# Media Credibility in the AI Age:

Bilinguals' Perceptual Analysis

Noga Broitman

A Thesis

in

The Department

of

Education

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Arts (Applied Linguistics) at Concordia University Montreal, Quebec, Canada

August 2024

© Noga Broitman, 2024

# **CONCORDIA UNIVERSITY**

# **School of Graduate Studies**

This is to certify that the thesis prepared

By:	Noga Broitman
Entitled:	Media Credibility in the AI Age: Bilinguals' Perceptual Analysis
1 1 %	
and submit	ted in partial fulfillment of the requirements for the degree of
	Master of Arts (Applied Linguistics)
-	ith the regulations of the University and meets the accepted standards with respect to and quality.
Signed by t	he final Examining Committee:
	Examiner
	Dr. Kim McDonough
	Supervisor
	Dr. Pavel Trofimovich
Approved b	y
	Dr. Kim McDonough, Graduate Program Director
	Dr. Pascale Sicotte, Dean of Faculty

#### **ABSTRACT**

Media Credibility in the AI Age: Bilinguals' Perceptual Analysis

# Noga Broitman

This study explores how media users determine credibility in the ever-evolving media landscape, centering the attention toward the impact language has on critical media literacy and how people interact and consider Artificial intelligence (AI)-generated content. Considering research on bilingualism describing a tendency for bilinguals to differ in their decision-making when using their first (L1) and second (L2) language, a phenomenon known as the foreign language effect (Costa et al., 2013, 2014; Keysar et al., 2012), this study examined 24 university students Spanish (L1)–English (L2) bilinguals' media literacy from the credibility perspective, with Facebook-like posts generated by ChatGPT in the two languages. Students reflected on the posts' overall credibility, sharing likelihood, and text's credibility in a mixed methods approach, through a questionnaire and a follow-up interview to elaborate on their perceptions of each post.

The findings suggest that students perceive English posts to be more credible than

Spanish posts, with students considering their ability to analyze credibility better in Spanish, due to superior intuitions in their L1 than L2. Students displayed their media literacy skills through highlighting of text elements for each post, indicating how the post's content and language changed their credibility perception of it. Students' susceptibility to deem English posts as more credible poses a serious implication for immigrants being more vulnerable to online scams.

Moreover, the results imply the existence of a confirmation bias, whereby students tend to consider the content they expect to see as more credible. This potential bias raises a concern as to people's media literacy at a time when information can be easily generated by AI for malicious

purposes (Hutchens, 2024). Supporting prior work exploring the relationship between the foreign language effect and identification of fake news (Caramancion, 2022; Muda et al., 2023), this study shows that bilinguals' reliance on their L2 experience might disadvantage them in terms of their critical media literacy skills, limiting their ability to critically engage with information online.

Apart from highlighting a major problem for media users' critical literacy skills in an increasingly bilingual world (Byers-Heinlein et al., 2019), this study illustrates a potential methodology for use as an awareness-raising intervention to foster people's media literacy in light of recent AI-focused technological advancements. This study showcases a new avenue of research on a highly important topic with relevance to society and contributes to existing literature on bilingualism.

Keywords: foreign language effect, bilingualism, media literacy, credibility perceptions, Artificial Intelligence

#### **Acknowledgments and Dedication**

I'm grateful for my Supervisor, Prof. Pavel Trofimovich for accompanying me in this learning and writing journey, highlighting room for improvement and celebrating significant milestones. Thank you for providing enlightening insights and your guidance throughout this process. I also want to thank my Prof. Kim McDonough and my peers at the Applied Linguistics Lab at Concordia for offering meaningful feedback and support. A big thank you to Melissa Rivosecchi, Concordia's Teaching & Research Librarian for the Social Sciences for helping with the literature review, formatting, and overall review of the thesis.

I want to thank the participants who were so kind to share their perceptions with me and contribute to this work. I enjoyed our conversations and your reflections, thank you for letting me learn from your experiences.

To my family and partner, friends, colleagues, peers, professors from all times, it is thanks to your contributions and your fierce belief in me that I managed to get to where I am today. Thank you for your support, your advice and perspective have made me wiser, I appreciate your love and care.

Lastly, while writing this thesis, the world has become a darker place for many, a place where humanity's cruelty rules. This thesis is dedicated wholeheartedly to those who try to cultivate hope during these dark times, thank you!

# **Table of Contents**

Chapter 1. Introduction	1
Chapter 2. Background Literature	3
Media Literacy	3
Bilingualism and Media Literacy	5
The Era of AI-Generated Content and Its Influence on Media Literacy	8
The Current Study	10
Chapter 3. Method	12
Participants	12
Materials	13
Procedure	17
Data Analysis	21
Chapter 4. Results	24
Perceptions of Credibility and Willingness to Share Posts	24
More Versus Less Credible Content	27
Chapter 5. Discussion and Implications	36
Information Credibility and the Foreign Language Effect	37
Bilingualism and Critical Media Literacy	41
Implications	43

Limitations and Recommendations	45
Chapter 6. Conclusion	47
References	49
Appendixes	64
Appendix A	64
Appendix B	68
Appendix C	70
Appendix D	71
Appendix E	72
Appendix F	78
Appendix G	84
Appendix H	85
Appendix I	86
Annendiy I	87

#### **Chapter 1. Introduction**

In his book *Media Literacy*, Potter (2021) highlights the staggering volume of information generated in our era, where data output, which doubles every year (Cisek, 2021), leaves people increasingly vulnerable to mass media influence. Potter underscores the importance of media literacy in regaining personal control, defining it as a skill that individuals must actively cultivate. Expressing a similar concern, Cook (2023) urges for the establishment of a multidisciplinary approach dedicated to the study of misinformation, to successfully address this intricate issue and especially to understand "how we evaluate information" (p. 9). Cherow-O'Leary (2023) similarly cites the need to re-evaluate human–machine relations and to establish an international coalition for media literacy. Given the major technological advancements and the need to cater to the rapidly expanding human needs under these novel circumstances, the time is ripe for a broader inspection of media literacy.

Inspired by Lazer et al.'s (2018) question, "How can we create a news ecosystem and culture that values and promotes truth?" (p. 1096), this study aims to address Cook's global challenge of understanding how we evaluate information. Even though Lazer et al.'s question is broad, it clearly identifies the task for language scholars, namely, to facilitate people's access to quality information across their multiple languages while providing the necessary tools for them to analyze it critically and independently. Consistent with this overall agenda, this study explores bilingual speakers' perception of information credibility, shedding light on their media literacy skills. This work is motivated by and expected to contribute to two distinct areas of prior research. One area concerns the foreign language effect, which refers to the finding that bilingual speakers often exhibit different patterns of reasoning and decision-making in their two languages (Costa et al., 2013, 2014; Hayakawa et al., 2017; Keysar et al., 2012). As its first objective,

therefore, this study examines whether bilinguals' perception of information credibility differs between their two languages, which would be consistent with the foreign language effect. The other area of relevant work pertains to the increasingly common use of AI in the present media landscape, where there is no longer a clear distinction between true and false (Ali et al., 2021; Arguedas & Simon, 2023). Because AI tools such as ChatGPT and Bard facilitate the generation of deceptive information (Aïmeur et al., 2023), emphasizing the need to foster people's media literacy skills, as its second objective, this study explores how bilinguals perceive credibility of AI-generated content (i.e., Facebook-like posts). Both objectives contribute to clarifying what it means to have credible information in the age of AI-generated media content.

#### **Chapter 2. Background Literature**

#### **Media Literacy**

Media literacy is notoriously difficult to define (Potter, 2022). For example, Karaduman (2015) defines this construct as "making a contribution to students' being not passive but active individuals against media" (p. 3040), thus characterizing media literacy as a skillset that shifts a consumer from a helpless position into a more proactive attitude, where the information delivered by the media is not taken for granted but is rather being questioned and dissected. According to Butler (2021), media literacy focuses on "what is considered newsworthy and how information is determined to be valid" (p. 4), which further highlights the importance of critical analysis of and reflection on media-driven data. Kellner et al. (2019) identifies six comprehensive areas of concern for critical media literacy, including social constructivism, languages/semiotics, audience/positionality, politics of representation, production/institutions, and social and environmental justice, with each area targeted through questions such as: Who are the individuals that have created this text? Why has this text been created and/or shared? Whom does this text advantages and/or disadvantages?

In a framework that discusses media literacy as tangible competences whose goal is to increase one's personal control, Potter (2021) identifies seven different skills such as analysis, evaluation, grouping, induction, deduction, synthesis, and abstraction. Of particular relevance to the present study are the key skills of analysis and evaluation. Analysis, as outlined by Potter, implies a person's ability to dissect a message into meaningful components, whereas evaluation concerns the ability to assess the value of each component, where the "judgment is made by comparing a message element to a standard" (p. 17). Therefore, for this study, media literacy is operationally defined as a person's ability to perform analysis of a text to derive its crucial

content components and to compare these components with existing knowledge to perform their evaluation.

Despite a lack of definitional agreement regarding media literacy, there is consensus that it represents a crucial skillset for people in the 21<sup>st</sup> century (Al Zou'bi, 2022; Bulger & Davidson, 2018; Butler, 2021; Chen, 2016; Kellner et al., 2019; Potter, 2022). Karaduman (2015) considers the outcome of media literacy to be people's "ability to interpret media messages, where they can distinguish between 'the reality' and 'the reality presented by media'" (p. 3040). With media punctuating almost every possible aspect of a person's daily life (De Zengotita, 2006; Hanson, 2022), it is of critical importance for consumers of information to be able to distinguish fake from real, by asking uncompromising, critical questions and being active yet careful media users.

There is a growing body of evidence suggesting that the critical media literacy skills of analysis and evaluation can be developed through dedicated instructional interventions. Al Zoui'bi (2022), for example, asked 100 undergraduate students from a Jordan university to judge 10 news titles (stemming from publications on Facebook, Instagram, Twitter, and other news sites) as either true or false, before and after these students received media and information literacy training. Following the training, the students showed a greater frequency and range of approaches that they used to identify fake news. In another recent study, Moore and Hancock (2022) asked participants to determine the credibility of 12 news headlines (6 real and 6 false) and to indicate, both at the pretest and posttest, whether they carried out any research to determine each headline's credibility. While the control group displayed rather modest improvements, the experimental group that completed an hour-long training module addressing

best practices for media literacy not only reliably distinguished between false and fake headlines but also reported to have conducted research before determining their credibility.

While targeted interventions are generally effective in improving people's media literacy, there is presently little assessment of people's media literacy skills. Understanding people's existing knowledge about media literacy, including their beliefs about news credibility, is crucial for the design and development of future interventions whose goal is to equip consumers of information with the necessary knowledge to foster their media literacy. A comprehensive understanding of media literacy thus requires a further examination of people's use of these skills in action, as people are confronted with information in real time.

# **Bilingualism and Media Literacy**

One way to examine people's media literacy in action is to explore it across a speaker's one or more languages, for instance, by observing the behavior of bilinguals when exposed to similar content in their two languages. The idea of potential differences in bilinguals' experience with media literacy in their two languages comes from the literature on the foreign language effect (Keysar et al., 2012). This effect describes differences in people's behavior and decision-making in their first language (L1), which is typically their earlier learned and more proficient language, versus their second language (L2), which is a language often learned later in life and to a lower degree of proficiency. Compared to when they use their L1, bilinguals using their L2 have been shown to engage in different behaviors, responding in distinct ways to essentially thesame scenario. In an early study, Keysar et al. (2012) presented bilinguals with a problem, asking them to choose a treatment for a disease endangering the lives of 600,000 people.

Bilinguals needed to choose between a safer (200,000 people will be saved) and a riskier (33.3% chance that all 600,000 will be saved, and a 66.6% chance that nobody will be saved), where the same dilemma was framed either in terms of potential gains (i.e., 200,000 people will be saved)

or potential losses (i.e., 400,000 people will die). Keysar et al. (2012) tested three groups of bilinguals (English–Japanese, Korean–English, English–French), and all displayed a similar pattern of asymmetry in their L1 decision-making. If the dilemma was presented in terms of gains, they would choose the safer option; but if it was presented in terms losses, they would more likely opt for the riskier alternative. In their L2, however, the results were overall symmetrical, where bilinguals opted for the two options at similar rates, regardless of how they were framed, demonstrating a reduced sensitivity to the framing bias (i.e., decision-making based on the options presented as positive or negative).

In a similar vein, Costa et al. (2014) investigated the potential role of language in bilinguals facing moral decisions. These researchers presented a diverse sample of bilinguals (English–Spanish, Korean–English, English–French, Spanish–English/Hebrew) with the footbridge dilemma, in which a fast train would kill five people stuck on the tracks unless one chose to sacrifice a single bystander to save the other five. Compared to the L1 scenario, the L2 scenario led bilinguals to more frequently opt for the utilitarian choice (i.e., saving five individuals by sacrificing one). In a further replication of this study with two groups of counterbalanced bilinguals (English–Spanish, Spanish–English), Costa et al. (2014) similarly showed that 18% of their participants opted for the utilitarian choice in the L1 scenario but 44% opted for the same option in the L2 scenario. In essence, when experiencing a moral dilemma, bilinguals tended to more frequently choose a more utilitarian response in their L2 than in their L1.

Despite some nuanced findings (Białek et al., 2020, 2022; Dylman & Champoux-Larsson, 2020), the foreign language effect is a robust phenomenon, whereby bilinguals tend to perceive threat as more beneficial than risky in their L2 than their L1 (e.g., Hadijchristidis et al., 2015; Keysar et al., 2012), consider themselves less susceptible to bias when tested in their L2

than their L1 (e.g., Niszczota et al., 2022), and demonstrate less tendency to treat two events as causally related in their L2 than in their L1 (e.g., Diaz-Lago & Matute, 2019). In fact, a recent meta-analysis of 180 effect sizes shows that bilinguals' tendency to opt for utilitarian choices when faced with a personal moral dilemma, such as pushing a person off the bridge to save others, is 1.64 times greater in their L2 than their L1 (Stankovic et al., 2022).

In recent years, several researchers have examined whether the foreign language effect could extend to people's ability to detect fake news. For instance, Caramancion (2022) investigated whether L1 and L2 speakers of English could detect disinformation in real versus fake news headlines which were either textual (e.g., plain text with no images) or visual (e.g., text along with an illustration). Compared to L2 speakers, L1 speakers were more accurate at distinguishing fake from real content, regardless of whether it was visual or textual, although both groups did not reach 100% detection rate. In another study, Fernández-López and Perea (2020) investigated bilinguals' susceptibility to fake news in their two languages. Whereas perceptions of news credibility did not depend on the language in which the information was presented, the emotionality of news content emerged as a key factor. The extracts of fake news were rated as more credible when the content contained high negative emotionality (e.g., a title discussing gender-based violence), regardless of the content's language. Finally, Muda et al. (2023) investigated the degree of credibility of news headlines presented to Polish–English bilinguals in their two languages. Bilinguals overall classified fake headlines as less credible than real ones, but distinguishing fake from real news was harder for bilinguals in their L2 than their L1. Furthermore, the credibility of news headlines was associated with the emotionality of their content, where more arousing news titles were considered more credible, regardless of whether they were presented in the bilinguals' L1 or L2.

Thus, with respect to bilinguals' ability to detect fake news, the role of language is unclear. In some situations, the use of L2 possibly impedes bilinguals' ability to detect fake news, compared to when they use their L1 (Caramancion, 2022; Muda et al., 2023). In other situations, however, language might not matter as much, and news credibility might be determined by the perceived emotionality of its content (Fernández-López & Perea, 2020).

### The Era of AI-Generated Content and Its Influence on Media Literacy

With AI now becoming ever more influential in people's lives (Kneusel, 2023), the alarming issue of fake news, which is easily generated through AI algorithms (Aïmeur et al., 2023), poses a particularly serious concern for governments and societies worldwide (Lazer et al., 2018), especially with respect to educating young generations on how to evaluate information online (McGrew et al., 2019). Considering that fake news is spreading faster than real news online (Vicario et al., 2016; Vosoughi et al., 2018) and that media consumers often do not fully read the content of media posts (Gabielkov et al., 2016; Pennycook et al., 2021), it is of key importance to equip people with the right tools to analyze online content and discern fake from real news.

The rise of AI also entails a significant challenge for language scholars, particularly in light of their past work identifying and describing the linguistic elements of fake news and fraudulent messaging (Anesa, 2022; Grieve, 2007; Grieve & Woodfield, 2023; Hancock et al., 2007; Zhou & Zafarani, 2020). Previously, real and fake content could be distinguished through a linguistic analysis of text. For instance, fake news created by humans tends to feature more negative emotionally laden words than real news (Long et al., 2017; Newman et al., 2003; Zhou et al., 2021), and fake news titles generated by people often contain simpler and more personalized language (e.g., fewer technical words, more frequent pronouns like *you*) as opposed to real titles (Horne & Adali, 2017). However, these linguistic metrics have become largely

irrelevant in the era of AI-driven content. The ability of AI to generate, in a matter of seconds, a sophisticated text conveying the required message in the most appealing way leaves people with very few opportunities to productively assess the likelihood of this information being misleading (Cardenuto et al., 2023; Kreps et al., 2022), especially in a climate where hoaxes, scams, and disinformation can be generated and spread faster than ever before (Blauth et al., 2022).

With the rise of popularity of AI-generated information, researchers have turned to investigating how people perceive this content. Labajová (2023), for instance, explored participants' live interaction with AI-generated content, without alerting them to the fact that it was created by AI. Of 100 participants, 87 reported that they were unable to differentiate between human and AI-generated texts, emphasizing the concern that the two are virtually indistinguishable and calling for the development and refinement of people's critical media literacy skills. Focusing on credibility of both human- and AI-generated texts, Huschens et al. (2023) found that their participants attributed the same degree of credibility to both types of content, while finding the AI-generated texts more engaging (see also Kreps et al., 2022). Taken together, these findings highlight the importance of more accessible and informative training to support critical consumption of AI-based information (Cook, 2023; Geeng et al., 2020; Stamboliev, 2023; Wu et al., 2020), to reinforce society's immunity to manipulation or deception. While there is abundant literature on how readers can benefit from media literacy training to enhance their ability to detect fake news (Al Zou'bi, 2022; Ford et al., 2023; Murrock et al., 2018), more insight is required to promote the understanding of how readers perceive this information in the era of AI, where texts can be generated to perfection instantly by automatic tools like ChatGPT.

#### **The Current Study**

The current rate of information generation is faster than ever before (Cisek, 2021; Potter, 2021), and this data climate serves as a particularly fruitful ground for misinformation (Bermes, 2021), allowing it to spread faster and more broadly (Vosoughi et al., 2018). It is of key importance, therefore, to understand how people—and especially bilinguals and multilinguals who represent the lion's share of today's highly mobile, multilingual, and multicultural world (Byers-Heinlein et al., 2019)—determine credibility of online content and what strategies they employ in doing so. This study investigated these issues in bilingual Spanish–English speakers, comparing their reactions to AI-generated online content in their two languages (Costa et al., 2014, 2019; Keysar et al., 2012). The study employed a mixed-methods design, with the quantitative phase followed by the qualitative phase. In the quantitative phase, bilinguals were asked to evaluate credibility of AI-generated Facebook-like posts and to highlight elements that they consider to be more versus less credible. They were also asked to indicate their willingness to share each post, as a measure of potential future behavior reflecting their credibility judgment (Majerczak & Strzelecki, 2022). In the qualitative phase, bilinguals were interviewed using two randomly selected posts as prompts (one in each language), to elicit their specific reasons for marking post elements as more versus less credible, with additional questions broadly targeting bilinguals' general media literacy. The study's overall goal was to understand the skillset of bilingual media consumers, exploring the role of language in their credibility judgements.

The present study extends previous work on media literacy and the foreign language effect by targeting AI-generated materials. Previous research has employed titles or brief text descriptions (as common in social media posts) to examine participants' reactions (Caramancion, 2022; Muda et al., 2023; Pennycook et al., 2021; Preston et al., 2021), with the target content designated as "fake" versus "real" based on verification checks through designated online

platforms such as Snopes (Pennycook et al., 2021). In contrast, the use of ChatGPT in this study to generate the target materials not only obviates the possibility that participants have previously been exposed to online content but also reflects the current reality, where content can be easily generated using automatic tools, facilitating online deception and spread of fake information. Thus, eliciting bilinguals' judgments of credibility for AI-generated content has the potential to uncover strategies that they use to discern information that is credible versus fraudulent. This study was guided by the following two questions:

- 1. Do Spanish–English bilinguals differ in their perceptions of credibility of AI-generated posts in their L1 (Spanish) versus their L2 (English)?
- 2. Which elements in AI-generated posts do Spanish–English bilinguals perceive as credible versus suspicious and do these perceptions differ across posts in L1 Spanish versus L2 English?

#### Chapter 3. Method

### **Participants**

The participants included 24 international university students (14 female, 10 male) who grew up in monolingual Spanish households in Colombia (7), Mexico (6), Dominican Republic (3), Ecuador (2), Argentina, Bolivia, Chile, El Salvador, Paraguay, and Peru (1 from each). All had been exposed to Spanish from birth and reported it as their L1. The students ( $M_{age} = 26.33$ years, SD = 6.37, range = 19–44) had resided in Canada for less than 5 years and could thus be considered newcomers to the country. When asked about how many years they had studied English, 17 students responded "more than 10 years," three students selected "5–10 years," and the remaining four indicated "3–5 years," suggesting that in their majority they were experienced L2 speakers. At the time of the study, they were pursuing undergraduate or graduate degree programs in English-medium universities (Concordia University, McGill University) in Montreal, a French-English bilingual city in the French-speaking province of Quebec. Both universities require students to obtain a minimum IELTS score of 6.0 (with no component under 5.5) for admission to an undergraduate program and a minimum score of 6.5 (or equivalent) for graduate studies. To ensure that the students had sufficient L2 English skills to understand the study materials and to address recent calls for a clearer reporting of participants' language skills in research on the foreign language effect (Del Maschio et al., 2022; Purpuri et al., 2024; Teiltelbaum-Dorfman et al., 2024), all students reported their IELTS scores (or equivalent) submitted for university admission. Two students did not report their exam results as they had provided proof of previous academic studies in English. Of the 22 remaining students, 10 took the Duolingo English test (M = 126.2, equivalent to an IELTS 7.0), eight took the TOEFL test (M= 102.37, equivalent to an IELTS 7.5), and four took the IELTS test (M = 7.5). Finally, when asked to describe their English skills (as basic, intermediate, or advanced), all 24 students

indicated their level of English to be advanced. Thus, all students overall reached an IELTS score of 7.0 (or above) and considered themselves advanced L2 English speakers.

The students reported primarily using English (51%) and Spanish (39%) daily; their daily use of French (9%) and other languages (1%) was low. As for their daily use of Facebook, 19 students (79%) reported using Facebook for less than an hour per day, while the remaining 21% used it for 1–2 hours daily. Facebook use was nearly equally split between English (46%) and Spanish (45%), whereas French (8%) and other languages (1%) were used less commonly on this platform. The most frequently reported functions of Facebook involved maintaining contact with family and friends (30%), sharing photos, videos, and stories (23%), and finding a place to live (16%). Less frequent uses of Facebook related to watching or reading the news (15%), looking for service providers (12%), and searching for jobs (4%).

#### **Materials**

#### **Experimental Materials**

Based on the recent data from the Canadian Anti-Fraud Centre (CAFC), two of the 10 most common frauds (CAFC, 2021, p. 8) were targeted in this study—employment and rental frauds. These two types of fraud also illustrate some of the most common content areas of concern for newly arrived students, considering that they often require housing and part- or full-time employment to support their studies. The target materials included two texts (henceforth, posts), one in English and one in Spanish, targeting each of the two frauds (for a total of four posts). In addition, two extra posts (one in each language) were created for practice, to familiarize the students with the experimental tasks. The posts were between 150–250 words in length, and all were checked through English (https://www.lextutor.ca/vp/eng) and Spanish (https://www.multilingprofiler.net) vocabulary profilers to ensure that they contain over 70% of the 2,000 most frequently used words in each language (see Table 1 for descriptive statistics). All

posts were created using the ChatGPT AI model developed by OpenAI (OpenAI, 2023) (see Appendixes A and B for the prompts used to generate each post). Released in November 2022 for public use, ChatGPT represents the latest evolution of language model research, allowing machine learning algorithms to analyze, interpret, and reproduce human language (Heaven, 2023).

**Table 1**Descriptive Statistics for the Target Posts

	Employm	ent posts	Rental posts		
Text statistic	Spanish	English	Spanish	English	
Length (words)	177	189	187	203	
1 K words (%)	119 (67.2%)	119 (63.0%)	127 (68.0%)	143 (70.4%)	
2 K words (%)	13 (7.4%)	19 (10.1%)	21 (11.2%)	13 (6.4%)	
3 AWL words (%)	6 (3.9%)	14 (7.4%)	9 (4.9%)	17 (8.4%)	

*Note.* AWL = academic word list.

To determine that the posts were comparable across the fraud types and languages, they were presented to seven L1 English speakers and six L1 Spanish speakers for prerating. They were asked to use a 7-point scale evaluating each post for representativeness ("How representative is the previous post of a **typical Facebook job posting**?"), with endpoints labeled as "not representative at all" and "highly representative," and for quality ("Would you say this is a good quality post?"), with endpoints labeled as "poor quality" and "high quality." As an additional experimental check of materials quality, the study participants (i.e., 24 students) also

responded to the same questions about each post's representativeness and quality, using the same scales. In terms of language differences, as shown in Table 2, the posts were generally comparable in representativeness and quality between Spanish and English, although the study participants tended to consider English posts to be of slightly better quality than Spanish posts. In terms of differences in post content (see Appendix C for statistical comparisons), however, the rental posts consistently elicited higher representativeness and quality ratings from the study participants than the employment posts within each language.

Table 2

Descriptive Statistics (Medians and Interquartile Ranges) and Comparisons (Wilcoxon Tests) of Representativeness and Quality for the Four Posts

	Employment posts		Comparison		Rental posts		Comparison	
Rater	Spanish	English	Z	p	Spanish	English	Z	p
	Representativeness							
Preraters	3.5 (3.00)	5.0 (2.00)	1.83	.067	4.5 (5.25)	6.0 (1.00)	1.10	.272
Participants	4.5 (3.00)	5.0 (3.25)	1.08	.280	6.0 (2.00)	7.0 (2.00)	0.66	.509
	Quality							
Preraters	4.5 (2.25)	6.0 (2.00)	0.67	.501	5.5 (2.50)	6.0 (0.00)	1.09	.275
Participants	4.5 (3.75)	5.0 (3.75)	2.30	.021	5.0 (1.00)	6.0 (2.00)	2.02	.044

In the experimental materials, each post was accompanied by a set of three main questions. The first question targeted the students' perception of the post's credibility ("How credible is the previous post?" ¿Qué tan creíble es la publicación anterior?). It required the students to indicate how credible they found the post (following Muda et al., 2023; Sousa &

Bates, 2021) using a 7-point scale, where 1 designates "extremely noncredible" (extremadamente increible) and 7 denotes "extremely credible" (extremadamente creible). The second question targeted the students' willingness to share each post ("Would you share this post on Facebook?" ¿Compartirias esta publicación en Facebook?). They used a similar 7-point scale, where 1 meant "not at all" (definitivamente no) and 7 meant "sure" (definitivamente si). The final question requested the students to highlight the specific elements (individual words or phrases) that make the post seem credible and noncredible to them, with each highlighted element colored either in green ("more credible," mas creible) or in red ("less credible," menos creible). The students' responses to this question were used as prompts during the interview phase of the experiment to elicit the reasoning underlying their credibility judgements and decisions.

#### **Interview Materials**

The post-task semi-structured interview (Richards, 2009) was guided by prompting questions derived from the study's research questions (Kvale, 2007). The interview set, which consisted of 10 questions (see Appendix D), followed the trunk-and-branch pattern outlined by Rubin and Rubin (2012), where two main questions (i.e., "What did you think about the highlighting of more credible and less credible parts?" ¿Que pensaron sobre la parte donde tenían que resaltar lo que les resulto mas creible o menos creible? and "In your daily use of social media, how do you decide if something is more or less credible?" En su uso diario de las redes sociales, ¿como defines si algo es mas creible o menos credible?) served as a branch intended to scaffold follow-up questions (e.g., "How did you distinguish between what you considered more or less credible?" ¿Como hicieron la distinción entre lo que consideraron mas o menos creible? and "Do you judge credibility differently in English versus Spanish?" ¿Califica la credibilidad de manera diferenta en ingles que en Espanol?) or probes (e.g., "If you were to

see the same post in English and Spanish, do you think you would judge their credibility differently depending on the language?" Si vieras la misma publicación en inglés y español, ¿crees que juzgarías su credibilidad de manera diferente según el idioma?) to elicit rich, nuanced answers from the students.

# Background Questionnaire

A background questionnaire (available in Appendix E) elicited information about the students' history of language learning and use (e.g., daily use of Spanish, English, and French), country of origin, date of arrival and length of residence in Montreal, formal education, and use of Facebook. The students indicated approximately how much of the content that they encountered on social media was in L1 Spanish versus L2 English or other languages and reported reasons for their Facebook use (e.g., looking for a job, keeping in touch with friends, etc.). The questionnaire's goal was to obtain a general profile of the students' language use, and to determine if it broadly plays a role in how they classify the posts.

All experimental materials were piloted with two Spanish–English bilinguals to ensure that the instructions were clear, that the experimental procedure ran smoothly, and that it did not exceed a total of 1 hour completion time, including the interview phase.

#### **Procedure**

All testing took place in an individually scheduled 1-hour session in a quiet research space at Concordia University, using a laptop to complete the tasks. The study was presented to the students as an experiment focusing on people's perceived credibility of Facebook posts, to ensure that they believed that the materials included actual Facebook posts. The researcher first shared with each student a brief presentation about the experiment, introducing the three main response categories following each post, while emphasizing that posts could appear in both English and Spanish (see presentation in Appendix F). The students were then asked to read and

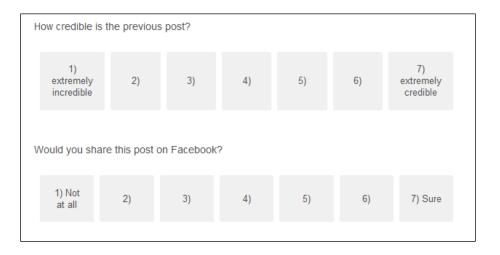
sign a consent form, in either Spanish or English depending on their preference, at which point any questions about their participation in the study and their right to withdraw from it were addressed (2 minutes).

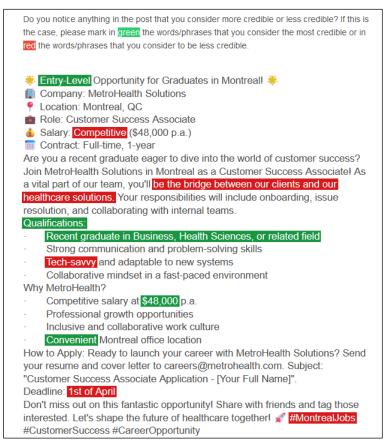
Once the students signed the consent form, they were taken to a separate room to complete the experimental tasks; the researcher sat in a different room to limit any potential effects of the researcher's presence on the students' performance. Using the laptop, the students accessed an online Qualtrics interface and proceeded first to the practice phase, followed by the experimental phase. For practice, they read a sample post in Spanish, followed by the three main questions and two additional questions focusing on post representativeness and quality (used to establish the quality of research materials). The students first read the instructions (Appendix G), before reading the post and responding to it; they then completed another practice post in English following the same procedure to ensure that they understood the task's requirements and practiced the procedure in both languages (3–5 minutes). If the students had no further questions about the procedure, they proceeded to the main task, where they responded to each of the four posts separately, first completing the Likert scale to assess each post's credibility, then evaluating the likelihood of sharing the post through Facebook, and finally highlighting the relevant sections in the post in green versus red, depending on whether they considered those as credible or suspicious (approximately 4–5 minutes per post, for a total of 16–20 minutes). The order of the posts was randomized, to resemble the way Facebook posts appear in real life, namely, controlled by the platform's algorithm. As the experiment included four posts, there were 24 possible random order combinations, ensuring that each student saw the four posts in a unique order. Each post appeared in isolation, and the students had unlimited time to read through and respond to it (see Figure 1 for a screenshot of the response interface). Upon completion of the

experimental phase, the students completed the background questionnaire using an online Qualtrics interface (6–8 minutes).

Figure 1

Three Main Questions Following Each Post





After completing the experimental task, the students engaged in an audio-recorded interview in their L1 (Spanish), to review their decisions to highlight certain text elements as credible versus suspicious (approximately 5–8 minutes discussion per post, for a total of 10–16 minutes of interview per student). Prior to the experiment, the researcher had generated a list containing two randomly chosen posts (one in Spanish, the other in English) for each student do discuss. During the interview, a screenshot of the students' actual responses to each of these two posts was shared with them on a PowerPoint slide and was used to guide the interview. The interview generally followed the structure outlined in Appendix D, so all students were asked the main 10 questions; however, several questions arose spontaneously as follow-up prompts, meaning that some interviews featured additional questions.

At the end of the testing session, the students were debriefed about the study's main goal, which was to understand what information bilingual readers identify as more or less credible, focusing in particular on any differences in their credibility judgements between Spanish and English (4–5 minutes). The purpose of the debrief was to provide the students with tools that might prevent them from falling victim to fake news, scams, and frauds, which are common in Canada (CAFC, n.d.). For the debrief, the researcher introduced the students to *The Little Black Book of Scams* in Spanish (Appendix H) and made the link to this resource available to them through email (see the email sent to participants in Appendix I and debriefing script in Appendix J). To illustrate the extent and impact of various types of frauds in Canada, the researcher also shared the periodically updated infographics from the Canadian Anti-Fraud Centre (see Figure 2).

Figure 2

Impact of Fraud in Canada: A Sample Infographic from the Canadian Anti-Fraud Centre



# **Data Analysis**

This study adopts the convergent parallel design (Creswell & Plano-Clark, 2011), whereby both the quantitative and qualitative data were collected and analyzed in parallel to support the interpretation of findings. The quantitative data include several measures derived from the students' responses to each post. The students' ratings of each post's credibility and their willingness to share each post are numeric values (1–7 scale), tabulated separately for the English and Spanish posts. The text elements highlighted by the students as more versus less credible were also tabulated, again separately for the English and Spanish posts, recording the frequency of each marked category (more vs. less credible) and the proportion of the marked lexical content per category (in words) out of the total word content in each post. Because the comparisons of the preraters' and the students' evaluations of each post's representativeness and quality revealed a difference in favor of the rental posts (see Appendix C), the rental and employment posts were analyzed separately (i.e., without averaging numeric responses for the

rental and employment posts in each language). All quantitative data were checked for the assumption of normality through visual inspection (Q–Q plots) and tests (Shapiro-Wilk). Most distributions revealed a negative skew, reflecting the students' tendency to provide positive assessments. Because numeric data were nonnormally distributed, all subsequent quantitative analysis were carried out using non-parametric statistics.

The qualitative data were derived from the recordings of interviews, which were first transcribed using TurboScribe (<a href="https://turboscribe.ai">https://turboscribe.ai</a>), then verified by the researcher along with the recording to correct transcription errors and misspellings. The content of each interview was analyzed using MAXQDA, which is a software application dedicated to qualitative analysis, with the students' comments about more versus less credible aspects of each post (coded separately for the English and Spanish posts) analyzed through bottom-up, iterative coding. For the first research question, the qualitative analysis focused on the students' description of potential differences in their credibility judgements in the two languages they speak. For the second research question, the analysis focused on the students' justification to select some post elements as more versus less credible. The purpose of qualitative analysis was to provide further insights into the students' perception of credibility through nuanced, detailed, first-person narratives.

Data coding followed open coding, axial coding, and selective coding, according to the grounded theory (Strauss & Corbin, 1998). During the first phase, an initial list of recurrent subcategories of reasons was created, using the students' verbatim descriptions. During the axial coding phase, related discrete-point subcategories were combined into more general thematic categories. Finally, during the selective coding, in the final step of the coding process, general categories were grouped (data permitting) under broader themes. To establish the trustworthiness of the thematic analysis (Nowell et al., 2017), another coder was asked to recode all data

following Braun and Clarke's (2006) concept of inductive analysis (i.e., preforming qualitative analysis with preexisting criteria), where the coder confirmed (or rejected) the recurring themes identified in the initial analysis. After the second coding was completed, both coders discussed all final coded categories and reached consensus on their composition through discussion. The results of the qualitative coding are presented descriptively, in relation to each research question, contrasting the students' responses to the Spanish versus English posts. All direct quotes reported in the Results sections are close translations from Spanish by the researcher.

#### **Chapter 4. Results**

### **Perceptions of Credibility and Willingness to Share Posts**

The first research question asked if Spanish–English bilinguals differ in their perceptions of credibility of AI-generated posts in their L1 versus L2. In terms of post credibility (summarized in Table 3), the students generally provided high ratings for both post types in Spanish and English, where the median values of the ratings reached or exceeded 5 on a 7-point scale. As shown in Table 3, Wilcoxon tests indicated that the students rated the credibility of both the rental and the employment post higher in English than in Spanish, with moderate to strong effect sizes (*r*), according to Cohen (1988).

Table 3

Descriptive Statistics (Medians and Interquartile Ranges) and Comparisons (Wilcoxon Tests) of

Credibility Ratings for the Employment and Rental Posts

	Spanish		English		Comparison		
Post type	Mdn	IQR	Mdn	IQR	Z	p	r
Employment	5.0	1.75	6.0	3.00	-2.51	.012	51
Rental	5.0	1.00	6.0	2.00	-2.00	.045	41

In terms of the students' willingness to share the posts (summarized in Table 4), the ratings were generally lower than the credibility assessments, with the median values around the scale midpoint. According to Wilcoxon tests, the students did not differ in their willingness to share either the employment or the rental post between Spanish and English, although the ratings tended to be higher for the English than the Spanish posts. To summarize, whereas the students

rated both types of English posts as more credible than Spanish posts, they did not differ in their rated willingness to share those posts between English and Spanish.

**Table 4**Descriptive Statistics (Medians and Interquartile Ranges) and Comparisons (Wilcoxon Tests) of Willingness to Share Ratings for the Employment and Rental Posts

	Spanish		English		Comparison		
Post type	Mdn	IQR	Mdn	IQR	$\overline{Z}$	p	r
Employment	2.0	2.50	3.0	4.00	-1.75	.080	36
Rental	4.0	3.00	4.0	3.00	-1.53	.126	31

These numeric patterns were reflected in the students' interview responses, where they generally considered information in English to be more credible than in Spanish with several reasons for this perceived difference. Some students, like P12, believed that more misinformation generally appears in Spanish than in English: "this might be a result of my upbringings in the Dominican Republic, a lot of the information I saw online was false, or false publications, so I'm more cautious when reading posts in Spanish." Others claimed that Spanish posts often target immigrants and therefore serve to spread fake information and take advantage of the people who are not proficient in either English or French, for instance, as explained by P15:

I already saw it in close cases, friends, that there are people who tell you "Ah, we are going to help you," and in the end they do not help you, on the contrary, they end up harming you, and they also take advantage of the fact that you do not speak English or French.

According to other students, Spanish is relatively uncommon in Montreal, where English and French are used frequently in the community, so posts in these languages are more prevalent and thus more credible: "because here everyone speaks English or French, so it is more difficult to deceive [people] that way" (P15). In fact, for some students, the use of Spanish in the sociolinguistic context of Montreal is already enough to create suspicion, as explained by P35: "And when I see an advertisement for something in Spanish in Montreal, I kind of start thinking, 'Well, why are they publishing something in Spanish, knowing that it is not a language [of the community]?" Finally, for some students, such as P18, any language other than Spanish is considered "superior" and therefore more credible: "I think the fact that it is in another language makes me believe it more... for the simple fact of thinking that perhaps it is a language [that is] in quotes 'superior' or something like that."

In terms of the information and skills that the students considered in their judgements, they generally attributed their stronger ability to judge credibility in Spanish to their extended familiarity and experience with it. For example, P18 spoke of how aware he is to how "Spanish or Latin American speakers sell smoke," and P16 claimed to be able to perform "a better analysis of [a post], simply because it is in [my] native language." Similarly, P25 claimed that "[i]n Spanish I already know that... they want to deceive me... in English let's say that I already recognize a little more but I feel more comfortable in Spanish than in English." A common thread emerging from the students' comments was that using their L1 allowed them to perform a deeper analysis of a post, as illustrated by P32, for whom the use of Spanish helped to "know the intentions [of the writer] by just reading a post," and by P19, for whom the use of Spanish triggered critical questions such as "I would ask myself, 'Why is this in Spanish? Who is this directed towards?""

Apart from language, the students also brought up various contextual and cultural variables which might determine a post's credibility. As explained by P20, "culture can play an important role, because if I see these types of publications in my country... I may be a little more critical than when I am in a different place, be it, for example, in England or here [Montreal]." For other students, like P23, language makes no difference in terms of credibility: "If I see like three posts that speak the same thing in three languages, I don't feel like I would change anything." A similar sentiment, for instance, was shared by P21, for whom other contextual details take precedence in credibility: "I rather look at where the information comes from, right? What is the means, the intention, etc.?"

#### **More Versus Less Credible Content**

The second research question asked which elements in AI-generated posts Spanish—
English bilinguals perceive as more versus less credible and if these perceptions differ between their L1 and L2. Table 5 summarizes the proportion of lexical content marked as more versus less credible in each post. For both the employment and rental posts, the students marked numerically a greater proportion of words as less credible in Spanish than in English. In contrast, at least for the employment post, they also marked a greater proportion of words as more credible in English than in Spanish.

Table 5

Descriptive Statistics (Means and Standard Deviations) for the Percent of Lexical Content

Marked as More Versus Less Credible in the Employment and Rental Posts

	Spa	nish	English		
Post type	Less credible More credible		Less credible	More credible	
Employment	9.27% (13.64)	9.83% (13.71)	4.49% (7.17)	13.11% (14.02)	
Rental	4.52% (4.94)	12.09% (14.23)	2.31% (2.33)	12.04% (13.96)	

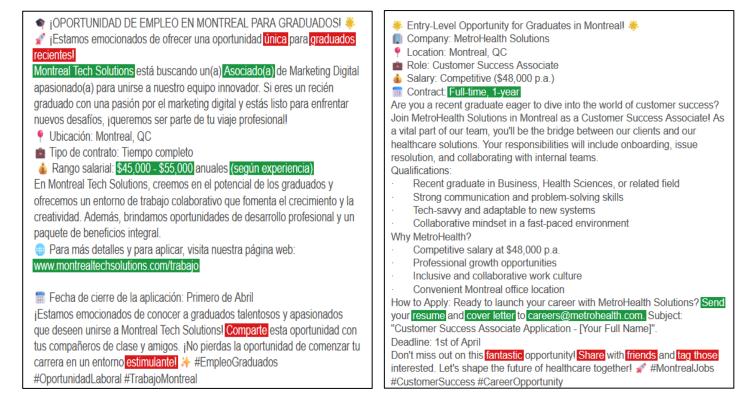
As summarized in Table 6, in terms of the specific lexical content marked as more versus less credible in the employment posts, more credible information included concrete, job-specific content such as the job application links (www.montrealtechsolutions.com/trabajo, careers@metrohealth.com), job descriptors such as \$45,000 (yearly salary), full-time, and such job qualifications as *según experiencia* (according to experience), *asociado(a)* (associate), resume, and cover letter. Less credible information generally included largely emotionally laden, descriptive content, including *estimulante!* (stimulating!), *unica* (unique), fantastic, *emocionados* (excited), as well as words related to sharing the job ad with friends, including *comparte* (share) and tag. Typical marking of more versus less credible content in the employment posts is illustrated in Figure 2.

**Table 6**Lexical Content Marked as More Versus Less Credible by at Least 6 Students (25%) in Spanish and English Employment Posts

Spa	nish	Eng	glish
Less credible	More credible	Less credible	More credible
estimulante (stimulating) (37.5%)	www.montrealtechsol utions.com/trabajo	tag (37.5%)	careers@metrohealth.
[oportunidad] única	(45.8%)	fantastic (33.3%) share (29.2%)	resume (45.8%)
(a unique opportunity) (33.3%)	según experiencia (according to	friends (29.2%)	full-time, 1 year (37.5%)
graduados recientes	experience) (37.5%)	[tag] those [interested] (25.0%)	send (37.5%)
(recent graduates) (33.3%)	\$45,000–\$55,000 (25.0%)	,	cover letter (37.5%)
comparte (share) (33.3%)	asociado(a) (associate) (25.0%)		
emocionados (excited) (29.2%)	Montreal Tech Solutions (25.0%)		

Figure 2

Typical Marking of More Credible (Green) Versus Less Credible (Red) Content in Spanish (Left Panel) and English (Right Panel) Employment Posts



Similarly, as summarized in Table 7, in the rental posts, more credible information involved location-specific descriptors, including *ubicación* (location), room size, utilities, *limpio* (clean), and *cercana* (close). Less credible content included such descriptors of anticipated roommates as amazing and professionals and the use of hyperbole to describe the rental, as in *una joya de espacio* (a gem of a space). Typical marking of more versus less credible content in rental posts is illustrated in Figure 3.

**Table 7**Top 10 Lexical Content Marked as More Versus Less Credible by at Least 6 Students (25%) in Spanish and English Rental Posts

Spa	anish	Eng	glish
Less credible	More credible	Less credible	More credible
<i>joya</i> (jewel) (58.3%)	ubicación (location)	professionals (25.0%)	monthly rent (41.7%)
de espacio (spacious)	(41.6%)		room size (37.5%)
(29.2%)	limpio (clean)		located (33.3%)
una (1 feminine)	(37.5%)		utilities (33.3%)
(25.0%)	ordenado (tidy) (37.5%)		conveniently (33.3%)
	mucha luz natural (a		
	lot of natural light)		
	(33.3%)		
	cercana (close to)		
	(33.3%)		
	\$800 (33.3%)		

Figure 3

Typical Marking of Credible (Green) Versus Noncredible (Red) Content in Spanish (Left Panel) and English (Right Panel) Rental Posts





The students' interview comments both extended and clarified these results. According to the students, more credible information in both languages generally involved specific content that they expected in each type of publication. In the employment post, for example, this content included the hiring company's name ("It may be a fraud, but I have the possibility of searching for that company on the Internet to see if it exists or not, or I can even find a number online and call," P35), an email address ("the way of contacting them and the ways of applying are very traditional and legitimate, such as sending them an email, which is safer, asking to send your resume, your cover letter and nothing else," P19), and a specific closing date ("they have a closing point and, that is, if you apply before... that's fine and if not, no way... it made me think it was credible because of that," P15). Similarly, in the rental post, more credible details included the information about a monthly rent and location ("I find that it is more credible for them to tell

you, look, this is what the apartment itself is, with what is included and if we are next door [to places nearby]," P26). The rental post was also perceived credible because of its detailed description of roommates. According to P21, because the post was written in Spanish, the information about the roommates was particularly credible ("And the most credible part was that they [the roommates] were talking about themselves, right? And they were in Spanish, maybe they are Latinos, friendly, respectful, I do believe them, because they are Latinos, not gringos, right?"). For P12, who was looking for a room at the time of the interview, the invitation to schedule a visit also greatly increased this post's credibility:

When you see the online publication, they directly ask you for a deposit to make a visit, and then I already know that it is not true... But if they ask you to schedule a visit, well, there is less risk for the person applying, let's say, of any scam. So that's why I found it more believable.

Other students such as P14 similarly shared their personal experiences looking for rooms for rent in the city, saying that the rental posts were similar to those they had seen online: "When I came here [Montreal], I have moved about twice, I have also searched for places to live on Facebook, and what appears there [in the rental posts] is very similar."

For other students, what made specific post elements more credible was that they resembled the content typically seen in their daily usage of social media ("I kind of feel that the housing ones seemed more credible to me because people do that a lot, like posting on Facebook that they are looking for someone or subleasing or something," P22) or that they were laid out as a typical job ad ("the formatting looks like something I normally see on Indeed [a job searching platform]," P13). The students also reported the welcoming and encouraging language employed, saying that it contributed to their perception of the post's credibility, including "words like

'collaborative,' 'innovative,' 'space' that gives you opportunities, encourages development... that all seemed quite in line with what a normal post about a job would be like" (P19).

In terms of less credible aspects of the two posts, the students reported the use of jargon, set expressions, excessive qualifiers, and hashtags or emojis as particularly suspicious, for instance, with such wordings as "the vital part of the team" and "competitive" perceived as red flags. For P25, the use of reader-focused, emotional language such as "we are looking for the perfect roommate" and "we look forward to get to know you and share this incredible space with you" emphasized the sense of urgency that these wordings entailed, making them less credible. The students considered the use of emojis as something that "a serious company would not put" (P30) and generally as unprofessional ("when I see that [emojis], the job posts, when they offer work, it doesn't seem so professional to me, I don't know, so it's like, I feel like it takes away a little bit of credibility," P23). Similarly, hashtags were perceived as incompatible with reputable posts, as explained by P25: "Generally I don't see hashtags, unless they are for entertainment...

But when I search for, for example, I don't know, something on Facebook, I don't see the hashtags."

For many students, a major concern about the post's credibility pertained to the quality of language used. For example, in the Spanish rental post, P17 considered the wording *una joya de espacio* (a gem of a space) as a poor automated translation, and for P29, the expression *envia un mensaje directo* (send a direct message) seemed to be "straight out of Google Translate, because normally it wouldn't be used... it will say 'private message' or 'in private,' 'DM', so this is literally the translation." For other students, such wordings were incompatible with the kind of language that is typically used in their age demographic, as explained by P32: "as if someone

literally copied what the translator said... I kind of feel like we young people don't use it as much as, 'oh, what a jewel!' I kind of feel like it's something my mom would say."

Finally, the students drew on different micro- and macro-level details to consider the content as less credible. For example, P14 noticed that the link to the company's career page appeared in Spanish, which was suspicious to this student "because [they] thought that if it is a company from Montreal, the link should not be in Spanish." P14 also considered the entire post as suspicious because, taken in its entirety, the employment information sounded overly generic: "it doesn't make sense to me that they open a job for graduates, since it's very generic, it doesn't say in what area."

## **Chapter 5. Discussion and Implications**

This study explored whether Spanish–English bilinguals perceive credibility of online information differently across their two languages, focusing in particular on Facebook-like posts generated by AI. The study's results illustrated a pattern whereby the students rated the credibility of posts in L2 English higher than in L1 Spanish, regardless of the post's topic (i.e., rental or employment). Even though they did not express a statistically significant preference in one language versus the other, the students' rated willingness to share posts was numerically greater in English than in Spanish. In terms of the elements they marked as more versus less credible, the students' credibility judgements also diverged between their L1 and L2. The general trend was that they identified numerically more words as more credible in the English posts and more words as less credible in the Spanish posts.

The students' qualitative comments shed light on their perceptions of credibility, providing insights into their critical media literacy. Many students attributed greater credibility to English posts for a variety of reasons, including their prior experience with news in Spanish, which they claimed often includes misinformation. For others, Spanish posts seemed less credible because, in the context of Montreal, most publications appear in English or French, so the use of Spanish made the posts suspicious. The students attributed their ability to judge a post's credibility to their superior intuition in their L1, which allows them to be more critical in their judgment of information quality, whereas this ability is lacking in their L2. And for a post to be considered credible, according to the students, it had to resemble similar posts online and to include expected details such as the listing of a monthly rent and location in the rental posts and the company name and application closing date in the employment posts. Less credible content included mainly the use of jargon, excessive qualifiers, emojis, and hashtags, and involved specific post details such as external links and particular wordings which were considered stilted

or poor translations. The students' attention to such detail was more pronounced in relation to the Spanish than English posts.

## Information Credibility and the Foreign Language Effect

This study was broadly conceptualized within the construct of the foreign language effect, which refers to the tendency for bilinguals to engage in riskier decision-making and to opt for more utilitarian choices in moral dilemmas in their L2 than in their L1 (e.g., Costa et al., 2014; Keysar et al., 2012). The present findings add to this literature by showing that there might be a similar language-based difference in how bilinguals evaluate the credibility of Facebooklike posts. Considering that the students in this study evaluated information in L2 English as more credible than in L1 Spanish, the use of L2 appears detrimental to bilinguals' decisionmaking, at least as far as their critical media literacy skills are concerned. This finding responds to recent calls to examine the role of language in bilinguals' decision-making (Hayakawa et al., 2016) and extends prior work which has shown bilinguals to be superior at detecting fake news headlines in their L1 versus their L2 (Caramancion, 2022; Muda et al., 2023). The students' bias to consider information more credible in the L2 than in the L1 was especially striking, given that all information had been generated by AI and included content details that were never seen before (e.g., names of companies, rental locations). All in all, the present findings underscore the concern that L2 speakers "can be less able to gain accurate information about their environment through online news sources partially populated by fake news" (Muda et al., 2023, p. 722). More importantly, these findings imply that this population of language speakers—particularly refugees, immigrants, and international students who might not be as proficient in their L2 as in their L1—could face a far greater risk to fall victim to frauds or scams, especially given the latest advancements in AI (Cardenuto et al., 2023; Kreps et al., 2022).

There are several reasons, including cognitive control, emotions, and social norms, which could explain the role of language in bilinguals' credibility judgments and critical media literacy skills as shown in this study (Circi et al., 2021). According to one explanation, for instance, bilinguals could engage with their two languages at different levels of emotionality, such that their experience in the L1 might involve a deeper, more emotional processing than in their L2 (Caldwell-Harris, 2014, 2015; Caldwell-Harris & Ayçiçeği-Dinn, 2009; Dylman & Champoux-Larsson, 2020; Fernández-López & Perea, 2020; Pavlenko, 2012; Vanek & Tovalovich, 2022). In this study, the students were asked to determine the credibility of rental and employment posts which contained little personal content and which therefore created little emotional involvement for the students other than requiring them to indicate whether they would believe such posts if they encountered them on social media. However, emotionality may have supported the students' rationalization. As Hayakawa et al. (2016) suggest, when "little time or few resources available to engage in careful deliberation, using a foreign language should lead to less optimal decisions" (p. 972). Because the students had to rely in their judgments only on the text, in the absence of further information, access to images, or additional search functions, they may have generally engaged in greater, deeper, or otherwise different emotional reasoning in their L1 than in their L2, which allowed them to exercise their critical media literacy skills to a greater extent in Spanish than in English.

Another explanation for the language difference in the students' performance might reflect L1–L2 differences in cognitive control, where L2 processing is more effortful than L1 processing (Kahneman, 2011). The students' credibility judgments may have therefore reflected an increase in cognitive load when reading in the L2 versus the L1. A lower cognitive load in Spanish thus allowed the students to perform a better analysis of each post's content, whereas a

more effortful processing in English may have made it harder for them to perform a careful analysis and ultimately resist the perception that the information in English was more credible. For instance, in a study examining bilinguals' decision-making across scenarios that involved various forms of bias (e.g., assessing decisions more favorably if they lead to positive outcomes), Vives et al. (2018) showed that a greater cognitive load experienced by bilinguals in their L2 made it harder for them to engage in rational thinking and to resist built-in task-specific biases. In fact, according to Hu et al. (2024), information processing in bilinguals' L2 is often not only slower but also qualitatively different from information processing in their L1, in that bilinguals experience particular challenges with evaluating and integrating previously encountered information. Against this backdrop, therefore, it is not altogether surprising that the students in this study may have engaged in more superficial, effortful, or otherwise less critical analysis of text credibility in their L2 than in their L1.

Yet another reason for the L1–L2 difference in the students' credibility judgments could also reflect various social forces. As the students indicated in their interviews, Spanish was perceived as less credible whereas English was described as "superior," which may have rendered the information in English to feel more credible than in Spanish, even though all posts featured AI-generated materials and were thus equally "made up." For example, Latin Americans' favorable perceptions of English are well documented, where having higher proficiency in English is associated with greater economic opportunity across multiple nations, including Puerto Rico (González-Rivera, 2021), Mexico (Despagne, 2010), and Argentina (Friedrich, 2009), with English as an international lingua franca considered an indicator of high socioeconomic status and professional growth (Rojas & Hernández-Fernández, 2018). In fact, the students' perceptions could broadly reflect a form of confirmation bias motivated by their

sociolinguistic reality and informed by their attitudes. In essence, the students tended to perceive the information in English as more credible than in Spanish because these perceptions conformed to their internalized beliefs that English is superior to Spanish, in that English affords greater economic and social opportunity and therefore is more legitimate or trustworthy. In the absence of attitudinal judgments collected from the participants, this explanation remains speculative at best. Nevertheless, it is likely that multiple variables interacted in their influence on the students' credibility judgments, in the sense that considering text quality in the L2 involved less emotional rationalization, greater cognitive effort, and stronger likelihood of engaging in socially constructed confirmation bias, compared to judging text quality in the L1.

Even though the students considered the posts presented in L2 English to be more credible than similar posts in L1 Spanish, they generally did not show a preference to share them differently in their L2 versus L1. These results are similar to those reported by Pennycook et al. (2020, 2021), who examined how readers judged the accuracy of political and Covid-19 headlines. For those participants, perceived accuracy of texts was unrelated to the likelihood of sharing them, meaning that people are similarly likely to share posts they would judge as accurate or inaccurate (Pennycook & Rand, 2021). Although there was no reliable statistical difference by language, the students nevertheless showed a numeric tendency to favor the English over the Spanish posts in their willingness to share them. This implies that bilinguals' credibility judgements might drive their real-world actions and that these judgments—and associated actions—might differ for bilinguals in terms of which language they use. Although tentative, this conclusion further reinforces the need for researchers and practitioners to consider bilinguals' media literacy skills and to understand the role of language in those skills, to build a

media-literate and media-savvy society (Lazer et al., 2018), especially in a country like Canada, where multilingualism is prevalent.

## **Bilingualism and Critical Media Literacy**

In terms of their stated reasons for considering specific content as more versus less credible, in their majority, the students appeared to rely on a set of personal criteria which they had developed through their individual experience of looking for jobs or apartment rentals. As discussed previously, this points to a potential confirmation bias, where people use established expectations to judge information credibility, thereby often disregarding the importance of new, text- or situation-specific information (Nickerson, 1998). In essence, the students demonstrated a pattern according to which they established credibility based on their often idiosyncratic prior experience, where the information that is normally or previously seen in similar contexts was used to judge a post's credibility. This confirmation bias is especially alarming given the current reality where information can be easily generated using AI tools with "a convincing illusion of human intelligence" (Hutchens, 2024, p. 152) and can be used to mislead, deceive, or otherwise harm people. As aptly pointed out by P12, "the truth is they all [posts in the experiment] could have been generated by AI, now that I think about it, but at the end of the day artificial intelligence is based on posts that have already existed." Thus, because generative AI is trained on human-made content, confirmation bias is amplified, and media users would most likely be exposed to what they already expect to see. Considering that media literacy skills vary widely from person to person and that these skills might differ across bilinguals' languages, it becomes highly problematic to rely only on prior individual experience (and associated confirmation bias) to establish the authenticity of information in the present-day media space, where AI-generated content can easily "fool" media users.

Apart from being influenced by potential confirmation bias, the students nevertheless demonstrated a clear awareness and use of specific critical media skills, as categorized by Kellner et al. (2019) into six comprehensive areas: social constructivism, languages/semiotics, audience/positionality, politics of representation, production/institutions, and social and environmental justice. In their interview comments, the students questioned the intentions of the presumed authors of the posts and their target audience. For example, they attended to how language and other symbols such as emojis were used in the posts (e.g., "The use of emojis, well, artificial intelligence is supposed to be able to do that very well too," P21), considered the post's audience and positionality ("Am I going to live with my landlord or am I going to live with a roommate?" P32), and raised issues of production/institutions, trying to understand why the post was shared or created ("Why is this in Spanish? Who is this directed towards?" P19). Moreover, as shown in their interview responses, the students appeared to analyze and compare each post's elements (e.g., language expressions, emojis, hidden meanings, etc.) to a certain standard which, according to Potter (2021), is an indication of tangible competencies for media literacy. As discussed previously, the students were able to deploy these competencies more effectively—or to a greater degree of depth and analysis—in their L1 Spanish than in their L2 English, likely in light of their greater familiarity with how information is produced and consumed in their L1 and their capacity to perform a better analysis of it, as illustrated by P25: "In Spanish I already know that when they say a phrase to me, I already know that they want to deceive me... In English I already recognize a little more but I feel more comfortable in Spanish than in English."

Among Kellner et al.'s (2019) six comprehensive areas of critical media literacy skills, the topics of social constructivism, politics of representation, and social and environmental justice did not come up explicitly in the students' interview responses. One reason for the

absence of these themes in the students' comments is that this study's materials did not directly focus on these aspects of media literacy, as the content presented to the students concerned rental and employment posts. Nevertheless, broadly speaking, the students did seem to be aware of these aspects of critical media literacy. For instance, when the students noted that posts written in Spanish are often meant to scam newcomers who do not speak English, they highlighted the social justice aspect of critical media literacy, by virtue of questioning who benefits from such information and who can be disadvantaged by it.

### **Implications**

All in all, the Spanish–English bilingual students in this study appeared to display a confirmation bias, likely driven by their language-specific prior experiences and beliefs, while also demonstrating the ability to employ some critical media literacy tools to analyze the information they encounter online. Nevertheless, a language-based difference in the students' media literacy skills was particularly salient in their marking of specific lexical content as more or less credible. The students marked a greater quantity of Facebook-like posts as problematic to a text's credibility in L1 Spanish than in L2 English, and their choices often diverged across the two languages. For example, whereas they seemed to attend to linguistic expressions in their L1 (e.g., una joya de espacio "a gem of a space," estimulante "stimulating"), they tended to focus on factual content details to determine credibility in their L2 (e.g., whether such relevant information as application deadline or contact details were included). Taken together, these findings seem to suggest that a potential way to improve bilingual students' media literacy would be to conduct literacy tasks in both languages, coupled with reflection phases, where students discuss what they notice in each language, which could facilitate the transfer of skills across language. Considering that the students in the study conceptualized media literacy somewhat differently in their two languages, any training focusing on critical media literacy therefore

should focus on the development of complementary, language-independent skills which should ideally extend beyond media literacy into decision-making in other domains like preventive medical care (Hayakawa et al., 2021) and business negotiations (Lee et al., 2024).

Even though the students in this study demonstrated some awareness and knowledge of critical media literacy (Kellner et al., 2019), including fact-checking to determine information credibility in their daily media use (Huschens et al., 2023; Labajová, 2023), it is important to stress that they were not typical immigrants and newcomers to Canada. According to Statistics Canada (2021), only 19.1% of Latino immigrants to Canada between 1980 and 2000 held a bachelor's degree or higher. Although this figure doubled to 46.2% in the next decade (Statistics Canada, 2023), this still means that more than half of Latino immigrants to Canada are less educated, possibly less aware of the benefits and pitfalls of AI, and thus potentially more susceptible to media-driven frauds and scams. Therefore, the students participating in the study must have benefited from exposure to AI and its capabilities as part of their academic and professional endeavors, a privilege other Latin-American immigrants do not necessarily have. This further highlights the importance of exploring issues of critical media literacy among different populations, preferably those who are not involved in post-secondary studies.

Finally, this study's methodology of asking participants to first review and then reflect on their credibility ratings, which included the highlighting of more versus less credible text elements, illustrated a possible instructional intervention to increase media users' awareness and help them become more vigilant in their processing of information online, especially with respect to content that could be generated by AI. As emphasized by Hutchens (2024), the main threat does not stem from the abilities of AI, but rather from the potentially malicious usage people could make of it. The present findings broadly suggest that the rapid advancements of AI and

potential threats posed by it are not—at least presently—matched with media users' critical literacy skills. This highlights not only the need to carry out further research on this issue but also to develop and implement dedicated instructional interventions to train people's media literacy when interacting with AI-generated content, with the goal of making media users resilient to misinformation, frauds, and scams online.

### **Limitations and Recommendations**

The present study employed a new paradigm to explore credibility in the age of AI. Nevertheless, as with any innovation, there are several limitations that should be recognized, so they can be addressed in future work. First, a relatively small sample size (24 bilinguals) limits the generalizability of findings, especially in light of nonnormal data distributions, which disallowed the use of parametric statistics. In terms of the materials, the participants pointed out that the lack of images in Facebook-like posts made the credibility judgements harder to establish, because they were used to relying on images in their judgements. Although the use of graphics would have taken the focus away from text, their use would have been more representative of real-life social media posts, which include images. Furthermore, the restricted number of posts (four in total) presents a limitation as the content was narrowed down to one post per language (e.g., one rental and one employment post in Spanish), which did not allow for a robust comparison of common themes within and across languages. Finally, the use of ChatGPT to generate the Spanish posts was another challenge, considering that the participants readily noted some unusual, stilted, or uncommon language use, which they described as the result of an "automatic translation." It remains possible, therefore, that the present findings for credibility of the Spanish posts reflected a generally poorer knowledge base of the ChatGPT tool used to generate the materials in Spanish.

In terms of various methodological choices, the use of Facebook as a social media platform for use with bilinguals aged 19–44 years old may have been problematic, given that half of the students indicated to use Facebook for only 1–2 hours per week, whereas the average use of social media in a relatively young demographic averages nearly 3 hours per day (Bottorf & Wong, 2023). Thus, because Facebook may not have been the students' primary social media platform, the study may have failed to capture their typical or representative credibility judgments for social media content. Moreover, the students appeared to use Facebook mainly to maintain contact with friends and family and share visual information between family and friends, rather than to explore rental or employment opportunities. The study may not have been fully successful, therefore, in addressing the main functions the students fulfill using Facebook.

Furthermore, another potential caveat related to the students was their extensive age range. There may be notable difference in how 19-year-old and 44-year-old students approach credibility of online content, so it is recommended that future studies opt for a restricted age group to better control or otherwise account for each demographic's life experiences. Our recommendation for future studies would be the recruitment of a larger sample size, targeting a homogenous age group. Moreover, the materials should be adapted to accommodate the social media platforms that participants use more often, which necessarily requires a pre-study review of the most commonly used media platforms for each age group. In addition, the materials developed for this study can be used to support educational interventions, to raise media users' awareness of the strengths and weaknesses of generative AI and to enable them reflect on and further develop their media credibility toolkit, fostering a more credibility-focused approach to media consumption.

Last but not least, the language pair chosen for this study may not have been ideal, considering that the city of Montreal is predominantly a French–English bilingual city even though Spanish is widely spoken. Additionally, the lack of a counterbalanced bilingual group such as an additional sample of English–Spanish bilinguals minimized the study's ability to determine the robustness of language effect on credibility. For instance, if the English–Spanish group displayed a different pattern in their credibility judgements, this would indicate that bilinguals' credibility judgments have more to do with the specific languages they speak rather than with an across-the-board L1–L2 effect. Therefore, future research would benefit from including other language pairs (e.g., English–French) and testing counterbalanced groups of bilinguals, to establish the magnitude of the foreign language effect and explore whether it indeed reproduces across bilinguals.

# **Chapter 6. Conclusion**

This study explored Spanish–English bilingual students' critical media literacy skills across their two languages, focusing in particular on AI-generated Facebook-like posts. The results indicate that the students were generally less critical in assessing content credibility in their L2 than in their L1. This language-based difference emerged in the students' judgments of credibility but also manifested through a non-significant trend in terms of the students' rated likelihood of sharing posts online. The students demonstrated a somewhat different pattern of analysis through which they approached their judgment of text credibility in their two languages, where they relied on factual details to determine credibility in English but referred to specific language expressions and use of emotional language to judge credibility in Spanish. The students in this study also appeared to exhibit what could be interpreted as confirmation bias, whereby they tended to determine credibility of content based on their prior, often person-specific

experiences, judging credibility through an expectation-driven process rather than a critical analysis of the actual information presented to them.

The students' tendency to rely on confirmation bias in judging credibility is concerning, considering various potentially malicious uses of AI which might target refugees, immigrants, and international students—including those from Latin America—as particularly vulnerable groups. Although the participants generally recognized the possibility of the study's posts being generated by AI, and some in fact claimed to have employed various fact-checking methods in their daily consumption of information online, distinguishing human texts from AI-generated content is nearly impossible, given the backend size and capability of AI tools. This study and its findings, which provide one example of how generative AI could be used in applied linguistics, call for more work on critical media literacy at the intersection of AI and bilingualism and multilingualism.

#### References

- Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*, 13(1). <a href="https://doi.org/10.1007/s13278-023-01028-5">https://doi.org/10.1007/s13278-023-01028-5</a>
- Al Zou'bi, R. M. (2022). The impact of media and information literacy on students' acquisition of the skills needed to detect fake news. *Journal of Media Literacy Education*, *14*(2), 58–71. https://doi.org/10.23860/JMLE-2022-14-2-5
- Ali, S., Dipaola, D., Lee, I., Sindato, V., Kim, G., Blumofe, R., & Breazeal, C. (2021). Children as creators, thinkers and citizens in an AI-driven future. *Computers and Education: Artificial Intelligence*, 2, 1–11. <a href="https://doi.org/10.1016/j.caeai.2021.100040">https://doi.org/10.1016/j.caeai.2021.100040</a>
- Arechar, A. A., Allen, J., Berinsky, A. J., Cole, R., Epstein, Z., Garimella, K., Gully, A., Lu, J. G., Ross, R. M., Stagnaro, M. N., Zhang, Y., Pennycook, G., & Rand, D. G. (2023). Author correction:

  Understanding and combatting misinformation across 16 countries on six continents. *Nature Human Behaviour*, 7(10), 1797. <a href="https://doi.org/10.1038/s41562-023-01676-9">https://doi.org/10.1038/s41562-023-01676-9</a>
- Arguedas, A. R., & Simon, F. M. (2023). Automating democracy: Generative AI, journalism, and the future of democracy. Balliol Interdisciplinary Institute, University of Oxford.
- Bermes A. (2021). Information overload and fake news sharing: A transactional stress perspective exploring the mitigating role of consumers' resilience during COVID-19. *Journal of Retailing and Consumer Services*, 61, Article 102555. https://doi.org/10.1016/j.jretconser.2021.102555

- Białek, M,. (2023). Why should we study the Foreign Language Effect: Debiasing through affecting metacognition? *Journal of Intelligence*, *11*(103), 103.

  <a href="https://doi.org/10.3390/jintelligence11060103">https://doi.org/10.3390/jintelligence11060103</a>
- Białek, M., Domurat, A., Paruzel-Czachura, M., & Muda, R. (2022). Limits of the foreign language effect: Intertemporal choice. *Thinking & Reasoning*, 28(1), 97–124. https://doi.org/10.1080/13546783.2021.1934899
- Białek, M., Muda, R., Stewart, K., Niszczota, P., & Pieńkosz, D. (2020). Thinking in a foreign language distorts allocation of cognitive effort: Evidence from reasoning. *Cognition*, 205. <a href="https://doi.org/10.1016/j.cognition.2020.104420">https://doi.org/10.1016/j.cognition.2020.104420</a>
- Blauth, T. F., Gstrein, O. J., & Zwitter, A. (2022). Artificial intelligence crime: An overview of malicious use and abuse of AI. *IEEE Access*, *10*. <a href="https://doi.org/10.1109/ACCESS.2022.3191790">https://doi.org/10.1109/ACCESS.2022.3191790</a>
- Bottorf, C., & Wong, B., (2023, May 18). *Top social media statistics and trends of 2024*. Forbes Advisor. Retrieved June 26, 2024, from https://www.forbes.com/advisor/business/social-media-statistics/#source
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
- Bulger, M., & Davison, P. (2018). The promises, challenges and futures of media literacy. *Journal of Media Literacy Education*, 10(1), 1–21.
- Butler, A. (2021). *Key scholarship in media literacy : David Buckingham* (Ser. Brill guides to scholarship in education, volume 5). Brill Sense.

- Byers-Heinlein, K., Esposito, A. G., Winsler, A., Marian, V., Castro, D. C., & Luk, G. (2019). The case for measuring and reporting bilingualism in developmental research. *Collabra: Psychology*, *5*(1). <a href="https://doi.org/10.1525/collabra.233">https://doi.org/10.1525/collabra.233</a>
- Caldwell-Harris, C. L. (2014). Kill one to save five? mais oui! *Scientific American Mind*, 25(5), 70–73. https://doi.org/10.1038/scientificamericanmind0914-70
- Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language:

  Implications for everyday life. *Current Directions in Psychological Science*, *24*(3), 214–219.

  https://doi.org/10.3389/fpsyg.2015.00437
- Caldwell-Harris, C. L., & Ayçiçeği-Dinn, A. (2009). Emotion and lying in a non-native language.

  \*International Journal of Psychophysiology, 71(3), 193–204.

  https://doi.org/10.1016/j.ijpsycho.2008.09.006
- Canadian Anti-Fraud Centre. (2021). *Annual Report 2021*. Government of Canada. <a href="https://publications.gc.ca/collections/collection\_2022/grc-rcmp/PS61-46-2021-eng.pdf">https://publications.gc.ca/collections/collection\_2022/grc-rcmp/PS61-46-2021-eng.pdf</a>
- Canadian Anti-Fraud Centre. (n.d.). *Canadian Anti-Fraud Centre*. <u>Government of Canada</u>. <u>https://www.antifraudcentre-centreantifraude.ca/index-eng.htm</u>
- Canadian Anti-Fraud Centre (2022, May 17). Frauds affecting individuals. Government of Canada.

  <a href="https://www.antifraudcentre-centreantifraude.ca/scams-fraudes/individual-individuelle-eng.htm">https://www.antifraudcentre-centreantifraude.ca/scams-fraudes/individual-individuelle-eng.htm</a>
- Caramancion, K. M. (2022, January). The role of user's native language in mis/disinformation detection:

  The case of English. In 2022 IEEE 12th Annual Computing and Communication Workshop and

  Conference (CCWC) (pp. 260–265). IEEE. https://doi.org/10.1109/CCWC54503.2022.9720812

- Cardenuto, J. P., Yang, J., Padilha, R., Wan, R., Moreira, D., Li, H., Wang, S., Andaló, F., Marcel, S., & Rocha, A. (2023). The age of synthetic realities: challenges and opportunities. *APSIPA Transactions on Signal and Information Processing*, 12(1). <a href="https://doi.org/10.1561/116.00000138">https://doi.org/10.1561/116.00000138</a>
- Chen, J. J. (2016). Educating English language learners for success in the 21<sup>st</sup> Century: Facilitating their acquisition of multiliteracies. In M. N. Yildiz & J. Keengwe (Eds.), *Handbook of research on media literacy in the digital age* (pp. 75–90). IGI Global.
- Cherow-O'Leary, R. (2023). Reconceptualizing ledia Literacy for the midtwenty-first century. In De Abreu, B. S. (Ed.), *Media literacy, equity, and justice* (pp. 257–280). Routledge. <a href="https://doi.org/10.4324/9781003175599">https://doi.org/10.4324/9781003175599</a>
- Circi, R., Gatti, D., Russo, V., & Vecchi, T. (2021). The foreign language effect on decision-making: A meta-analysis. *Psychonomic Bulletin & Review*, 28(4), 1131–1141.

  https://doi.org/10.3758/s13423-020-01871-z
- Cisek, G. (2021). The triumph of artificial intelligence: How artificial intelligence is changing the way we live together. Springer. https://doi.org/10.1007/978-3-658-34896-0
- Cohen J. (1988). Statistical power analysis for the behavioral sciences. Routledge Academic.
- Competition Bureau Canada (2018). *The little black book of scams* (2nd ed.). Government of Canada. <a href="https://ised-isde.canada.ca/site/competition-bureau-canada/en/little-black-book-scams-2nd-edition">https://ised-isde.canada.ca/site/competition-bureau-canada/en/little-black-book-scams-2nd-edition</a>
- Concordia University. (n.d.). *Admissions*. Concordia University.

  https://www.concordia.ca/academics/graduate/calendar/current/admissions.html

- Concordia University. (n.d.). English language proficiency. Concordia University Undergraduate admissions. <a href="https://www.concordia.ca/admissions/undergraduate/requirements/english-language-proficiency.html">https://www.concordia.ca/admissions/undergraduate/requirements/english-language-proficiency.html</a>
- Cook, P. (2023). Beyond "Fake News": Misinformation studies for a postdigital era. In L. Parker (Ed.0, *Education in the age of misinformation: Philosophical and pedagogical explorations* (pp. 9–32). Palgrave Macmillan. <a href="https://doi.org/10.1007/978-3-031-25871-8">https://doi.org/10.1007/978-3-031-25871-8</a>
- Costa, A., Duñabeitia, J. A., & Keysar, B. (2019). Language context and decision-making: Challenges and advances. *Quarterly Journal of Experimental Psychology*, 72(1), 1–2. https://doi.org/10.1177/1747021818789799
- Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2013). "Piensa" twice: On the foreign language effect in decision making. *Cognition*, *130*(2), 236–254.

  <a href="https://doi.org/10.1016/j.cognition.2013.11.010">https://doi.org/10.1016/j.cognition.2013.11.010</a>
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *Plos One*, *9*(4), 94842.

  <a href="https://doi.org/10.1371/journal.pone.0094842">https://doi.org/10.1371/journal.pone.0094842</a>
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications.
- de Zengotita, T. (2006). *Mediated: How the media shapes your world and the way you live in it.*Bloomsbury.

- Del Maschio, N., Crespi, F., Peressotti, F., Abutalebi, J., & Sulpizio, S. (2022). Decision-making depends on language: A meta-analysis of the Foreign Language Effect. *Bilingualism*, 25(4), 617-630. https://doi.org/10.1017/S1366728921001012
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554–559. https://doi.org/10.1073/pnas.1517441113
- Despagne, C. (2010). The difficulties of learning English: Perceptions and attitudes in Mexico.

  Canadian and International Education, 39(2), 55–74.
- Diaz-Lago, M., & Matute, H. (2019). Thinking in a foreign language reduces the causality bias.

  \*\*Quarterly Journal of Experimental Psychology, 72(1), 41–51.

  \*\*https://doi.org/10.1177/1747021818755326\*\*
- Dylman, A. S., & Champoux-Larsson, M. F. (2020). It's (not) all Greek to me: Boundaries of the foreign language effect. *Cognition*, 196, 104148–104148.

  <a href="https://doi.org/10.1016/j.cognition.2019.104148">https://doi.org/10.1016/j.cognition.2019.104148</a>
- Dylman, A. S., & Champoux-Larsson, M.-F. (2020). It's (not) all greek to me: Boundaries of the foreign language effect. *Cognition*, 196, 104148–104148.

  https://doi.org/10.1016/j.cognition.2019.104148
- Fernández-López, M., & Perea, M. (2020). Language does not modulate fake news credibility, but emotion does. *Psicológica Journal*, 41(2), 84–102. <a href="https://doi.org/10.2478/psicolj-2020-0005">https://doi.org/10.2478/psicolj-2020-0005</a>

- Ford, T. W., Yankoski, M., Facciani, M., & Weninger, T. (2023). Online media literacy intervention in Indonesia reduces misinformation sharing intention. *Journal of Media Literacy Education*, *15*(2), 99–123. https://doi.org/10.23860/JMLE-2023-15-2-8
- Friedrich, P. (2003). English in Argentina: Attitudes of MBA Students. World Englishes, 22(2), 173–184.
- Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social clicks what and who gets read on twitter? *ACM Sigmetrics Performance Evaluation Review*, *44*(1), 179–192. https://doi.org/10.1145/2964791.2901462
- Geeng, C., Yee, S., & Roesner, F. (2020, April). Fake news on Facebook and Twitter: Investigating how people (don't) investigate. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–14). https://doi.org/10.1145/3313831.3376784
- González-Rivera, M. (2021). Language attitudes towards Spanish and English in Puerto Rico. *Revista de Filología Y Lingüística de La Universidad de Costa Rica*, 47(2).

  <a href="https://doi.org/10.15517/rfl.v47i2.47006">https://doi.org/10.15517/rfl.v47i2.47006</a></a>
- Hadjichristidis, C., Geipel, J., & Savadori, L. (2015). The effect of foreign language in judgments of risk and benefit: The role of affect. *Journal of Experimental Psychology. Applied*, 21(2), 117–29. https://doi.org/10.1037/xap0000044
- Hanson, R. E. (2022). Mass communication: Living in a media world (8th ed.). SAGE.
- Hayakawa, S., Costa, A., Foucart, A., & Keysar, B. (2016). Using a foreign language changes our choices. *Trends in Cognitive Sciences*, 20(11), 791–793. https://doi.org/10.1016/j.tics.2016.08.004

- Hayakawa, S., Pan, Y., & Marian, V. (2021). Using a foreign language changes medical judgments of preventative care. *Brain Sciences*, 11(10). https://doi.org/10.3390/brainsci11101309
- Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychological Science*, 28(10), 1387–1397. <a href="https://doi.org/10.1177/0956797617720944">https://doi.org/10.1177/0956797617720944</a>
- Heaven, W. D. (2023, February 8). *ChatGPT is everywhere. Here's where it came from*. MIT Technology Review. <a href="https://www.technologyreview.com/2023/02/08/1068068/chatgpt-is-everywhere-heres-where-it-came-from/">https://www.technologyreview.com/2023/02/08/1068068/chatgpt-is-everywhere-heres-where-it-came-from/</a>
- Higdon, N., & Huff, M. (2022). Let's agree to disagree: A critical thinking guide to communication, conflict management, and critical media literacy. Routledge, Taylor & Francis Group.
- Horne, B.D., & Adali, S. (2017). This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. *Proceedings of the International AAAI Conference on Web and Social Media, 11*(1), 759-766.

  https://doi.org/10.1609/icwsm.v11i1.14976
- Hu, Z., Suitner, C., & Navarrete, E. (2024). Uncovering the role of foreign language on acquiescence.

  \*Bilingualism: Language and Cognition, 1–7. doi:10.1017/S1366728924000178
- Hu, Z., Suitner, C., & Navarrete, E. (2024). Uncovering the role of foreign language on acquiescence.

  \*Bilingualism: Language and Cognition, 1–7. https://doi.org/10.1017/S1366728924000178
- Huschens, M., Briesch, M., Sobania, D., Rothlauf, F. (2023). Do You Trust ChatGPT? -- Perceived Credibility of Human and AI-Generated Content. *arXiv*.

https://doi.org/10.48550/arXiv.2309.02524

- Hutchens, J. (2024). The language of deception weaponizing next generation AI. Wiley.
- Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Giroux.
- Karaduman, S. (2015). The role of critical media literacy in further development of consciousness of citizenship. *Procedia Social and Behavioral Sciences*, *174*, 3039–3043. https://doi.org/10.1016/j.sbspro.2015.01.1095
- Kellner, D., Share, J., & Luke, A. (2019). *The critical media literacy guide: Engaging media and transforming education* (Ser. Brill guides to scholarship in education, volume 2). Brill Sense. <a href="https://doi.org/10.1163/9789004404533">https://doi.org/10.1163/9789004404533</a>
- Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*(6), 661–8. https://doi.org/10.1177/0956797611432178
- Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science*, 9(1), 104–117. <a href="https://doi.org/10.1017/XPS.2020.37">https://doi.org/10.1017/XPS.2020.37</a>
- Kvale, S. (2007). Conducting an interview. In *Doing Interviews* (pp. 52-66). SAGE Publications, Ltd, <a href="https://doi.org/10.4135/9781849208963">https://doi.org/10.4135/9781849208963</a>
- Labajová, L. (2023). The state of AI: Exploring the perceptions, credibility, and trustworthiness of the users towards AI-generated content [Master's thesis, Malmö University]. DiVA. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1772553&dswid=-6301
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R.,

- Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <a href="https://doi.org/10.1126/science.aao2998">https://doi.org/10.1126/science.aao2998</a>
- Lee, J. H., Elfenbein, H. A., & Bottom, W. P. (2024). Foreign language effect in negotiations:

  Negotiation language and framing effect on contract terms and subjective outcomes.

  International Journal of Conflict Management, 35(2), 270–286. <a href="https://doi.org/10.1108/IJCMA-03-2023-0052">https://doi.org/10.1108/IJCMA-03-2023-0052</a>
- Long, Y., Lu, Q., Xiang, R., Li, M., Huang, C.R. (2017). Fake news detection through multi-perspective speaker profiles. *Proceedings of the Eighth International Joint Conference on Natural Language Processing* (Volume 2: Short Papers), pp. 252–256.
- Majerczak, P., & Strzelecki, A. (2022). Trust, media credibility, social ties, and the intention to share towards information verification in an age of fake news. *Behavioral Sciences*, 12(2), 51. <a href="https://doi.org/10.3390/bs12020051">https://doi.org/10.3390/bs12020051</a>
- McGill University (n.d.). *English language proficiency*. Future Graduate Students.

  <a href="https://www.mcgill.ca/gradapplicants/how-apply/you-apply-mcgill/proficiency">https://www.mcgill.ca/gradapplicants/how-apply/you-apply-mcgill/proficiency</a>
- McGill University (n.d.). *English language proficiency*. Undergraduate Admissions. https://www.mcgill.ca/undergraduate-admissions/apply/english-proficiency
- McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology*, 89(3), 485–500. <a href="https://doi.org/10.1111/bjep.12279">https://doi.org/10.1111/bjep.12279</a>
- Milczarski, W., Borkowska, A., Paruzel-Czachura, M., & Białek, M. (2023). Using a foreign language does not make you think more: Null effects of using a foreign language on cognitive reflection

- and numeracy. *International Journal of Bilingualism*. https://doi.org/10.1177/13670069231206701
- Moore, R. C., & Hancock, J. T. (2022). A digital media literacy intervention for older adults improves resilience to fake news. *Scientific Reports*, 12(1), 6008–6008. <a href="https://doi.org/10.1038/s41598-022-08437-0">https://doi.org/10.1038/s41598-022-08437-0</a>
- Muda, R., Pennycook, G., Hamerski, D., & Białek, M. (2023). People are worse at detecting fake news in their foreign language. *Journal of Experimental Psychology. Applied*, (20230508). https://doi.org/10.1037/xap0000475
- Murrock, E., Amulya, J., Druckman, M., & Liubyva, T. (2018). Winning the war on state-sponsored propaganda: Results from an impact study of a Ukrainian news media and information literacy program. *Journal of Media Literacy Education*, 10(2), 53–85.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5), pp. 665-675.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175
- Niszczota, P., Pawlak, M., & Białek, M. (2022). Bilinguals are less susceptible to the bias blind spot in their second language. *International Journal of Bilingualism*, (20220711). https://doi.org/10.1177/13670069221110383
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1). https://doi.org/10.1177/1609406917733847

- OpenAI. (2023). ChatGPT (September 24 version) [Large language model]. https://chat.openai.com/
- Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology*, 47(6), 405–428. https://doi.org/10.1080/00207594.2012.743665
- Pennycook, G. (2023). A framework for understanding reasoning errors: From fake news to climate change and beyond. *Advances in Experimental Social Psychology*, 67, 131–208. https://doi.org/10.1016/bs.aesp.2022.11.003
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <a href="https://doi.org/10.1016/j.tics.2021.02.007">https://doi.org/10.1016/j.tics.2021.02.007</a>
- Pennycook, G., Binnendyk, J., Newton, C., & Rand, D. G. (2021). A practical guide to doing behavioral research on fake news and misinformation. *Collabra: Psychology*, 7(1). <a href="https://doi.org/10.1525/collabra.25293">https://doi.org/10.1525/collabra.25293</a>
- Pennycook, G., Cheyne, J. A., Koehler, D. & Fugelsang, J.A. (2020). On the belief that beliefs should change according to evidence: Implications for conspiratorial, moral, paranormal, political, religious, and science beliefs. *Judgment and Decision Making*, 15, 476–498.

  <a href="https://doi.org/10.1017/S1930297500007439">https://doi.org/10.1017/S1930297500007439</a></a>
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590–595. https://doi.org/10.1038/s41586-021-03344-2
- Potter, W. J. (2021). Media literacy. SAGE Publications, Incorporated.
- Potter, W. J. (2022). Analysis of definitions of media literacy. *Journal of Media Literacy Education*, 14(2), 27–43. https://doi.org/10.23860/JMLE-2022-14-2-3

- Preston, S., Anderson, A., Robertson, D. J., Shephard, M. P., Huhe, N., & Fernández-Berrocal, P. (2021).

  Detecting fake news on Facebook: the role of emotional intelligence. *Plos One*, *16*(3).

  <a href="https://doi.org/10.1371/journal.pone.0246757">https://doi.org/10.1371/journal.pone.0246757</a>
- Purpuri, S., Vasta, N., Filippi, R., Wei, L., & Mulatti, C. (2024). Does language shape the way we think?

  A review of the foreign language effect across domains. *International Journal of Bilingualism*.

  https://doi.org/10.1177/13670069231225374
- Purpuri, S., Vasta, N., Filippi, R., Wei, L., & Mulatti, C. (2024). The Foreign Language Effect on tolerance of ambiguity. *Bilingualism: Language and Cognition*, 27(1), 16–24. doi:10.1017/S1366728923000469
- Richards, K. (2009). Interviews. In Heigham, J. & Croker, R. A. (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp.182–199). Palgrave Macmillan. DOI: 10.1057/9780230239517
- Rojas, J. & Hernández-Fernández, J. (2018). The importance of English as a Second Language in Latin America. In J. Hernández-Fernández & J. Rojas (Eds.), *English public policies in Latin America:*Looking for innovation and systemic improvement in quality English language teaching (pp.10–30). British Council Mexico.
- Rubin, H. J., & Rubin, I. S. (2012). Structuring the interview. In *Qualitative Interviewing: The art of hearing data* (2nd ed., pp. 129–151). SAGE Publications, Inc. https://doi.org/10.4135/9781452226651
- Sarrion, E. (2023). Exploring the power of chatgpt: applications, techniques, and implications. Apress. <a href="https://doi.org/10.1007/978-1-4842-9529-8">https://doi.org/10.1007/978-1-4842-9529-8</a>

- Sousa, S., & Bates, N. (2021). Factors influencing content credibility in Facebook's news feed. *Human-Intelligent Systems Integration*, 1–10. https://doi.org/10.1007/s42454-021-00029-z
- Stamboliev, E. (2023). Proposing a postcritical AI literacy: Why we should worry less about algorithmic transparency and more about citizen empowerment. *Media Theory Journal*, 7(1), 201–232. <a href="http://mediatheoryjournal.org/">http://mediatheoryjournal.org/</a>
- Stankovic, M., Biedermann, B., & Hamamura, T. (2022). Not all bilinguals are the same: A meta-analysis of the moral foreign language effect. *Brain and Language*, 227, 105082–105082. https://doi.org/10.1016/j.bandl.2022.105082
- Statistics Canada. (2021). Census Profile, 2021 Census of Population. Government of Canada. <a href="https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E">https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E</a>
- Statistics Canada. (2023). A portrait of educational attainment and occupational outcomes among racialized populations in 2021. Government of Canada. <a href="https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021011/98-200-X2021011-eng.cfm">https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021011/98-200-X2021011-eng.cfm</a>
- Strauss, A. L., & Corbin, J. M. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). Sage Publications.
- Teitelbaum Dorfman, F., Kogan, B., Barttfeld, P., & García, A. M. (2024). Bilinguals on the footbridge:

  The role of foreign-language proficiency in moral decision making. *Bilingualism: Language and Cognition*, 1–16. https://doi.org/10.1017/S1366728924000312
- Vanek, N., & Tovalovich, A. (2022). Emotionality ratings and electrodermal responses to university-related expressions in a native and a non-native language. *International Journal of Bilingual Education and Bilingualism*, 25(8), 2817–2833. https://doi.org/10.1080/13670050.2021.1978924

- Vives, M.L., Aparici, M., Costa, A., & Sigman, M. (2018). The limits of the foreign language effect on decision-making: The case of the outcome bias and the representativeness heuristic. *Plos One*, 13(9). <a href="https://doi.org/10.1371/journal.pone.0203528">https://doi.org/10.1371/journal.pone.0203528</a>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <a href="https://doi.org/10.1126/science.aap9559">https://doi.org/10.1126/science.aap9559</a>
- Wu, Y., Mou, Y., Li, Z., & Xu, K. (2020). Investigating American and Chinese subjects' explicit and implicit perceptions of ai-generated artistic work. *Computers in Human Behavior*, 104.
  <a href="https://doi.org/10.1016/j.chb.2019.106186">https://doi.org/10.1016/j.chb.2019.106186</a>
- Zhou, C., Li, K., & Lu, Y. (2021). Linguistic characteristics and the dissemination of misinformation in social media: the moderating effect of information richness. *Information Processing and Management*, 58(6). <a href="https://doi.org/10.1016/j.ipm.2021.102679">https://doi.org/10.1016/j.ipm.2021.102679</a>
- Zhou, X., & Zafarani, R. (2021). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, *53*(5), 109.

## **Appendixes**

## Appendix A

AI-generated Facebook post in Spanish about a recent graduate job in Montreal (150-250 words).



♠ ¡OPORTUNIDAD DE EMPLEO EN MONTREAL PARA GRADUADOS!



¡Estamos emocionados de ofrecer una oportunidad única para graduados recientes!

Montreal Tech Solutions está buscando un(a) Asociado(a) de Marketing Digital apasionado(a) para unirse a nuestro equipo innovador. Si eres un recién graduado con una pasión por el marketing digital y estás listo para enfrentar nuevos desafíos, ¡queremos ser parte de tu viaje profesional!

📍 Ubicación: Montreal, QC

illimitation Tipo de contrato: Tiempo completo

& Rango salarial: \$45,000 - \$55,000 anuales (según experiencia)

En Montreal Tech Solutions, creemos en el potencial de los graduados y ofrecemos un entorno de trabajo colaborativo que fomenta el crecimiento y la creatividad. Además, brindamos oportunidades de desarrollo profesional y un paquete de beneficios integral.

Para más detalles y para aplicar, visita nuestra página web: www.montrealtechsolutions.com/trabajo

Fecha de cierre de la aplicación: Primero de Abril

¡Estamos emocionados de conocer a graduados talentosos y apasionados que deseen unirse a Montreal Tech Solutions! Comparte esta oportunidad con tus compañeros de clase y amigos. ¡No pierdas la oportunidad de comenzar tu carrera en un entorno estimulante! \*#
#EmpleoGraduados #OportunidadLaboral #TrabajoMontreal

ك Like	○ Comment	Send Send
LINC.	C comment	O sella

AI-generated Facebook post in Spanish about housing for students in Montreal (150-250 words).



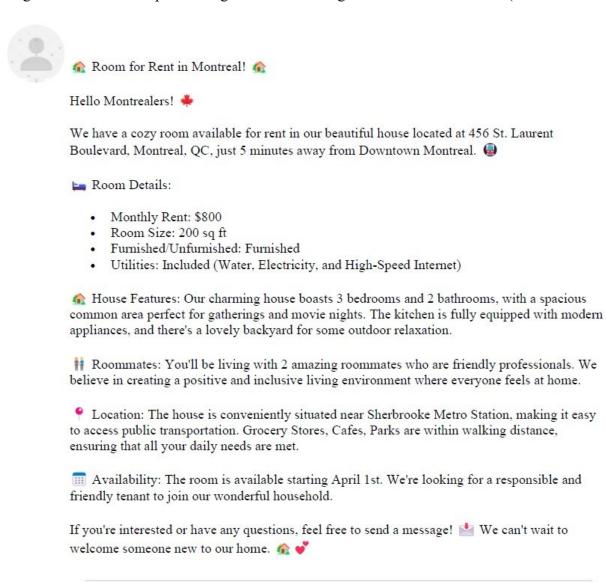
- 🏠 ¡Oportunidad de alquiler en el corazón de Montreal! 🏠
  - 🌞 Habitación Disponible para Alquiler 🌞

¡Hola Montrealenses! Tenemos una joya de espacio disponible en nuestro acogedor hogar y estamos emocionados de encontrar el compañero de cuarto perfecto para compartirlo.

- Ubicación: 123 Rue Belle, Montreal
- Habitación: Amplia habitación con mucha luz natural y armario empotrado.
- Casa: Hermosa casa con cocina totalmente equipada, sala de estar cómoda y dos baños.
- Tarifa de alquiler mensual: \$800 (incluye servicios públicos y Wi-Fi)
- iii Disponibilidad: La habitación estará disponible a partir del 1 de abril.
- Compañeros de cuarto: Somos un grupo amigable y respetuoso de profesionales jóvenes que disfrutan de la buena compañía y la convivencia armoniosa. Todos nos tomamos en serio nuestras responsabilidades y valoramos un hogar limpio y ordenado.
- Proximidad: A poca distancia de tiendas locales, restaurantes y parques encantadores. ¡La estación de metro más cercana está a solo 5 minutos a pie, facilitando tus desplazamientos por la ciudad!

△ Like	Comment	₹ Send
--------	---------	--------

AI-generated Facebook post in English about housing for students in Montreal (150-250 words).



Comment

ל' Like

Y/ Send

AI-generated Facebook post in English about a job opportunity for students in Montreal (150-250 words).



Are you a recent graduate eager to dive into the world of customer success? Join MetroHealth Solutions in Montreal as a Customer Success Associate! As a vital part of our team, you'll be the bridge between our clients and our healthcare solutions. Your responsibilities will include onboarding, issue resolution, and collaborating with internal teams.

## Qualifications:

- · Recent graduate in Business, Health Sciences, or related field
- Strong communication and problem-solving skills
- · Tech-savvy and adaptable to new systems
- Collaborative mindset in a fast-paced environment

## Why MetroHealth?

- Competitive salary at \$48,000 p.a.
- · Professional growth opportunities
- Inclusive and collaborative work culture
- Convenient Montreal office location

How to Apply: Ready to launch your career with MetroHealth Solutions? Send your resume and cover letter to careers@metrohealth.com. Subject: "Customer Success Associate Application - [Your Full Name]".

## Deadline: 1st of April

Don't miss out on this fantastic opportunity! Share with friends and tag those interested. Let's shape the future of healthcare together! 
#MontreaUobs #CustomerSuccess #CareerOpportunity

0	O =	
△ Like	Comment	Send Send

## Appendix B

## English job post prompt

Anonymous

create a 150-1250 words post resembling a Facebook job post appearing in Montreal based groups. The post should include specific details: company's name, office location, salary (according to the standards in Montreal based Facebook groups), and the contract type. The job is for graduate students.

ChatGPT

🌞 Exciting Job Opportunity for Graduate Students in Montreal! 🌞

## English rental post prompt

Anonymous

Please generate a post of 150-250 words about a room for posting on Facebook. The post needs to resemble a Facebook post about a room for rent in Montreal-based groups. The post should include specific details like: monthly rent rate, the exact address, a description of the house and the room, brief information about the roommates, proximity to facilities and public transportation. The post's length should represent the average length of a room for rent post on Facebook in Montreal groups.

ChatGPT

🞪 Room for Rent in Montreal! 🎪

## Spanish Rental post prompt

Anonymous

Genere una publicación de 150 a 250 palabras sobre una sala para publicar en Facebook. La publicación debe parecerse a una publicación de Facebook sobre una habitación en alquiler en grupos con sede en Montreal. La publicación debe incluir detalles específicos como: tarifa de alquiler mensual, dirección exacta, descripción de la casa y la habitación, información breve sobre los compañeros de habitación, proximidad a las instalaciones y transporte público. La longitud de la publicación debe representar la duración promedio de una publicación de alquiler de habitación en Facebook en grupos de Montreal.

ChatGPT

🎪 ¡Oportunidad de alquiler en el corazón de Montreal! 🎪

## Spanish job post prompt



## **Anonymous**

Genere una publicación de 150 a 250 palabras sobre una oferta de trabajo en Facebook en español. La publicación debe parecerse a una publicación de Facebook sobre una oferta de trabajo en grupos con sede en Montreal. La publicación debe incluir detalles específicos como: rango salarial (ya sea anual o por horas, lo que sea más común en los grupos de Facebook de Montreal), incluir la URL de la empresa, el tipo de contrato y la ubicación de la oficina de la empresa. La duración de la publicación debe representar la duración promedio de una publicación de trabajo en Facebook en los grupos de Montreal.

## ChatGPT



Descriptive Statistics (Medians and Interquartile Ranges) for Representativeness and Quality and Employment–Rental Post Comparisons (Wilcoxon Tests) Within Language

Appendix C

	Spani	ish	Comp	arison	Engli	sh	Comp	arison
Rater	Employment	Rental	Z	p	Employment	Rental	Z	p
			R	epresen	tativeness			
Preraters	3.5 (3.00)	4.5 (5.25)	0.92	.357	5.0 (2.00)	6.0 (1.00)	1.86	.063
Participants	4.5 (3.00)	6.0 (2.00)	2.96	.003	5.0 (3.25)	7.0 (2.00)	2.19	.029
				Qua	ality			
Preraters	4.5 (2.25)	5.5 (2.50)	0.45	.655	6.0 (2.00)	6.0 (0.00)	1.41	.157
Participants	4.5 (3.75)	5.0 (1.00)	3.00	.003	5.0 (3.75)	6.0 (2.00)	2.14	.032

## Appendix D

## Preguntas para la encuesta

- 1. Como les fue el experimento? ¿Les resulto facil o dificil?
- 2. ¿Que pensaron sobre la parte donde tenían que resaltar lo que consideran como más creíble o menos creíble?
- 3. ¿Como hicieron la distinción entre lo que consideraron más creíble y menos creíble?
- 4. ¿Porque marcaron esta frase en verde?/ ¿ porque marcaron esta frase en rojo?
- 5. ¿Que tiene esta frase/palabra que les hace pensar que es más creíble o menos creíble?
- 6. ¿En su uso diario de las redes sociales, como definen si algo es más creíble o menos creíble?
- 7. ¿Califica la credibilidad de manera diferente en inglés que en español?
- 8. ¿Que sabes a cerca de la inteligencia artificial?
- 9. ¿Cree que la inteligencia artificial puede contribuir o perjudicar la credibilidad del contenido online?
- $10.\ {}_{\hbox{$\rlap{\'e}$}}$  Crees que las publicaciones que has visto fueron escritas por humanos o generadas por IA?

## **Interview Questions**

- 1. How did the experiment go for you? Did you find it easy or difficult?
- 2. What were your thoughts on the part where you had to highlight what you thought to be more credible or less credible?
- 3.How did you make the distinction between what you considered to be more credible or less credible?"
- 4. Why did you highlight this sentence in green? / Why did you highlight this sentence in red?
- 5. What does this sentence/word have that makes you think it is more credible or less credible?
- 6. In your daily use of social media, how do you decide if something is more credible or less credible?
- 7. Do you judge credibility differently in English versus Spanish?
- 8. What do you know about artificial intelligence?
- 9. Do you think artificial intelligence can contribute to or harm the credibility of online content?
- 10. Do you think the posts you have seen were written by humans or generated by AI?

## Appendix E

Por favor escriba su numero de participante	e (esta al costa	ado de la compu	tadora portátil)
por favor indique su fecha de nacimiento			
	Mes	día	año
Por favor indique:	~		~
¿Cuál es su género?			
Femenino			
Masculino			
Transgénero			
No binario			
Prefiero no contestar			
Por favor escriba el nombre de su país de o	origen		

	¿vives en Canada actualmente?	
	Si	
	No	
	¿Es el español tu lengua materna?	
	No	
	Si	
	$\rightarrow$	
En que	e universidad estudias?	
Conco	ordia University	
McGill	I University	

¿En

¿Cuál es tu actual grado de escolaridad? Ninguno Primaria Secundaria Bachillerato BA/BsC MA/MsC PhD ¿Cómo definirías tu nivel de inglés? Avanzado Intermedio básico ¿Has tomado algun examen de Ingles para ingresar a su universidad en Montreal? Si

No

Usted ha indicado que tomo el examen de ¿ Cual fue su nota total en el examen?

¿Cuántos años has estudiado Inglés?

	Menos de un año	
	Un año	
	2 años	
	3 años	
	4 años	
	5 años	
	6 años	
	7 años	
	8 años	
	9 años	
	10 años	
	Mas de 10 años	
¿Cuál es el porce	entaje promedio en que usas los siguientes idiomas en tu	día-día?
Español		0
Inglés		0
Francés		0
Otros idiomas		0
Total		0

## ¿Cuánto tiempo usas Facebook al día?

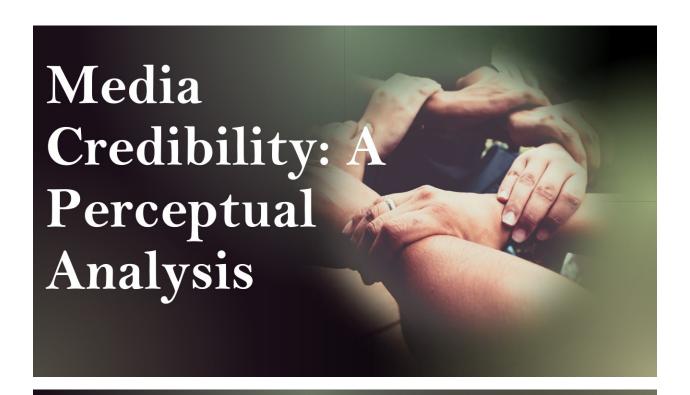
Menos de una hora
1-2 horas
3-4 horas
Mas de 5 horas
¿Cuánto tiempo usas Facebook <b>a la semana</b> ?
1-2 horas
3-5 horas
5-7 horas
7-9 horas
Mas de 10 horas
¿Para qué utilizas Facebook principalmente? ordena las frases de forma ascendente según la frecuencia con la que las usas.
compartir videos, imágenes, historias
Mantenerse en contacto con amigas y familiares
Buscar trabajo
Buscar vivienda
Buscar servicios
Para ver noticias

estás en Facebook?	
compartir videos, imágenes, historias	0
Mantenerse en contacto con amigos y familiares	0
Buscar trabajo	0
Buscar vivienda	0
Buscar servicios	0
Para ver noticias	0
Total	0
¿Qué porcentaje del contenido que ves en Facebook está en los siguientes idiom	ias?
Español	0
Inglés	0
Francés	0
Otros idiomas	0
Total	0

¿Cuál es el tiempo proporcional que dedicas a cada una de estas actividades mientras

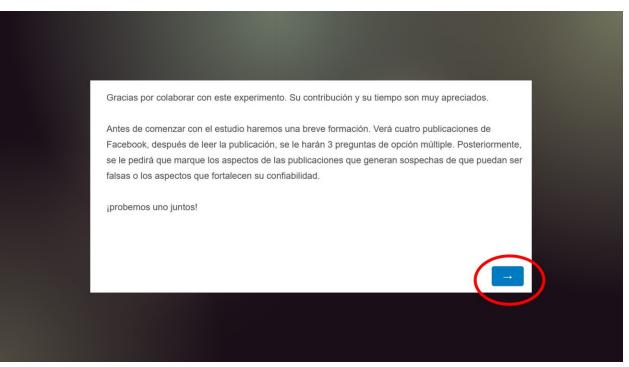
We thank you for your time spent taking this survey. Your response has been recorded.

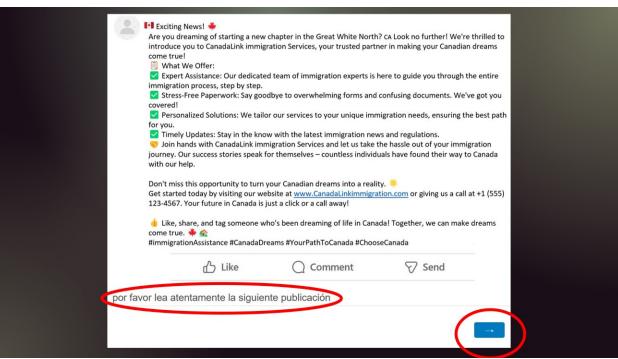
## Appendix F



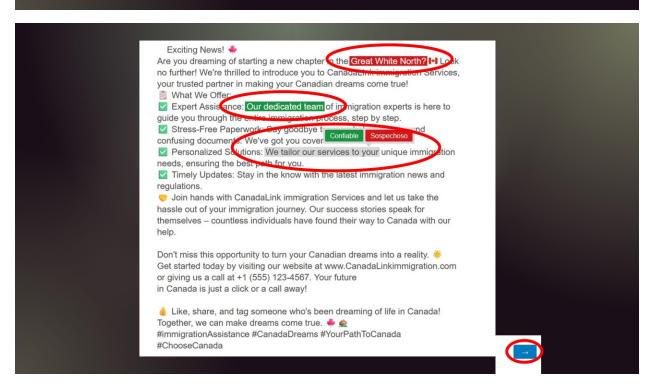
## El plan para hoy

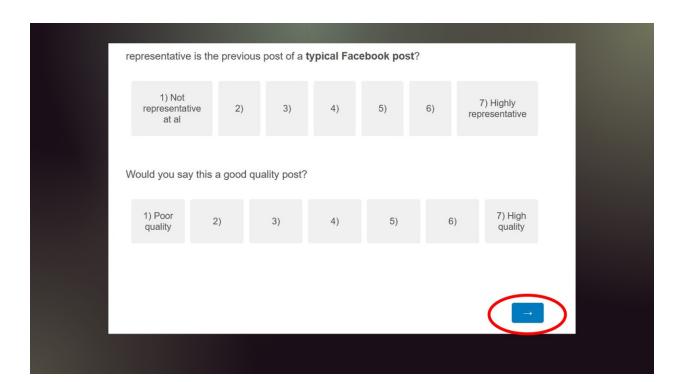
- Breve introducción (2-3 minutos)
- Prueba (3-5 minutos)
- Experimento (20-30)
- Entrevista (15 minutos)
- Resumen (5 minutos)





How credible is t	he previo	us post?				
1) Not at all credible	2)	3)	4)	5)	6)	7) extremely crdible
Would you share	this post	on Facebook	?			
1) Not at all	2)	3)	4)	5)	6)	7) Sure





Gracias por responder la primera encuesta. Ahora pasaremos a la siguiente encuesta.

Para continuar, por favor haga click en el siguiente link:

https://concordia.yul1.qualtrics.com/jfe/form/SV\_cSezn6EIDh9DDtY

¿alguna pregunta?

## FORMULARIO DE INFORMACIÓN Y CONSENTIMIENTO



## Appendix G

Gracias por colaborar con este experimento. Su contribución y su tiempo son muy apreciados. Antes de comenzar con el estudio haremos una breve formación. Verá cuatro publicaciones de Facebook, después de leer la publicación, se le harán 3 preguntas de opción múltiple. Posteriormente, se le pedirá que marque los aspectos de las publicaciones que generan sospechas de que puedan ser falsas o los aspectos que fortalecen su confiabilidad. ¡probemos uno juntos!

Thank you for collaborating with this experiment. Your contribution and time are greatly appreciated. Before starting the study we will do a brief training. You will see four Facebook posts, after reading the post you will be asked 3 multiple choice questions. Subsequently, you will be asked to mark the aspects of the publications that raise suspicion that they may be false or the aspects that strengthen their reliability. Let's try one together!

Gracias por colaborar con este experimento. Su contribución y su tiempo son muy apreciados.

Antes de comenzar con el estudio haremos una breve formación. Verá cuatro publicaciones de Facebook, después de leer la publicación, se le harán 3 preguntas de opción múltiple. Posteriormente, se le pedirá que marque los aspectos de las publicaciones que generan sospechas de que puedan ser falsas o los aspectos que fortalecen su confiabilidad.

¡probemos uno juntos!

 $\rightarrow$ 

## Appendix H



## Appendix I

Dear participant,

Thank you for participating in the study, we highly appreciate you taking the time and effort to help us with our scientific work. As follow up to the experiment, please find below a few resources that can help to overcome a fraud or scam attempt:

- Concordia's IT services Security & Privacy page: <a href="https://www.concordia.ca/it/security.html">https://www.concordia.ca/it/security.html</a>
   What is it? A great resource for learning how to identify and protect oneself from online scammers, as well as best practices to adopt in case you were targeted. It also provide relevant contact information.
- 2. Report a fraud (a federal resource): <a href="https://www.antifraudcentre-centreantifraude.ca/report-signalez-eng.htm">https://www.antifraudcentre-centreantifraude.ca/report-signalez-eng.htm</a>
- 3. SPVM page on fraud reporting: https://spvm.qc.ca/en/Seniors/Fraud--
- 4. Concordia's Counselling & Psychological services: <a href="https://www.concordia.ca/health/mental-health/counselling.html">https://www.concordia.ca/health/mental-health/counselling.html</a>
- 5. The link to the English version of *The Little Black Book of Scams*, with best practices to protect yourselves from scams online: <a href="https://publications.gc.ca/collections/collection\_2018/isde-ised/lu54-42-2018-eng.pdf">https://publications.gc.ca/collections/collection\_2018/isde-ised/lu54-42-2018-eng.pdf</a>

Once again thank you for your collaboration, we wish you a great rest of your day!

Noga Broitman,

Master's student, Concordia department of Education

## Appendix J

In this experiment, all posts that we presented to you were created using ChatGPT, a language model developed by OpenAI. This shows that AI tools like ChatGPT can produce a realistic post that anyone is likely to see online. We chose not to tell you that the posts that you saw during the experiment were created by ChatGPT for one important reason. We wanted to see how bilingual speakers like you—those who speak Spanish and English—might react to such realistic-looking content online without knowing that the content is created by ChatGPT. In other words, we wished to re-create a possible situation, where readers like you and me can encounter a post online that is created through ChatGPT, not by a human.

As you can imagine, ChatGPT can be used for all kinds of illegal actions, including online frauds and scams, to create realistic-looking content. So it could be challenging for internet users like you and me to distinguish between authentic and manipulated content. Therefore, one challenge for researchers and educators is to understand how to train people to become more careful readers of online information, so that they can distinguish information that is more credible versus less credible.

- Do you have any questions about the purpose of our study?
- Do you have any concerns about us using your data?

Because it is important for people to recognize frauds and scams when they occur, we would like to provide you with further information which can help you combat the risks of online scams, including those that might rely on ChatGPT.

First, here is some information about online fraud statistics in Canada:

# The impact of fraud so far this year As of December 31, 2023 Reports processed: 62,365 (92,078 in 2022) Victims of fraud: 41,111 (57,578 in 2022) Lost to fraud: \$554 M (\$531 M in 2022)

The Canadian Anti-Fraud Centre (CAFC) website includes updated data and recommendations on how to protect yourself. We suggest visiting the site regularly for data and advice: <a href="https://www.antifraudcentre-centreantifraude.ca/index-eng.htm">https://www.antifraudcentre-centreantifraude.ca/index-eng.htm</a>

The government of Canada also provides a resource in the form of a booklet that includes the most common scams in the country and tips to prevent them. The book is called "The Little Black Book of Scams." You can find the version of this book in various languages, including English and Spanish (available in English in the following link): <a href="https://publications.gc.ca/collections/collection\_2018/isde-ised/lu54-42-2018-eng.pdf">https://publications.gc.ca/collections/collection\_2018/isde-ised/lu54-42-2018-eng.pdf</a>



Lastly, we would kindly like to ask you to **avoid** disclosing or discussing the purpose or procedure of this experiment with participants who haven't taken the experiment yet. We appreciate your understanding, and thank you for your collaboration!

En este experimento, todos los mensajes que le presentamos fueron creados utilizando ChatGPT, un modelo de lenguaje desarrollado por OpenAl. Esto demuestra que herramientas de inteligencia artificial como ChatGPT pueden producir mensajes realistas que cualquiera podría ver en línea. Optamos por no decirle que los mensajes que vio durante el experimento fueron creados por ChatGPT por una razón importante. Queríamos ver cómo hablantes bilingües como usted, aquellos que hablan español e inglés, podrían reaccionar ante contenido con apariencia realista en línea sin saber que el contenido fue creado por ChatGPT. En otras palabras, deseamos recrear una situación posible, donde lectores como usted y yo pueden encontrar un mensaje en línea creado a través de ChatGPT, no por un humano.

Como puede imaginar, ChatGPT puede ser utilizado para todo tipo de acciones ilegales, incluyendo fraudes y estafas en línea, para crear contenido con apariencia realista. Por lo tanto, podría ser desafiante para usuarios de Internet como usted y yo distinguir entre contenido auténtico y manipulado. Por lo tanto, un desafío para los investigadores y educadores es entender cómo entrenar a las personas para que sean lectores más cuidadosos de la información en línea, de modo que puedan distinguir entre información más creíble y menos creíble.

- ¿Tiene alguna pregunta sobre el propósito de nuestro estudio?
- ¿Tiene alguna preocupación sobre el uso de sus datos?

Dado que es importante que las personas reconozcan fraudes y estafas cuando ocurren, nos gustaría proporcionarle más información que puede ayudarlo a combatir los riesgos de estafas en línea, incluidas aquellas que podrían depender de ChatGPT.

Primero, aquí hay información sobre las estadísticas de fraude en línea en Canadá:

# The impact of fraud so far this year As of December 31, 2023 Reports processed: 62,365 (92,078 in 2022) Victims of fraud: 41,111 (57,578 in 2022) Lost to fraud: \$554 M (\$531 M in 2022)

El sitio web del Canadian Anti-Fraud Centre (CAFC) incluye datos actualizados y recomendaciones de cómo protegerse, le sugerimos visitar el sitio regularmente para obtener datos y consejos: https://www.antifraudcentre-centreantifraude.ca/index-eng.htm

El gobierno de Canadá también ofrece un recurso en forma de una liberta que incluye las estafas más comunes en el país y consejos para prevenirlos. El libro se llama "El pequeño libro negro de las estafas". Podrán encontrar la versión de este libro en diversas lenguas, entre ella en ingles e español (disponibles en el siguiente link): <a href="https://ised-isde.canada.ca/site/competition-bureau-canada/en/little-black-book-scams-2nd-edition">https://ised-isde.canada.ca/site/competition-bureau-canada/en/little-black-book-scams-2nd-edition</a>



Por último, nos gustaría pedirle que evite revelar o discutir el propósito o procedimiento de este experimento con participantes que aún no lo hayan realizado. ¡Agradecemos su comprensión y gracias por su colaboración!