzed inductively [9] by three coders. Here, we highlight insights relevant to interpreting the attentional and cognitive load findings.

Several participants noted that one-line disclosures were ambiguous about AI's specific role, leaving room for different interpretations. For instance, one participant remarked: "[One-line disclosure] *was a bit general, I don't know if my mother would read an article and read AI was used as a final layer, I don't know what she would understand from that. So I think it's a bit specific for our generation that knows a bit more.*" (P3). Participants also mentioned that one-line disclosures lacked information they considered important, such as assurance of human oversight and contact for reporting errors. In contrast, detailed disclosures were perceived as more transparent, giving readers a clearer picture of which steps AI was involved in. However, participants (N=6) noted that detailed disclosures were too long and likely to be skipped, drawing parallels with cookie banners and terms-of-use agreements.

When asked about their disclosure preferences, a majority of participants (N=22) preferred detailed disclosures. Interestingly, many participants (N=15) suggested ideal disclosure designs with a detail-on-demand feature, where a short disclosure is presented by default with the option to obtain more detail. This included three-fourths of the participants who preferred one-line (N=11) disclosures. Suggested designs included a clickable icon next to a short disclosure that expands into a detailed statement, and special icons or codes that link to a public detailed statement on the outlet's website explaining how AI was involved. The main reasoning was that readers should have agency over the level of disclosure and can choose depending on the topic.

Interview insights also suggest that interest and curiosity played a key role in participants' engagement with articles. Participants reported being curious about unfamiliar lifestyle topics, which may have increased engagement and arousal for those articles.

5 Discussion

5.1 Key Findings

Our findings show that AI-use disclosures influence readers' attentional load, even when cognitive load remains largely unchanged. Eye-tracking measures revealed differences across disclosure conditions that subjective measures did not capture, highlighting the value of real-time gaze sensing for evaluating disclosure designs.

One-line disclosures imposed the highest attentional burden, as reflected in both fixation duration and saccade count, an effect particularly pronounced for AI-edited articles. Detailed disclosures, despite containing more information, did not differ substantially from the no-disclosure baseline. This observation is in line with the Information-Gap theory [35]. One-line disclosures likely led to a 'gap' between readers' awareness of AI use and their understanding of its role. This unresolved uncertainty appears to have triggered active information-seeking and visual search. Interview data further support this interpretation, highlighting that one-line disclosures introduce ambiguity by omitting details about the specific steps in which AI was involved. This can cause readers to spend additional attentional effort to infer the extent and implications of AI's role, similar to a verification process [52]. In contrast, detailed disclosures appear to reduce this uncertainty by explicitly clarifying AI's role and confirming human oversight, thereby mitigating attentional demands despite presenting more text. Notably, although such eye-tracking patterns may be associated with reading complex texts [14, 26], this effect was pronounced for AI-edited articles despite their easier readability compared to partially AI-generated articles (Flesch readability: AI-edited M=40.18, AI-generated M=26.43; see Table 2), suggesting the effect is driven by disclosure ambiguity rather than text complexity.

The attentional load effect was also more pronounced for lifestyle articles than for political articles. Although political articles elicited higher fixation durations and saccade counts overall, this difference was primarily driven by news type rather than disclosures. For lifestyle articles, where baseline attentional load was lower, the introduction of a one-line disclosure produced a more noticeable increase, suggesting that attentional effects of disclosures may be more salient when readers are processing less demanding content.

NASA-TLX scores and pupil diameter showed no significant differences across conditions, suggesting that AI-use disclosures do not impose cognitive burden regardless of the level of detail. NASA-TLX scores showed higher load for political articles than lifestyle, which is in line with the readability scores. However, pupil diameters were similar across news types, which may be partly explained by interest and curiosity. Interview insights suggest that participants were less familiar with some lifestyle topics, which led to unexpected curiosity and interest. Arousal stemming from curiosity/interest may have led to higher pupil dilation [10] for lifestyle articles, narrowing the expected difference with cognitively demanding political articles.

Higher attentional load from one-line disclosures may have practical consequences beyond the reading experience. Prior work has shown that higher attentional demands during news reading can reduce information retention and increase susceptibility to misinformation [3, 15]. Whether the attentional load observed in our study translates to such downstream effects remains an open question for future work.

The divergence between attentional and cognitive load measures has methodological implications for the HCI community. Subjective measures and physiological indicators of cognitive load (pupil diameter) were not sensitive to disclosure-induced processing differences, whereas attentional measures (fixation duration, saccade count) were. This underscores that attentional and cognitive load are disconnected constructs in this context: disclosures can affect how readers distribute their attention without increasing their cognitive load. For interface designers, this means that relying solely on subjective measures or arousal indicators may miss meaningful differences in how users process transparency information.

5.2 Implications for Adaptive Disclosure Design

Our findings have direct implications for the design of adaptive disclosure interfaces. While detailed disclosures perform better on attentional measures, participants noted they would pay less attention to lengthy disclosures over time, drawing parallels with cookie banners [16, 44]. This tension motivates detail-on-demand

disclosure designs, where a concise one-line disclosure is presented by default with the option to expand for more detail [30, 50]. Gaze signals provide empirical grounding for such adaptive interfaces, potentially triggering progressive disclosures when a higher attentional load is detected. Such interfaces could dynamically balance transparency and attentional burden based on readers' real-time gaze patterns and news context.

Connecting to the broader AI disclosure literature, studies have found that more transparency or detailed disclosures tend to reduce trust [42, 44]. Taken together, these findings present a design challenge: detailed disclosures reduce attentional load but risk reducing trust, while one-line disclosures maintain trust but increase attentional load. Detail-on-demand designs, informed by real-time gaze sensing, represent a promising path to navigate this trade-off.

5.3 Limitations and Future Work

This study has several limitations. First, the participant sample was campus-based and reported high, relatively homogeneous AI literacy scores (M=26.21, SD=3.19, out of 30), which limits generalizability to broader populations with varying levels of AI familiarity. Readers with lower AI literacy may respond differently to AI-use disclosures, and future work should investigate this with more diverse samples.

Second, the lab-based setting differs from real-world news consumption, which typically involves skimming, multitasking, and reading across multiple devices, reducing ecological validity. The controlled environment may have also encouraged more careful reading than would occur naturally.

Third, the no-disclosure condition was always presented first to avoid priming participants about AI involvement. However, the disclosure detail was manipulated as a between-subjects factor. Consequently, any effects of fatigue or interface familiarity would be distributed equally across the one-line and detailed disclosure groups. Since our primary findings emerge from the comparison between these independent groups, the ordering of the baseline does not compromise the validity of the disclosure-detail effects.

Fourth, eye-tracking metrics were computed over the full reading block, including both the disclosure text and article body. While this captures the holistic user experience, future work could employ Area of Interest (AOI) analysis to isolate the specific contributions of the disclosure region. However, we note that detailed disclosures resulted in lower overall attentional load, despite containing more text than one-line disclosures. Since our findings cannot be explained by disclosure lengths, they are not a result of simple disclosure reading time.

Future work should investigate detail-on-demand disclosure designs and explore how gaze-based adaptive disclosures can be implemented and evaluated in news interfaces. Building and testing a prototype that uses real-time gaze signals to trigger progressive disclosures would be a natural next step. Additionally, longitudinal studies examining how readers' attentional responses to disclosures change with repeated exposure would provide insights into habituation effects relevant to real-world deployment.

6 Conclusion

We investigated how AI-use disclosure detail level and AI's role in content production affect readers' attentional and cognitive load through a $3 \times 2 \times 2$ mixed-factorial study, combining NASA-TLX and eye-tracking measures. One-line disclosures increased attentional load, as reflected in higher fixation durations and saccade counts, particularly for AI-edited and lifestyle articles. Detailed disclosures did not impose additional burden despite containing more information. NASA-TLX scores and pupil diameter showed no differences across conditions, indicating that cognitive load remains unchanged regardless of disclosure detail. Interview data, together with Information-Gap Theory, suggest that the ambiguity of one-line disclosures is a plausible reason for higher attentional load, while detailed disclosures reduce uncertainty by clarifying AI's role. Furthermore, participants' ideal disclosure involved detail-on-demand designs, reinforcing the need for adaptive approaches that balance transparency with attentional burden. Our findings demonstrate the value of eye-tracking for evaluating disclosure designs and motivate gaze-informed adaptive disclosure interfaces that dynamically adjust detail level based on readers' attentional patterns and news context.

Safe and Responsible Innovation Statement

This research promotes responsible innovation in AI transparency by investigating non-invasive eye-tracking measures in a controlled lab setting. No deception was used; participants were fully informed about the study before providing their written informed consent, and were free to withdraw at any time. The study was approved by an ethical committee, and all data were stored anonymously in accordance with GDPR guidelines. Potential biases due to the campus-based sample are acknowledged, and future work aims to address this through more diverse populations. In real-world applications of gaze-based adaptive disclosure interfaces, careful attention would be necessary to ensure inclusivity, prevent dark patterns, and avoid overreliance on automated disclosures.

Acknowledgment

Except for creating the news article stimuli, generative AI (Claude Opus 4.6) was used only for rephrasing parts of this manuscript.

References

[1] Ashwaq Alhargan, Neil Cooke, and Tareq Binjammaz. 2017. Multimodal affect recognition in an interactive gaming environment using eye tracking and speech signals. In *Proceedings of the 19th ACM international conference on multimodal interaction*. 479–486.

[2] Sacha Altay and Fabrizio Gilardi. 2024. People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. *PNAS nexus* 3, 10 (2024), pgae403.

[3] Oberiri Destiny Apuke, Bahiyah Omar, Elif Asude Tunca, and Celestine Verlumun Gever. 2024. Information overload and misinformation sharing behaviour of social media users: Testing the moderating role of cognitive ability. *Journal of Information Science* 50, 6 (2024), 1371–1381.

[4] Oswald Barral, Sébastien Lallé, Grigorii Guz, Alireza Iranpour, and Cristina Conati. 2020. Eye-tracking to predict user cognitive abilities and performance for user-adaptive narrative visualizations. In *Proceedings of the 2020 international conference on multimodal interaction*. 163–173.

[5] Michael Barz, Roman Bednarik, Andreas Bulling, Cristina Conati, and Daniel Sonntag. 2024. HumanEYEze 2024: Workshop on Eye Tracking for Multimodal Human-Centric Computing. In *Proceedings of the 26th International Conference on Multimodal Interaction*. 696–697.

[6] BBC. 2025. *How we're designing user-centred AI labels at the BBC.* https://www.bbc.co.uk/mediacentre/articles/user-centred-ai-labels

[7] Kim Björn Becker, Felix M Simon, and Christopher Crum. 2025. Policies in parallel? A comparative study of journalistic AI policies in 52 global news organisations. *Digital Journalism* 13, 9 (2025), 1578–1598.

[8] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57, 1 (1995), 289–300.

[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[10] Garvin Brod and Jasmin Breitwieser. 2019. Lighting the wick in the candle of learning: generating a prediction stimulates curiosity. *NPJ science of learning* 4, 1 (2019), 17.

[11] Cecilia Ka Yuk Chan and Wenjie Hu. 2023. Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International journal of educational technology in higher education* 20, 1 (2023), 43.

[12] Jingruo Chen, Tung-Yen Wang, Marie Williams, Natalia Andrea Jordan, Mingyi Shao, Linda Zhang, and Susan R Fussell. 2025. Examining the Impact of Label Detail and Content Stakes on User Perceptions of AI-Generated Images on Social Media. In *Companion Publication of the 2025 Conference on Computer-Supported Cooperative Work and Social Computing*. 270–275.

[13] Francesco Chiossi, Ekaterina R Stepanova, Benjamin Tag, Monica Perusquia-Hernandez, Alexandra Kitson, Arindam Dey, Sven Mayer, and Abdallah El Ali. 2024. PhysioCHI: Towards Best Practices for Integrating Physiological Signals in HCI. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[14] Anne-Françoise de Chambrier, Marco Pedrotti, Paolo Ruggeri, Jasinta Dewi, Myrto Atzemian, Catherine Thevenot, Catherine Martinet, and Philippe Terrier. 2023. Reading numbers is harder than reading words: An eye-tracking study. *Acta Psychologica* 237 (2023), 103942.

[15] Nicolas Debue and Cécile Van De Leemput. 2014. What does germane load mean? An empirical contribution to the cognitive load theory. *Frontiers in psychology* 5 (2014), 1099.

[16] Abdallah El Ali, Karthikeya Puttur Venkatraj, Sophie Morosoli, Laurens Naudts, Natali Helberger, and Pablo Cesar. 2024. Transparent AI disclosure obligations: Who, what, when, where, why, how. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–11.

[17] Rudolph Flesch. 1948. A new readability yardstick. *Journal of applied psychology* 32, 3 (1948), 221.

[18] Dilrukshi Gamage, Dilki Sewwandi, Min Zhang, and Arosha K Bandara. 2025. Labeling Synthetic Content: User Perceptions of Label Designs for AI-Generated Content on Social Media. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–29.

[19] Fabrizio Gilardi, Sabrina Di Lorenzo, Juri Ezzaini, Beryl Santa, Benjamin Streiff, Eric Zurfluh, and Emma Hoes. 2024. Willingness to Read AI-Generated News Is Not Driven by Their Perceived Quality. *arXiv preprint arXiv:2409.03500* (2024).

[20] Parul Gupta, Komal Chugh, Abhinav Dhall, and Ramanathan Subramanian. 2020. The eyes know it: Fakeet-an eye-tracking database to understand deepfake perception. In *Proceedings of the 2020 international conference on multimodal interaction*. 519–527.

[21] Christian Hansen, Casper Hansen, Jakob Grue Simonsen, Birger Larsen, Stephen Alstrup, and Christina Lioma. 2020. Factuality checking in news headlines with eye tracking. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*. 2013–2016.

[22] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[23] Alexander Heimerl, Pooja Prajod, Silvan Mertes, Tobias Baur, Matthias Kraus, Ailin Liu, Helen Risack, Nicolas Rohleder, Elisabeth André, and Linda Becker. 2024. The ForDigitStress dataset: A multi-modal dataset for automatic stress recognition. *IEEE transactions on affective computing* 16, 2 (2024), 1219–1234.

[24] Natali Helberger and Nicholas Diakopoulos. 2023. The European AI act and how it matters for research into AI in media and journalism. *Digital Journalism* 11, 9 (2023), 1751–1760.

[25] Mohammad Sohorab Hossain, Joshua D Clapp, and Vesna D Novak. 2025. Effects of Algorithmic Transparency on User Experience and Physiological Responses in Affect-Aware Task Adaptation. *IEEE Transactions on Affective Computing* (2025).

[26] Chaudhary Muhammad Aqdus Ilyas, Sifat-E Noor, Ashkan Tashk, Bart Cooreman, Sofie Beier, and Per Bækgaard. 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*. 1–7.

[27] Nordic AI Journalism in collaboration with Utgivarna. 2024. *AI transparency in journalism.* https://www.nordicaijournalism.com/ai-transparency

[28] Fengfeng Ke, Ruohan Liu, Zlatko Sokolikj, Ibrahim Dahlstrom-Hakki, and Maya Israel. 2024. Using eye-tracking in education: review of empirical research and technology. *Educational technology research and development* 72, 3 (2024), 1383–1418.

[29] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W Woźniak. 2023. A survey on measuring cognitive workload in human-computer interaction. *Comput. Surveys* 55, 13s (2023), 1–39.

[30] Amber Kusters, Pooja Prajod, Pablo Cesar, and Abdallah El Ali. 2026. More Human or More AI? Visualizing Human-AI Collaboration Disclosures in Journalistic News Production. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. 1–22.

[31] Matteo Lavit Nicora, Pooja Prajod, Marta Mondellini, Giovanni Tauro, Rocco Vertechy, Elisabeth André, and Matteo Malosio. 2024. Gaze detection as a social cue to initiate natural human-robot collaboration in an assembly task. *Frontiers in Robotics and AI* 11 (2024), 1394379.

[32] Robin Leuppert, Carina Weinmann, and Jan Eiden. 2025. AI-reporters as the future of journalism? Investigating recipients' credibility evaluations of AI-authored news. *Journalism* (2025), 14648849251382484.

[33] Ailin Liu, Yesmine Karoui, Fiona Draxler, Frauke Kreuter, and Francesco Chiossi. 2026. Sensing What Surveys Miss: Understanding and Personalizing Proactive LLM Support by User Modeling. *arXiv preprint arXiv:2602.00880* (2026).

[34] Yuhan Liu, Shuining Wang, and Guoming Yu. 2023. The nudging effect of AIGC labeling on users' perceptions of automated news: evidence from EEG. *Frontiers in Psychology* 14 (2023), 1277829.

[35] George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological bulletin* 116, 1 (1994), 75.

[36] Jacob A. Long, Tabitha Oyewole, Maryam Goli, Jacqueline M. Keisler, Saud Alyaqout, Michael D. Rodgers, and Arielle N'Diaye. 2025. The Disclosure Dilemma: How AI Attribution Affects Reactions to Public Health Messages. Poster presented at the 108th Annual Conference of the Association for Education in Journalism and Mass Communication (AEJMC).

[37] Chiara Longoni, Andrey Fradkin, Luca Cian, and Gordon Pennycook. 2022. News from generative artificial intelligence is believed less. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 97–106.

[38] Raphael Menges, Chandan Kumar, and Steffen Staab. 2019. Improving user experience of eye tracking-based interaction: Introspecting and adapting interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 6 (2019), 1–46.

[39] Sophie Morosoli, Emma van der Goot, Valeria Resendez, Claes de Vreese, and Natali Helberger. 2025. The Transparency Dilemma: An Experiment on How AI Disclosures Affect Credibility Perceptions and Engagement Across Topics. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Vol. 8. 1748–1757.

[40] Haruka Murakami and Tetusnari Inamura. 2025. Inferring User State from Gaze Dynamics in a VR Throwing Task: Toward Adaptive and User-Centered Rehabilitation Support. In *Companion Proceedings of the 27th International Conference on Multimodal Interaction*. 235–239.

[41] Andreas Nanz, Alice Binder, and Jörg Matthes. 2025. AI in the Newsroom: Does the Public Trust Automated Journalism and Will They Pay for It? *Journalism Studies* (2025), 1–20.

[42] Vu Minh Ngo. 2025. Balancing AI transparency: Trust, Certainty, and Adoption. *Information Development* (2025), 02666669251346124.

[43] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. 2022. The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Computing Surveys (CSUR)* 55, 3 (2022), 1–39.

[44] Pooja Prajod, Hannes Cools, Thomas Röggla, Karthikeya Puttur Venkatraj, Amber Kusters, Alia ElKattan, Pablo Cesar, and Abdallah El Ali. 2026. Full Disclosure, Less Trust? How the Level of Detail about AI Use in News Writing Affects Readers' Trust. *arXiv preprint arXiv:2601.09620* (2026).

[45] Pooja Prajod, Matteo Lavit Nicora, Matteo Malosio, and Elisabeth André. 2023. Gaze-based attention recognition for human-robot collaboration. In *Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments*. 140–147.

[46] Pooja Prajod, Matteo Lavit Nicora, Marta Mondellini, Matteo Meregalli Falerni, Rocco Vertechy, Matteo Malosio, and Elisabeth André. 2024. Flow in human-robot collaboration—multimodal analysis and perceived challenge detection in industrial scenarios. *Frontiers in Robotics and AI* 11 (2024), 1393795.

[47] Richard J Senter and Edgar A Smith. 1967. *Automated readability index.* Technical Report.

[48] Li Shi, Nilavra Bhattacharya, Anubrata Das, and Jacek Gwizdka. 2023. True or false? Cognitive load when reading COVID-19 news headlines: an eye-tracking study. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*. 107–116.

[49] Vasileios Skaramagkas, Giorgos Giannakakis, Emmanouil Ktistakis, Dimitris Manousos, Ioannis Karatzanis, Nikolaos S Tachos, Evanthia Tripoliti, Kostas Marias, Dimitrios I Fotiadis, and Manolis Tsiknakis. 2021. Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE reviews in biomedical engineering* 16 (2021), 260–277.

[50] Aaron Springer and Steve Whittaker. 2020. Progressive disclosure: When, why, and how do users want algorithmic transparency information? *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10, 4 (2020), 1–32.

[51] Ömer Sümer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, and Enkelejda Kasneci. 2021. FakeNewsPerception: An eye movement dataset on the perceived believability of news stories. *Data in brief* 35 (2021), 106909.

[52] Xin Sun, Rongjun Ma, Shu Wei, Pablo Cesar, Jos A Bosch, and Abdallah El Ali. 2025. Understanding trust toward human versus AI-generated health information through behavioral and physiological sensing. *International Journal of Human-Computer Studies* (2025), 103714.

[53] Xin Sun, Shu Wei, Ting Pan, Yajing Wang, Jos A Bosch, Isao Echizen, Abdallah El Ali, and Saku Sugawara. 2026. Eyes Can't Always Tell: Fusing Eye Tracking and User Priors for User Modeling under AI Advice Conditions. *arXiv preprint arXiv:2604.01741* (2026).

[54] Benjamin Toff and Felix M Simon. 2025. "Or they could just not use it?": The dilemma of AI disclosure for audience trust in news. *The International Journal of Press/Politics* 30, 4 (2025), 881–903.

[55] Sebastián Valenzuela, Ingrid Bachmann, Porismita Borah, and Natalia Solís Valdés. 2026. The Effects of Generative AI in News on Media Credibility and Selectivity: Evidence from a Conjoint Experiment in Chile. *Digital Journalism* (2026), 1–22.

[56] Lian A. Wu, Pooja Prajod, Abdallah El Ali, and Pablo Cesar. 2025. *Characterizing Physiological and Behavioral Responses to Human- and AI-Generated Real and Fake News.* https://sites.google.com/view/newsfutures/accepted-submissions News Futures Workshop at the 2025 CHI Conference on Human Factors in Computing Systems.

[57] Yan Zhang, Yawen Xu, Li Li, Tianyi Wen, and Jie Ding. 2024. You Can't Beat the Affects: How Emotional News Written by AI Affect the Psychophysiological Responses of Young Readers. *Available at SSRN 4950484* (2024).

[58] Jessica Zier and Nicholas Diakopoulos. 2024. Labeling AI-Generated News Content: Matching Journalist Intentions with Audience Expectations. In *Proceedings of the Computafion and Journalism Symposium 2024.*