

Registered Report: How Language Shapes Belief in Misinformation: A Study Among Multilingual Speakers in Ukraine

Kevin Aslett^a, Aaron Erlich^b, Jonathan Nagler^{a,c}, and Joshua A. Tucker^{a,c}

^aCenter for Social Media and Politics, New York University

^bDepartment of Political Science, McGill University

^cWilf Family Department of Politics, New York University

1 Introduction

Our cumulative knowledge about belief in misinformation predominantly comes from surveying English-speaking Americans about misinformation written in English from American media sources (Allcott and Gentzkow 2017; Pennycook and Rand 2020; Clayton et al. 2019; Pennycook, Cannon, and Rand 2018).¹ However, the global media environment is complexly multilingual. Half of the global population uses two or more languages or dialects in their daily life (Ansaldo et al. 2008; Grosjean 2010) and, therefore, likely consumes media, including misinformation, in multiple languages from both within and outside of their borders. As consumption of cross-border media increases and multilingual media with it (PwC-UK 2016), the distribution of false or misleading news in different languages poses substantial political consequences. Reporting news in different languages has the capacity to differentially mobilize populations (Onguny 2019), sometimes on the basis of false or misleading evidence that escalates political violence (Ismail and Deane 2008). Differential belief in misinformation strengthens ethno-political divides (Somerville 2009) and increase levels of affective polarization (Lau et al. 2017; Suhay, Bello-Pardo, and Maurer 2018; Stewart, Plotkin, and McCarty 2021), which weaken the foundations of liberal democracy (Kuklinski et al. 2000). If misperceptions are based on ethnic divisions, policies advocated by politicians will likely accentuate ethnic and economic inequality and may, in some cases, lead to the disenfranchisement of minority groups. Given both these troubling consequences for democracy and the rise of multilingual speakers in a single media market consuming news in different languages, it is imperative we develop a more comprehensive understanding of how news consumers perceive misinformation in different languages. In this registered report we propose a study that addresses the broad research question: Are individuals more or less susceptible to misinformation in their non-dominant language?

¹There are a few, but growing list of exceptions (Mujani and Kuipers 2020; Badrinathan 2020; Rosenzweig et al. 2021)

Specifically, we plan to test whether credibility cues associated with a specific language affect belief in misinformation. Previous research on language and belief in misinformation has focused on identifying a language proficiency effect, but we propose that language also signals *information credibility* which may affect belief in misinformation written in that language. For example, a minority group with a distinct language may be more skeptical of news that is written in a different language spoken by the majority, because they have become distrustful of media that communicate using that language. Credibility cues unique to a language and context, such as these, may explain why past studies measuring the effect of language proficiency on belief in misinformation have reported inconsistent results across different languages.

We propose testing whether domestic or international factors do indeed create language credibility cues that individuals use when evaluating misinformation by conducting a survey experiment in Ukraine. In our proposed survey experiment we will randomly assign bilingual respondents to evaluate false and true news stories in either their dominant or less preferred language (Russian or Ukrainian) directly after the publication of an article. We seek to answer three non-mutually exclusive research questions: (1) In places in which most news is reported in two languages, are multilingual individuals who prefer to consume media in the minority language more skeptical of misinformation produced in their less preferred language (the dominant language)? (2) Are multilingual individuals more skeptical of misinformation produced in their less preferred language if that language is the same as a foreign power associated with a foreign disinformation campaign? (3) Does this “external power” language effect exist in areas that disinformation campaigns specifically target?

To present our expected results for this pre-registered experiment we use data from a 1-week pilot study to simulate the results of a full, (proposed) 10-week study. Using this pilot data we present the predicted effects of reading misinformation in one’s less preferred language for different credibility cues as well as how our results will be presented in the final paper once we complete the full ten-week study. Our simulated data suggests that minority groups may be more skeptical of misinformation written in a majority language. It also suggests that individuals are not more skeptical of misinformation written in the language of a foreign power associated with disinformation campaigns, unless they live in areas specifically targeted by these disinformation campaigns.

This registered report consists of four more sections. The next section presents the theory that motivates our research questions and hypotheses. The final three sections present the experimental design, proposed methods and expected results. The discussion of the results will be added upon the completion of our proposed study.

2 Theory and Hypotheses

Much of the work on language and misinformation has focused on identifying a language proficiency effect by investigating two modes of cognition: an effortless mode based on heuristics and a more reflective mode based on deliberation (Keysar, Hayakawa, and An 2012; Costa, Vives, and Corey 2017; Corey et al. 2017). Muda et al. (2021) have suggested that individuals who read news in their less proficient language often use their more reflective mode of cognition and rely less on heuristics. Given that higher levels of cognitive reflection lowers individuals' belief in misinformation (Pennycook and Rand 2019), individuals who evaluate misinformation in their less proficient language could be less likely to believe misinformation, but studies testing language's effect on the belief in misinformation report mixed evidence (Fernández-López and Perea 2020; Muda et al. 2021). Inconsistent results suggest that reading news in one's less preferred language may not have the same effect across each language and context. Features of news identified by other work, such as credibility cues (Flanagin and Metzger 2000; Althaus and Tewksbury 2000; Fogg et al. 2001; Flanagin and Metzger 2007) may explain why some individuals are more susceptible to believing misinformation in their non-native language than others. Building on this work we move past this proficiency argument and focus on what language signals in terms of information credibility. Credibility cues associated with a language can be particularly influential for those reading in their less preferred language, because individuals are less familiar with news in this language and are therefore more likely to rely on these credibility cues when evaluating suspect information. We therefore focus on the effect of evaluating misinformation in one's less preferred language relative to their dominant language.

Language may pass on different credibility cues for numerous reasons. For example, language may pass on a cue to the source of the information, which could affect belief in misinformation. News consumers often use the source of information to aid their evaluation of news stories' veracity (Sundar and Nass 2001). Moreover, the information source's country of provenance (which is often imparted by language) can change the effect of information on public opinion (Dragojlovic 2015).² Language also has particularly strong priming effects (Pérez 2016), particularly on bilingual and bicultural individuals (Ross, Xun, and Wilson 2002; Trafimow et al. 1997).³ In a political context, speaking in a language associated with a minority population can increase the political salience of ethnic divisions (Pérez and Tavits 2019) and the language of political ads can pass on cues to voters that affect their sentiment towards the targeted political candidate (Flores and Coppock

²Unsurprisingly, Dragojlovic (2015) finds that individuals are more likely to be persuaded by new information from foreign sources that they view more positively. Other work has not found any foreign source effect in any direction when individuals have a negative view of the foreign source (Fisher 2020)

³Using a different language also has a powerful effect on people's interpretations of the world (Boroditsky 2006), time (Boroditsky 2001) or even the same event (Fausey et al. 2010).

2018). The effects of language on consuming information clearly depend on the contextual cues given by domestic or international conditions. In the following two sections we develop the theoretical frameworks for two possible language credibility cues that could affect belief in misinformation: (1) the *language of the majority* effect caused by domestic conditions and (2) the *language of external power* caused by international conditions.

2.1 Language of the Majority (Domestic Conditions)

Within a country, minority groups (often ethnic groups) with distinct languages are often skeptical of news that is written in the language spoken by the majority, because mainstream news in the majority language often portray minority groups in a negative light and over-represent the dominant group that prefers to speak the majority language (Keshishian 2000; Mastro 2009; Tukachinsky, Mastro, and Yarchi 2015). This skepticism can push those in the linguistic minority to consume sectoral or extranational media that often use different languages (Tsfati and Peri 2006). Minority groups' divergence in trust can gain prominence during periods of crisis. For example, after the 2014 Ukrainian crisis trust in majority language news dropped dramatically among minority Russian speakers in Estonia and Latvia (Vihalemm, Juzefovičs, and Leppik 2019). Indeed, it is likely that during a crisis mainstream news in the majority language can promote separate narratives that alienate minorities already skeptical of news in the majority language. Therefore, in such situations, we expect that those who prefer to speak a minority language are less likely to believe misinformation if it is written in the majority language spoken (ie., their less-preferred language).

In Ukraine, we plan to test this “language of the majority” hypothesis among those that prefer to speak Russian and measure the effect reading misinformation in their less-preferred language, Ukrainian, has on belief in that misinformation. Although the status of Russian is debated in Ukraine, over 40% report Russian as their language of private life and take opinion surveys in Russian, while 60% report Ukrainian as their preferred language (Onuch and Hale 2018). Generally, almost the entire population of Ukraine report high reading proficiency in both languages, given their linguistic similarity.⁴ Given their minority-status in Ukraine, we expect Russian-dominant speakers in Ukraine to be more skeptical of misinformation written in Ukrainian than in Russian.

(H1) Respondents who are *Russian-dominant* speakers in Ukraine are less likely to believe a false/misleading article written in their less preferred language (Ukrainian) than in their more

⁴In our pilot study, 86% of our respondents living in Ukraine (230 respondents evenly divided Russian and Ukrainian speaking preferences) self-reported equivalent reading proficiency levels in Ukrainian and Russian.

preferred language (Russian)

2.2 Language of External Power (International Conditions)

International conditions could also create language credibility cues. For example, a sustained disinformation campaign by a foreign power that employs a domestic minority language could lower belief in misinformation written in the language. Often disinformation campaigns target a minority ethnic group that uses their own preferred language (Snegovaya 2015) that differs from the majority-spoken language. Reading news in the language associated with a foreign disinformation campaign may cue lower credibility among those who speak the majority language and prime those that do not prefer to speak this language to look at the information more skeptically. If this is the case, we would expect that those who prefer to speak the language of the majority would be less likely to believe misinformation written in the “language of external power” than if it was written in the majority language.

Russian language news in Ukraine fits these criteria (StratCom 2015): Ukrainian news consumers who are most proficient in Ukrainian (60 percent) (Vorobiov 2015) and normally consume news in Ukrainian may associate news in Russian with a foreign actor’s disinformation campaign targeting the linguistically-Russian diaspora in Ukraine. This credibility cue could reduce belief in reported information in that language. Previous research has partially identified this phenomenon and found that those in Ukraine that prefer to speak Ukrainian are less likely to believe popular pro-Kremlin disinformation news stories (which are usually but not exclusively written in Russian) than those that prefer to speak Russian (Erich and Garner 2021).⁵ We, therefore, test this “language of external power” hypothesis among those that prefer to speak Ukrainian in Ukraine and measure the effect reading misinformation in their less-preferred language, Russian, has on belief in that misinformation. We expect that Ukrainian-dominant speakers in Ukraine are more skeptical of misinformation written in Russian (the language of the foreign disinformation campaigns) than in Ukrainian.

(H2) Respondents who are *Ukrainian-dominant* are less likely to believe a false/misleading article written in their less preferred language (Russian) than in their more preferred language (Ukrainian)

We also test if this hypothesis holds solely among Ukrainian-dominant speakers that live in areas targeted by Russian disinformation (those in Southern and Eastern Ukraine) where this “language of external power”

⁵However, this work does not randomly assign respondents to the language in which they consume news stories or test the proficiency side of the theory we suppose.

effect should be strongest.⁶

(H3) Respondents who are *Ukrainian-dominant* speakers in the Southern and Eastern regions Ukraine are less likely to believe a false/misleading article written in their less preferred language (Russian) than a false/misleading article in their more preferred language (Ukrainian)

In addition to testing these three hypotheses we will also run four exploratory analyses that: (1) determine if the effects we find hold across articles with different ideological perspectives, (2) test each hypotheses using strictly evaluations of true articles, (3) test if there is a measurable difference in the “majority language” effect (H1) and the “foreign power” effect (H2), and (4) determine if the effects we find hold across all articles. A full description of these exploratory analyses is located in the Supplementary Materials, Section L.

3 Experimental Design

To test our hypotheses we will carry out a 10-week survey experiment in Ukraine. Ukraine is an ideal case to test both the “language of the majority” and “language of external power” hypotheses, because: (1) The majority of its citizens are bilingual news consumers in Ukrainian and Russian, but prefer one language; (2) News consumed in Ukraine is produced in both Ukrainian and Russian, but the majority of news is in Ukrainian; (3) Russian disinformation campaigns in Ukraine heavily target the Russian-speaking diaspora (Snegovaya 2015) through Russian language news websites and media companies.⁷

Each week, we will sample Ukrainian respondents to ensure geographic, ideological, and linguistic balance among our respondents. Information on who we sample can be found in the Supplementary Materials, Section D. Human subjects in Ukrainian samples will be recruited by *Qualtrics*.⁸ Previous research measuring belief in misinformation has yet to integrate important findings about how individuals consume misinformation, limiting inference from these studies. Specifically, misinformation is consumed very quickly after publication (Vosoughi, Roy, and Aral 2018; Starbird et al. 2018), but most research asks respondents to evaluate months- or years-old fact-checked, misinformation (Bronstein et al. 2019; Clayton et al. 2019; Pennycook and Rand 2020). To address this limitation, we create a transparent, replicable, and pre-registered news article selection

⁶Given that Russian disinformation targets Russian speakers in Ukraine (Snegovaya 2015), it is likely that this “foreign power” effect is strongest in areas where Russian is most often use. These oblasts we identify from Ukrainian census data (SSCU 2001 and are delineated in Section E, Supplementary Materials

⁷For example, the Russian-controlled news agency Sputnik has engaged in the spread of political misinformation (Satariano 2019).

⁸Section D, Supplementary Materials details the recruitment of respondents.

process that sources popular false/misleading and true articles within 24 hours of their publication and then distributes the full articles for evaluation to respondents in Ukraine. Our respondents evaluate these popular articles within 72 hours of publication. This process ensures that we are measuring the effect of language on belief in popular misinformation in the time period that individuals are most likely to consume this misinformation.

Our selection method also removes potential researcher selection effects. Past studies investigating interventions affecting belief in misinformation have asked survey respondents to evaluate either popular pieces of misinformation chosen by the researcher (Allcott and Gentzkow 2017; Bronstein et al. 2019; Clayton et al. 2019; Pennycook and Rand 2020), or researcher-designed synthetic news articles (Pennycook, Cannon, and Rand 2018; Moravec, Minas, and Dennis 2018). Both methods risk article selection effects and sampling articles that are unlikely to be representative of popular misinformation consumed. This introduces limitations for properly quantifying the effect of interventions on misinformation encountered online.

For each of the study’s 10 weeks, we will collect and distribute a new group of five articles for each respondent to evaluate in randomized order. Three of these articles come from political websites known to produce low credibility news; two articles come from mainstream news sources. Section J, Supplementary Materials fully describes the news sources we select articles from and this selection process.

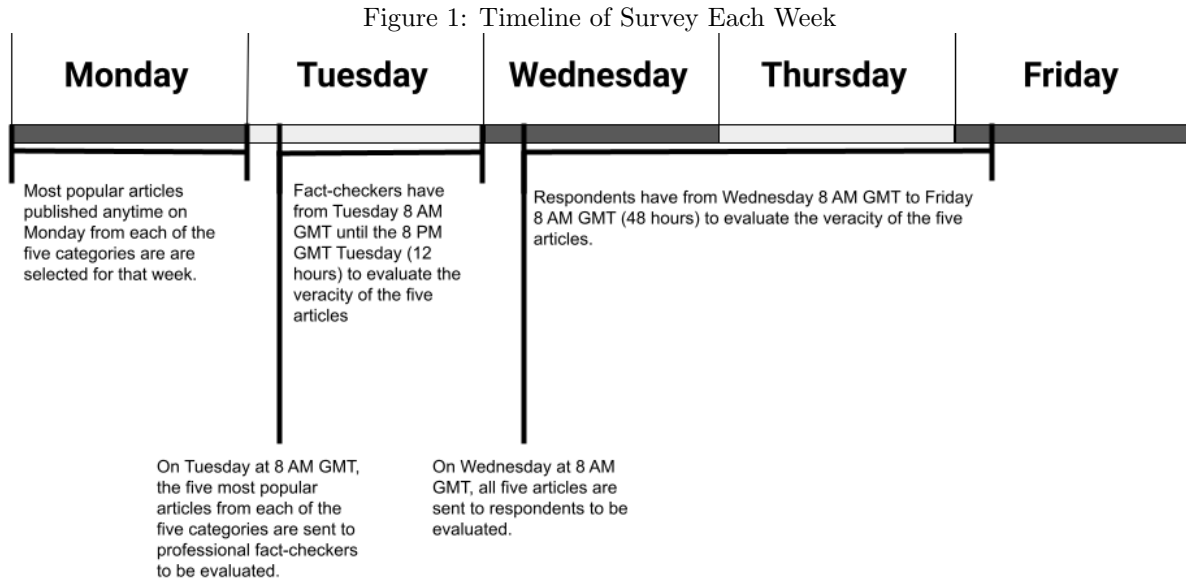
Figure 1 displays the weekly process of selecting articles and distributing these articles to both respondents and professional fact-checkers. Every week, on Tuesday morning we will source the most popular article published on Monday from each of the five lists of sources.⁹ The most popular article published in each of the five lists of online news websites are selected and then distributed to professional fact-checkers to be evaluated by a team of three professional fact-checkers to ascertain whether each article is true, false/misleading, or “could not determine” based on the materials in the article. We use the modal response of the three professional fact checkers to determine whether an article is coded as true, false/misleading, or ‘could not determine’.¹⁰ Then, on Wednesday morning we will send out the articles to be evaluated by respondents, who will have until Friday morning to complete their evaluations. For each article evaluation, we randomly assign the language in which the respondent reads the article (either Ukrainian or Russian).¹¹ Using this process, respondents evaluate articles within 48-96 hours of publication.¹²

⁹We exclude some select articles. The exclusion protocol can be found in Section A, Supplementary Materials.

¹⁰Professional fact-checkers rate articles first, so that we can inform our respondents of the veracity of each article they evaluated at the end of the survey.

¹¹Articles are translated to Ukrainian or Russian if the article is only available in one language on the website. They are translated by one translator and then checked by another translator.

¹²An example of how the article is viewed by respondents can be found in Section F, Supplementary Materials.



Given that we focus on belief in misinformation, we only use evaluations of articles that are rated as “false/misleading” by professional fact-checkers. Evaluations of all other articles are not utilized in the main analysis.¹³ We leverage the random assignment of the language in which the story is read by respondents, to assign article evaluations to a control or treatment group. Every evaluation of an article read in a respondent’s non-dominant language is a ‘treated’ observation, while each evaluation of an article in a respondent’s dominant language is ‘not treated.’ In table 1 we display the evaluations that we consider in the treatment and control group among Ukrainian-dominant and Russian-dominant speakers.

Table 1: Assignment of Treatment by Preferred Language of Respondent and Language Article is Written In

Preferred Language of Respondent		
Language of Article Evaluated	Ukrainian	Russian
Ukrainian	Control	Treatment
Russian	Treatment	Control

¹³In the exploratory analyses section, we test if the estimated effects hold across articles of different ideological perspectives and true articles.

Given respondent i and news article a , we test each hypothesis using a 4-point ordinal scale ($Y_ordinal_4_{ia}$).¹⁴ Additionally, before respondents assess articles, we randomly present half of the respondents with “tips” to help spot false news stories (Section C, Supplementary Materials contains the full set of tips). Analyzing the “tips” treatment is not a part of this study, but we control for it in robustness checks to ensure that it does not affect our results. We also ask respondents to pass an attention check at five points throughout the survey. After respondents reads each of the five articles they are shown in the survey, we ask each them to tell us if they were able to access the article. If they respond “No”, but also answer the evaluation question, it is clear that they are untruthfully (and inattentively) answering these questions. Removing inattentive respondents threatens internal and external validity (Berinsky, Margolis, and Sances 2014), so we control for their attention in our primary regression models.

4 Proposed Methods

We test our hypotheses using OLS regression models with article-level fixed effects (α_a in the models) and two-way standard errors¹⁵ clustered at the respondent and article level to predict belief in misinformation (i.e., rating a false or misleading article as true). We only use evaluations of articles rated as false/misleading by the modal number of fact-checkers. For each respondent i and news article a , we test each hypothesis using one main dependent variable that measures perceived veracity on a four-point ordinal scale ($Y_ordinal_4_{ia}$). For the results relating to our three primary hypotheses (H1-H3), we will present the false discovery rate (a false discovery rate at 0.10) corrected p -values (Benjamini and Hochberg 1995).

In the following models, the treatment variable ($Lang_Treatment_{ia}$) is a dummy variable (1 if the article evaluated is written in the respondent’s less preferred language, 0 otherwise).¹⁶ The γ coefficient represents the estimate of interest. We specify \mathbf{Z} as a collection of covariates, which will be minimal (the attention check) in most models since observable variables should be balanced. We will use a larger set of pre-treatment covariates identified in previous research to correlated with individuals’ perceived veracity of information in robustness check II. These include education, age, income, region fixed effects, gender, whether they were

¹⁴Question: “To the best of your knowledge, how accurate is the claim in the above article?
Answers: (A) Very accurate (4); (B) Somewhat accurate (3); (C) Not very accurate (2); (D) Not at all accurate (1).” Many studies on belief in misinformation, use this measure (e.g., Guess and Munger 2020; Pennycook, Cannon, and Rand 2018; Pennycook and Rand 2019), but as a robustness check we also test each hypothesis using a categorical measure (Y_true_{ia}) detailed in Section B, Supplementary Materials.

¹⁵We use HC2 robust standard errors in all analyses and report p -values from two-tailed t -tests.

¹⁶Proficiency is closely-related but distinct from language usage (Andreenkova 2019). We measure language preference by grouping each respondent by the language in which they choose to take the survey (itself an exercise in reading, which likely mimics print media consumption). However, in robustness check I, we use other self-identified measures of Ukrainian and Russian proficiency (see Section C, Supplementary Materials).

treated with the digital media literacy intervention, whether the article was translated by our hired translator and not taken directly from the website, and ideological congruence with the evaluated article’s ideological lean.

$$Y_Outcome_{ia} = \alpha_a + \gamma * Lang_Treatment_{ia} + \Gamma Z + \epsilon \quad (1)$$

In robustness check (III), we will run the analysis only including respondents who were not presented with the tips to identify misinformation before they evaluated the articles, which is part of a different study. This treatment is randomly assigned, so it is unlikely to affect results, but we will confirm this is the case. As a final robustness check (IV), we will also test each hypotheses using a categorical measure (Y_true_{ia}),¹⁷ rather than the 4-point ordinal scale ($Y_ordinal_4_{ia}$) specified earlier.¹⁸

5 Expected Results

To preview what our results may look like, we simulated data for a ten-week study using a pilot study of 230 respondents carried out the week of June 7th, 2021.¹⁹ We randomly sampled from these 230 responses with replacement to create 1,000 simulated datasets of 2,300 hypothetical respondents who evaluated two articles rated as true, one as “could not determine”, and two as false/misleading. As specified by our design, we only use evaluations of the false/misleading articles for the main hypotheses. These articles were the most popular articles from the “Not Anti-Russia low-quality” and “Russia-based mainstream” news streams. The false/misleading article from the “Not Anti-Russia low-quality” news stream described a new tax proposed by Ukrainian President Zelensky that would hurt the Ukrainian economy. The false/misleading article from the “Russia-based mainstream” news stream stated that the Ukrainian soccer team will use a nationalist symbol during the European Soccer Championship that assumes Crimea is a part of Ukraine even though it officially is not.²⁰

These two stories yield 4,600 simulated evaluations of false/misleading articles, similar to the number we expect to collect during the main study, though we plan to have many more diverse articles during the main

¹⁷Even for our dichotomous variable (Y_true_{ia}) an OLS regression is preferred, because it provides unbiased, reliable estimates of a variable’s average effect (Allison 1999; Hellevik 2009; Baetschmann, Staub, and Winkelmann 2015)

¹⁸Section H and G of the Supplementary Materials contains questions used to measure variables used in all models.

¹⁹We will not use any of this data in the final analysis, but a balance table for this pilot study is located in Section G, Supplementary Materials. In the final version of the paper, this section will be replaced with one that reports the results from the actual study, as opposed to the simulated study we describe here.

²⁰All the articles used in the pilot study are located in Section I, Supplementary Materials

study. Using 1,000 simulated datasets of 4,600 evaluations we run our pre-registered models and present our expected results for models testing Hypotheses 1 through 3 and exploratory analyses. The simulated results we present in this section are suggestive, but we present findings in a fashion we plan to once the full ten-day study is complete.²¹

Figure 2 displays the estimated effect of reading misinformation written in one’s less preferred language on belief of misinformation in the three distinct contexts tested in Hypotheses 1 through 3. Each line represents a confidence interval that covers 95 percent of the point estimates reported using these simulated datasets. Each dot represents the median point estimate. Our simulated results suggest support for Hypothesis 1: Russian-dominant speakers on average (median) belief in misinformation was reduced by -0.23 (95% confidence interval: [-0.15,-.31]) on a 4-point ordinal score when reading it in their less-preferred (Ukrainian) language. Among Ukrainian-dominant speakers, we find the opposite effect. On average (median), belief in misinformation increased by 0.13 (95% confidence interval: [0.05,0.20]) on a 4-point ordinal score when Ukrainian-dominant speakers read misinformation in their less-preferred language (Russian). This evidence from our simulated datasets does not appear to support Hypothesis 2.

Although our simulated data does not support our second hypothesis, we further investigate (Hypothesis 3) whether Ukrainian-dominant speakers are less likely to believe misinformation written in their less-preferred language (Russian) in areas known to be targeted by Russian disinformation (South and East regions of Ukraine). Ukrainian-dominant speakers in the South and East regions of Ukraine, on average (median) believe misinformation when reading it in their less-preferred language (Ukrainian) less by 0.39 (95% confidence interval: [-0.22,-0.57]) on a 4-point ordinal score.²²

5.1 Exploratory Analyses

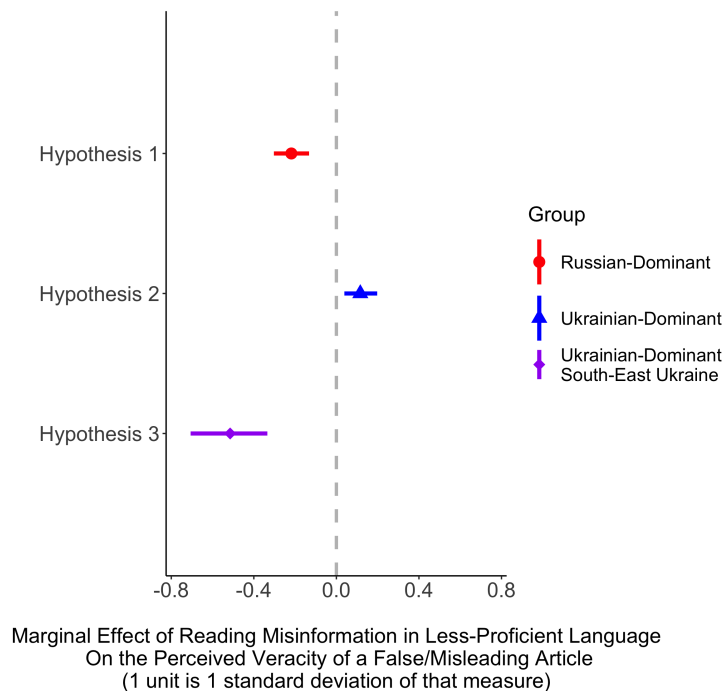
We do not have sufficient data to present the results from Exploratory Analysis 1, but we will present the results after completing the full ten-day survey. In our second exploratory analysis, we present the marginal effect of reading *true* news in their less-preferred language among Ukrainian-dominant and Russian-dominant speakers. Our simulated results indicate that the effects we identified among false/misleading articles are **not** present for true articles.²³

²¹The code we wrote in the R statistical program to carry out these analyses on our simulated datasets can be found in Section K of the Supplementary Materials.

²²Robustness checks in Section H of the Supplementary Materials also support the findings for each of the hypotheses presented.

²³Figures presenting the coefficients and confidence intervals are available in Section H of the Supplementary Materials.

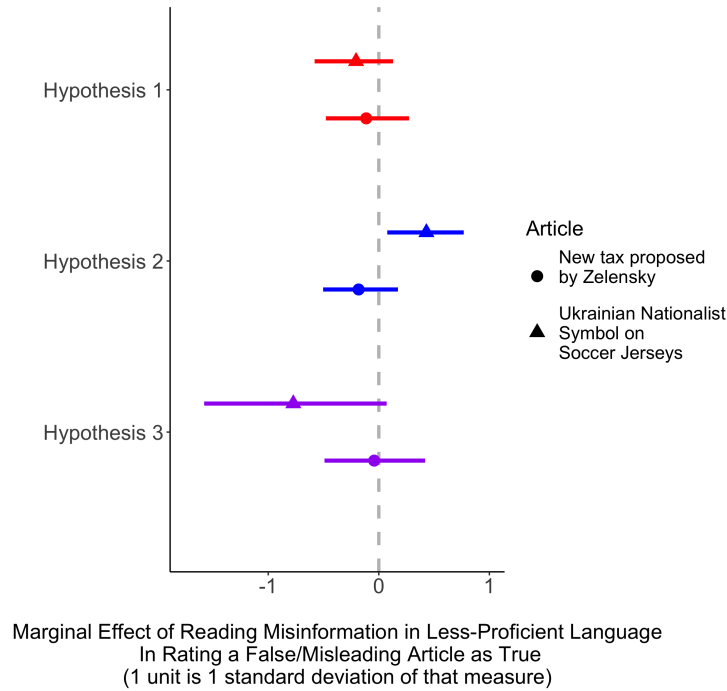
Figure 2: Predicted Pooled Marginal Effect of Evaluating *False/Misleading* News Articles in One’s Less Preferred Language



For our third exploratory analysis we will strictly report the p -value of the interaction term comparing the language effects of reading misinformation in one’s less preferred language on Ukrainian-Dominant and Russian-Dominant speakers. We found that in almost every simulated analysis, the language effect of reading misinformation in one’s less preferred language is stronger for Russian-dominant speakers than Ukrainian speakers and statistically significant at the 95 percent level.

In our fourth and final exploratory analysis, we report the effects of reading misinformation in one’s less preferred language for Hypotheses 1 through 3 for each false/misleading article. Figure 3 displays the effects of reading misinformation in one’s less preferred language on simulated data and two false/misleading articles from the pilot testing each hypothesis using simulated data for each article. This is a disaggregated analysis where we are not concerned about the statistical significance, given the small sample of evaluations per article. Rather we are interested if the marginal effects are in a similar direction across each article. We find that the median effect identified for hypothesis 1 is relatively consistent for both articles, but the predicted effects are mixed for hypothesis 2 and 3 across articles.

Figure 3: Predicted Marginal Effect of Evaluating Misinformation in One’s Less Preferred Language for each False/Misleading article



Overall, simulated data using our pilot study suggests that linguistic minorities are less likely to believe misinformation written in the majority language and that there is potentially an effect of reading news in languages associated with foreign disinformation campaign, but only in areas where these campaigns are most visible. Our explanatory analyses also reveal that this language effect may solely apply to misinformation and not true news.

6 Discussion

This section will be added in the final paper to discuss the implications of our findings and directions for future research.

7 References

Allcott, Hunt and Matthew Gentzkow (2017). “Social media and fake news in the 2016 election”. *Journal of economic perspectives* 31.2, pp. 211–36.

- Allison, Paul D (1999). “Comparing logit and probit coefficients across groups”. *Sociological methods & research* 28.2, pp. 186–208.
- Althaus, Scott L and David Tewksbury (2000). “Patterns of Internet and traditional news media use in a networked community”. *Political communication* 17.1, pp. 21–45.
- Andreenkova, Anna V. (2019). “How to Choose Interview Language in Different Countries”. In: *Advances in Comparative Survey Methods*. Ed. by Timothy P. Johnson et al. 1st ed. John Wiley & Sons, Ltd, pp. 314–324.
- Ansaldo, Ana Inés et al. (2008). “Language therapy and bilingual aphasia: Clinical implications of psycholinguistic and neuroimaging research”. *Journal of Neurolinguistics* 21.6, pp. 539–557.
- Badrinathan, Sumitra (2020). “Educative Interventions to Combat Misinformation: Evidence from a Field Experiment in India”. *American Political Science Review*, pp. 1–17.
- Baetschmann, Gregori, Kevin E Staub, and Rainer Winkelmann (2015). “Consistent estimation of the fixed effects ordered logit model”. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pp. 685–703.
- Benjamini, Yoav and Yosef Hochberg (1995). “Controlling the false discovery rate: a practical and powerful approach to multiple testing”. *Journal of the Royal statistical society: series B (Methodological)* 57.1, pp. 289–300.
- Berinsky, Adam J, Michele F Margolis, and Michael W Sances (2014). “Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys”. *American Journal of Political Science* 58.3, pp. 739–753.
- Boroditsky, Lera (2001). “Does language shape thought?: Mandarin and English speakers’ conceptions of time”. *Cognitive psychology* 43.1, pp. 1–22.
- (2006). “Linguistic relativity”. *Encyclopedia of cognitive science*.
- Bronstein, Michael V et al. (2019). “Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking”. *Journal of Applied Research in Memory and Cognition* 8.1, pp. 108–117.
- Clayton, Katherine et al. (2019). “Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media”. *Political Behavior*, pp. 1–23.
- Corey, Joanna D et al. (2017). “Our moral choices are foreign to us.” *Journal of experimental psychology: Learning, Memory, and Cognition* 43.7, p. 1109.
- Costa, Albert, Marc-Lluis Vives, and Joanna D Corey (2017). “On language processing shaping decision making”. *Current Directions in Psychological Science* 26.2, pp. 146–151.

- Dragojlovic, Nick (2015). “Listening to outsiders: The impact of messenger nationality on transnational persuasion in the United States”. *International Studies Quarterly* 59.1, pp. 73–85.
- Erlich, Aaron and Calvin Garner (2021). “Is pro-Kremlin Disinformation Effective? Evidence from Ukraine”. *The International Journal of Press/Politics*.
- Fausey, Caitlin M et al. (2010). “Constructing agency: the role of language”. *Frontiers in psychology* 1, p. 162.
- Fernández-López, Maria and Manuel Perea (2020). “Language does not modulate fake news credibility, but emotion does”. *Psicológica Journal* 41.2, pp. 84–102.
- Fisher, Aleksandr (2020). “Demonizing the enemy: the influence of Russian state-sponsored media on American audiences”. *Post-Soviet Affairs* 36.4, pp. 281–296.
- Flanagin, Andrew J and Miriam J Metzger (2000). “Perceptions of Internet information credibility”. *Journalism & Mass Communication Quarterly* 77.3, pp. 515–540.
- (2007). “The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information”. *New media & society* 9.2, pp. 319–342.
- Flores, Alejandro and Alexander Coppock (2018). “Do bilinguals respond more favorably to candidate advertisements in English or in Spanish?” *Political Communication* 35.4, pp. 612–633.
- Fogg, Brian J et al. (2001). “What makes web sites credible? A report on a large quantitative study”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 61–68.
- Grosjean, François (2010). *Bilingual*. Harvard university press.
- Guess, Andrew and Kevin Munger (2020). “Digital Literacy and Online Political Behavior”. *Charlottesville: OSF Preprints*. Retrieved April 13, p. 2020.
- Hellevik, Ottar (2009). “Linear versus logistic regression when the dependent variable is a dichotomy”. *Quality & Quantity* 43.1, pp. 59–74.
- Ismail, Jamal Abdi and James Deane (2008). “The 2007 general election in Kenya and its aftermath: The role of local language media”. *The International Journal of Press/Politics* 13.3, pp. 319–327.
- Keshishian, Flora (2000). “Acculturation, communication, and the US mass media: The experience of an Iranian immigrant”. *Howard Journal of Communications* 11.2, pp. 93–106.
- Keysar, Boaz, Sayuri L Hayakawa, and Sun Gyu An (2012). “The foreign-language effect: Thinking in a foreign tongue reduces decision biases”. *Psychological science* 23.6, pp. 661–668.
- Kuklinski, James H et al. (2000). “Misinformation and the currency of democratic citizenship”. *The Journal of Politics* 62.3, pp. 790–816.
- Lau, Richard R et al. (2017). “Effect of media environment diversity and advertising tone on information search, selective exposure, and affective polarization”. *Political Behavior* 39.1, pp. 231–255.

- Mastro, Dana (2009). “Racial/ethnic stereotyping and the media”. *Media processes and effects*, pp. 377–391.
- Moravec, Patricia, Randall Minas, and Alan R Dennis (2018). “Fake news on social media: People believe what they want to believe when it makes no sense at all”. *Kelley School of Business Research Paper* 18-87.
- Muda, Rafal et al. (2021). *People Are Worse at Detecting Fake News in Their Foreign Language*. URL: [doi:10.31219/osf.io/p8su6](https://doi.org/10.31219/osf.io/p8su6).
- Mujani, Saiful and Nicholas Kuipers (2020). “Who Believed Misinformation during the 2019 Indonesian Election?” *Asian Survey* 60.6, pp. 1029–1043.
- Onguny, Philip (2019). “Electoral violence in Kenya 2007-2008 - the role of vernacular radio”. *Journal of African Elections* 18.1, pp. 86–107.
- Onuch, Olga and Henry E. Hale (May 2018). “Capturing ethnicity: the case of Ukraine”. en. *Post-Soviet Affairs* 34.2-3, pp. 84–106.
- Pennycook, Gordon, Tyrone D Cannon, and David G Rand (2018). “Prior exposure increases perceived accuracy of fake news.” *Journal of experimental psychology: general* 147.12, p. 1865.
- Pennycook, Gordon and David G Rand (2019). “Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning”. *Cognition* 188, pp. 39–50.
- (2020). “Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking”. *Journal of personality* 88.2, pp. 185–200.
- Pérez, Efrén O (2016). “Rolling off the tongue into the top-of-the-head: Explaining language effects on public opinion”. *Political Behavior* 38.3, pp. 603–634.
- Pérez, Efrén O and Margit Tavits (2019). “Language heightens the political salience of ethnic divisions”. *Journal of Experimental Political Science* 6.2, pp. 131–140.
- PwC-UK (2016). “The rise of cross-border news”.
- Rosenzweig, Leah R et al. (2021). “Happiness and surprise are associated with worse truth discernment of COVID-19 headlines among social media users in Nigeria”. *Harvard Kennedy School Misinformation Review*.
- Ross, Michael, W. Q. Elaine Xun, and Anne E. Wilson (Aug. 2002). “Language and the Bicultural Self”. *Personality and Social Psychology Bulletin* 28.8, pp. 1040–1050.
- Satariano, Adam (2019). “Facebook identifies Russia-linked misinformation campaign”. *The New York Times*. Retrieved March 1, p. 2019.
- Snegovaya, Maria (2015). “Putin’s information warfare in Ukraine”. *Soviet Origins of Russia’s Hybrid Warfare*, *Russia Report* 1, pp. 133–135.

- Somerville, Keith (2009). “British media coverage of the post-election violence in Kenya, 2007–08”. *Journal of Eastern African Studies* 3.3, pp. 526–542.
- SSCU (2001). *All-Ukrainian Population Census 2001*. State Statistics Committee of Ukraine. Accessed August 03, 2021. URL: <http://2001.ukrcensus.gov.ua/eng/results/general/language/>.
- Starbird, Kate et al. (2018). “Engage early, correct more: How journalists participate in false rumors online during crisis events”. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–12.
- Stewart, Andrew, Joshua Plotkin, and Nolan McCarty (2021). “Inequality, identity, and partisanship: How redistribution can stem the tide of mass polarization”. *Proceedings of the National Academy of Sciences* 118.
- StratCom (2015). *Analysis of Russias Information Campaign against Ukraine*. Tech. rep. NATO StratCom Centre of Excellence.
- Suhay, Elizabeth, Emily Bello-Pardo, and Brianna Maurer (2018). “The polarizing effects of online partisan criticism: Evidence from two experiments”. *The International Journal of Press/Politics* 23.1, pp. 95–115.
- Sundar, S Shyam and Clifford Nass (2001). “Conceptualizing sources in online news”. *Journal of communication* 51.1, pp. 52–72.
- Trafimow, David et al. (Jan. 1997). “The Effects of Language and Priming on the Relative Accessibility of the Private Self and the Collective Self”. *Journal of Cross-Cultural Psychology* 28.1.
- Tsfati, Yariv and Yoram Peri (2006). “Mainstream media skepticism and exposure to sectorial and extranational news media: The case of Israel”. *Mass Communication & Society* 9.2, pp. 165–187.
- Tukachinsky, Riva, Dana Mastro, and Moran Yarchi (2015). “Documenting portrayals of race/ethnicity on primetime television over a 20-year span and their association with national-level racial/ethnic attitudes”.
- Vihalemm, Triin, Jānis Juzefovičs, and Marianne Leppik (2019). “Identity and media-use strategies of the Estonian and Latvian Russian-speaking populations amid political crisis”. *Europe-Asia Studies* 71.1, pp. 48–70.
- Vorobiov, Ievgen (2015). “Why Ukrainians Are Speaking More Ukrainian”. *Foreign Affairs*. URL: <https://foreignpolicy.com/2015/06/26/why-ukrainians-are-speaking-more-ukrainian>.
- Vosoughi, Soroush, Deb Roy, and Sinan Aral (2018). “The spread of true and false news online”. *Science* 359.6380, pp. 1146–1151. URL: <https://science.sciencemag.org/content/359/6380/1146>.