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]

[a]Center for Social Media and Politics, New York University
[b]Department of Political Science, McGill University
[c]Wilf Family Department of Politics, New York University

Introduction

Our cumulative knowledge about belief in misinformation predominantly comes from surveying 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 increases levels of affective polarization (Lau et al. 2017; Suhay, Bello-Pardo, and Maurer 2018), 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. Specifically, we address the question of whether groups are more or less susceptible to misinformation in their non-dominant language?

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

Much of the work on language and misinformation has focused on identifying a proficiency effect of language by investigating the 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 more 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 have lower belief in 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 ones less preferred language may not have the same effect across each language. 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; Aslett et al. n.d.) may explain the differential effects across languages.

Therefore, we move past this proficiency argument and focus on the effect of what language signals in terms of information credibility. While proficiency is important, oftentimes citizens are have similar levels of proficiency in two or more languages they speak.[2] In these cases, credibility cues are more likely than proficiency to be the source of any language effect on belief in misinformation. Given this, we focus on the effect of evaluating misinformation in one's less preferred language relative to their dominant language. Credibility cues are likely stronger 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.

The source of information also affects individuals' belief in misinformation. News consumers 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 also change the effect of information on public opinion (Dragojlovic 2013; Dragojlovic 2015). 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 2018). The effects of language on consuming information clearly depend on

[2]Examples include many inhabitants of the former Soviet Union (including Ukrainians), Spanish speakers in the U.S., and large numbers of urban Sub-Saharan Africans who often speak both an indigenous and colonial settler language (e.g., French-English).

[3]Using 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).

the contextual cues given by domestic or international conditions between groups that speak each language. We test if domestic or international conditions create credibility cues that affect belief in misinformation in one's less preferred language.

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 most prevalent language. Therefore, in such situations, we expect that those that prefer to speak a minority language should be less likely to believe misinformation if it is written in the majority language spoken (ie., their less-preferred language).

We test this "language of the majority" hypothesis among those that prefer to speak Russian in Ukraine and test what 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 has some proficiency in both languages, given their linguistic similarity.[4]

International conditions could also provide credibility cues. A sustained disinformation campaign by a foreign power that employs a domestic minority language could lower belief in misinformation among those who prefer to consume media in the majority language. Often disinformation campaigns target a minority ethnic group that uses their own preferred language (Snegovaya 2015) that differs from the majority-spoken language. 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 reading news in the "language of external power" to decrease the likelihood that one would believe misinformation in that language among

[4]In 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.

those who prefer to speak the language of the majority.

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 the perception of the reported information's veracity in that language. Previous research has partially identified the latter half of 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 (Erlich and Garner 2021). However, 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. We, therefore, test this "language of external power" hypothesis among those that prefer to speak Ukrainian in Ukraine and test what effect reading misinformation in their less-preferred language, Russian, has on belief in that misinformation. In addition, we test if this effect is strongest in areas where Russian disinformation is most prevalent.

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 campaigns heavily target the Russian-speaking diaspora (Snegovaya 2015) through Russian language news websites and media companies.[5] This news in Russian is often associated with Russia's disinformation campaign.

To test both the "language of the majority" and "language of external power" hypotheses we conduct a survey in Ukraine in which we randomly assign 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. With the survey experiment 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 the foreign power associated with a foreign disinformation campaign? (3) Does this "external power" language exist in areas that disinformation campaigns specifically target?

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

Hypotheses

To answer our three research questions we measuring the effect of reading misinformation in one's less preferred language in three different situations. First, we measure the "majority language" effect by measuring the effect of reading news in the language spoken by the majority when one's preferred language is a language spoken by the minority. In Ukraine, we test this by measuring the effect of reading misinformation in Ukrainian among Russian-dominant speakers. Second, we measure the language effect when one's less preferred language passes on a "external power" cue and is associated with a foreign disinformation campaign. In Ukraine, we test this by measuring the effect of reading misinformation in Russian among Ukrainian-dominant speakers across all of Ukraine. To answer our third and final research question we test this theory solely among Ukrainian-dominant speakers that live in areas targeted by Russian disinformation (those in Southern and Eastern Ukraine).[6].

Given that we focus on the belief in misinformation, our main hypotheses only use evaluations of articles that are rated as "false/misleading" by professional fact-checkers, and evaluations of all other articles are not utilized in the main analysis.[7] We leverage the random assignment of the language in which the story is read by respondents, to assign 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

Below we list our main hypotheses. For the rest of this paper we define those that are "Ukrainian-dominant" as those who prefer to read in Ukrainian and "Russian-dominant" as those who prefer to read in

[6]Given 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 J of the Supplementary Materials

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

Russian.

First, we test if evaluating misinformation in one's less preferred language increases belief in misinformation when ones less preferred language is the majority language of the country. Therefore, we conduct test on respondents who are Russian-dominant (the minority linguistic group in Ukraine). We expect Russian-dominant speakers in Ukraine to be skeptical of news written in Ukrainian. Consequently, reading misinformation in their less preferred language (Ukrainian) will make them less likely to believe false/misleading news.

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

Second, we test if the language effect we describe above also holds when ones less preferred language passes on a "foreign aggressor" cue. When ones less preferred language is associated with a foreign disinformation campaign it is likely that one is less likely to believe a false/misleading news story in that language. In the case of Ukraine, we can test this hypotheses strictly among respondents who are Ukrainian-dominant. We expect the Ukrainian-dominant speakers in Ukraine are skeptical of news written in Russian (the language of the foreign disinformation campaigns). Consequently, they are less likely to be believe misinformation written in their less preferred language (Russian).

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

In addition, we test if the language effect tested in Hypothesis 2 holds in regions in which Russian disinformation is most prevalent (Southern and Eastern Ukraine).

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

Exploratory Analyses

Exploratory Analysis 1: To determine if the effects we find hold across articles with different ideological perspectives, we will test each of the hypotheses listed in the previous section strictly using sets of evaluations pooled by articles with different ideological perspectives present in Ukraine.[8] These may be under-powered if there are only some false/misleading articles with certain ideological perspectives. If they are under-powered we may not choose to present all of these analyses in the main body of the paper, but we will report them in the supplementary materials.

Exploratory Analysis 2: We will test each hypotheses using articles that are rated as true by professional fact-checkers. The dependent variable in these models will be replaced by belief in true news.

Exploratory Analysis 3: We will test if there is a measurable difference in the "majority language" effect (H1) and the "foreign power" effect (H2) on those evaluating a false/misleading article in one's less preferred language.

Exploratory Analysis 4: To determine if the effects we find hold across all articles, we will test each of the hypotheses listed in the previous section strictly using sets of evaluations pooled by each unique false/misleading article.

Experimental Design

To answer our research questions we will carry out a 10-week survey experiment in Ukraine. 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 Section I of the Supplementary Materials. Human subjects in Ukrainian samples are recruited by *Qualtrics*.[9]

Research measuring beliefs in misinformation has yet to integrate important findings about how individuals consume misinformation, limiting inference from these studies. Specifically, fake news 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, fake news articles (Bronstein et al. 2019; Clayton et al. 2019; Pennycook and Rand 2020). To address this limitation, we create a transparent, replicable, and

[8] We will classify articles into the following ideological perspectives: Anti-Russia, Not Anti-Russia, and Unclear. We specifically explain why these ideological perspectives are chosen in the Experimental Design section.

[9] Section I, Supplementary Materials details the recruitment of respondents.

pre-registered news article selection 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, which introduce limitations for properly quantifying the effect of interventions on misinformation encountered online. We are unsure if the sampling frames for misinformation are representative of popular misinformation consumed.

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 R, Supplementary Materials fully describes the news sources we select articles from and this selection process.

Given respondent i and news article a, we test each hypothesis using our main dependent variable ($Y\_ordinal\_4_{ia}$).[10] Many studies on belief in misinformation, use this measure (e.g., Guess and Munger 2020; Pennycook, Cannon, and Rand 2018; Pennycook and Rand 2019). As a robustness check we also test each hypotheses using a categorical measure (Y_true_{ia}) detailed in Section F, Supplementary Materials.

Additionally, before respondents assess articles, we randomly present half of the respondents with "tips" to help spot false news stories (Section G, 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.

[10] Question: "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)."

Methods and Results (Plan for Hypotheses)

We test our hypotheses using OLS regression models with article-level fixed effects (α_a in the models) and two-way standard errors[11] clustered at the respondent and article level to predict belief in misinformation (i.e., rating a false or misleading article as true). 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 Benjamini and Hochberg 1995 p-values.

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).[12] 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 treated with the digital media literacy intervention, and ideological congruence with the evaluated article's ideological lean.

$$Y\_Outcome_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon \tag{1}$$

In robustness check (III), we will run the analysis only including respondents who were not presented with the fake news tips 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. In additional models, we will control for whether the article that was evaluated was translated by one of our translators. We will use a dummy variable ($Translated_a$) assigned 1 if the article was translated by our hired translator and not taken directly from the website. As a final robustness check (IV), we will also test each hypotheses using a categorical measure (Y_true_{ia}),[13] rather than the 4-point ordinal scale ($Y\_ordinal\_4_{ia}$) specified earlier.[14]

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

[12] Proficiency 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 G, Supplementary Materials).

[13] Even 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)

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

Expected Results

To preview what our results will 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.[15] 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 (these are the pilot articles detailed above). 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 apart of Ukraine even though it officially is not.

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 would have many more diverse articles during the main study. Using these datasets we run our pre-registered models to display how our results for models testing Hypotheses 1 through 3 and exploratory analyses may look after the full 10-day experiment. We present findings in a fashion we we would expect to present the main study's results.

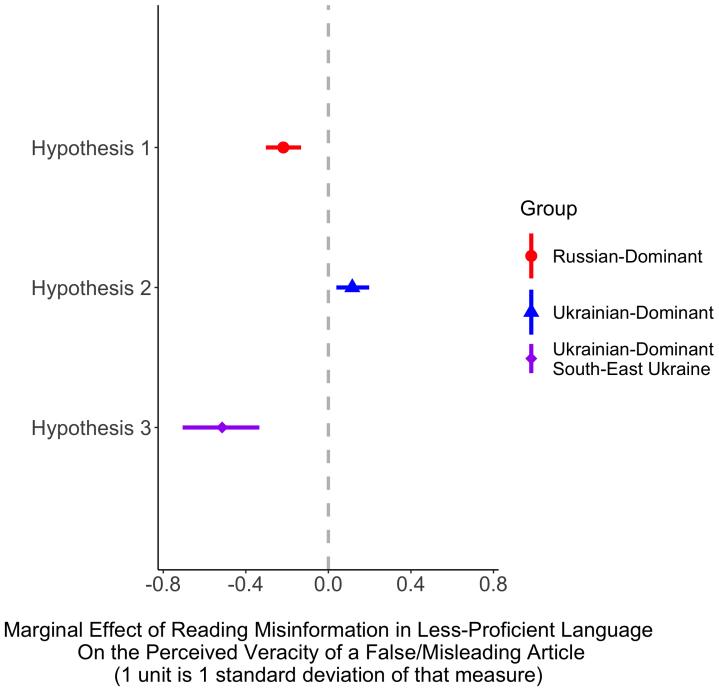
Figure 1 displays the median marginal effects on belief of misinformation written in one's less preferred language and confidence intervals that cover 95 percent of the point estimates reported using these simulated datasets for our models testing Hypotheses 1 through 3. Our simulated results suggest 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 dataset 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

[15] We will not use this data in the final analysis.

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.[16]

Figure 1: Predicted Pooled Marginal Effect of Evaluating *False/Misleading* News Articles in One's Less Preferred Language



Exploratory Analyses

We do not have sufficient data to present results from Exploratory Analysis 1 in the main body, but after collecting data from the full ten-day survey we will display the results.

In our second exploratory analysis, we present the marginal effect of reading *true* news 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. Ukrainian-dominant speakers were slightly less likely to believe true news when using the main 4-point ordinal scale; however, this effect does not hold across all three measures of belief in misinformation. We also do not find that Russian-dominant speakers were more or less likely to believe true news in their less preferred language (Ukrainian).[17]

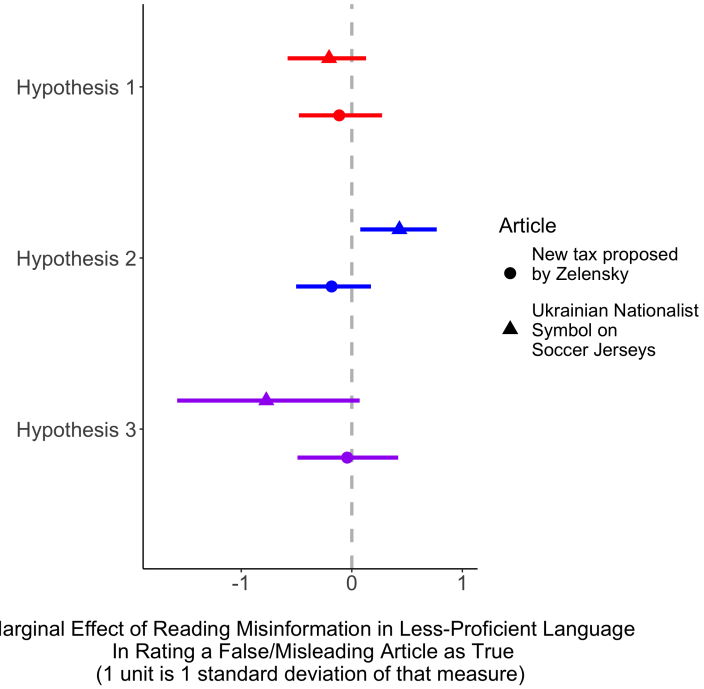
[16]Robustness checks in Section M of the Supplementary Materials also support the findings for each of the hypotheses presented.

[17]Figures presenting the coefficients and confidence intervals are available in Section M of the Supplementary Materials.

For our third exploratory analysis we 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 our simulated main analysis, the language effect of reading misinformation in one's less preferred language is stronger for Russian-dominant speakers than Ukrainian speakers by an average -0.35 (0.38 of a standard deviation of that measure for belief in misinformation) on a four-point ordinal scale.

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. In Figure 2 we display the effects of 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 2: Predicted Marginal Effect of Evaluating Misinformation in One's Less Preferred Language for each False/Misleading article



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 prevalent. Our explanatory analyses also reveal that this language effect may solely apply to misinformation and not true news. The simulated results we present in this section are suggestive and serve as an outline for how we will present our results once we run the full ten-week study.

Discussion

TBD post-experiment.

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.

Aslett, Kevin et al. (n.d.). "Testing The Effect of Additional Information on Discerning the Veracity of News in Real-Time" ().

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 (2013). "Leaders without borders: Familiarity as a moderator of transnational source cue effects". *Political Communication* 30.2, pp. 297–316.

— (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, Marıa and Manuel Perea (2020). "Language does not modulate fake news credibility, but emotion does". *Psicológica Journal* 41.2, pp. 84–102.

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.

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.

Pennycook, Gordon and David G Rand (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.

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`.

Supplementary Methods and Materials: Effect of Browser Extension on Behavior and Trust

Kevin Aslett[a], Aaron Erlich[b], Joshua A. Tucker[a,c], Jonathan Nagler[a,c], and Richard Bonneau[a,d]

[a]Center for Social Media and Politics, New York University
[b]Department of Political Science, McGill University
[c]Wilf Family Department of Politics, New York University
[d]Department of Biology, New York University

Contents

A Five News Streams

Table 1: Ukrainian-High Credibility News Domains

Domain
24tv.ua
bykvu.com
censor.net.ua
fakty.com.ua (No RSS Feed)
fakty.ua
gordonua.com
interfax.com.ua
lb.ua
liga.net
nv.ua
pravda.com.ua
rbc.ua
segodnya.ua
tsn.ua
ukrinform.ru (No RSS Feed)
unian.net
zn.ua

Table 2: Russian-High Credibility News Domains

Domain
aif.ru
dni.ru
kommersant.ru
kp.ru
lenta.ru
lentainform.com (No RSS Feed)
life.ru
newsru.com
pravda.ru
regnum.ru
riafan.ru
russian.rt.com
slovodel.com
svpressa.ru (No RSS Feed)
tass.ru
vz.ru

Table 3: Not Anti-Russia-Low Credibility News Domains:

Domain
fraza.ua
mignews.com.ua
sharij.net
vesti.ua
vz.ua
antimaydan.info (No RSS Feed)
c-inform.info
dnr-pravda.ru
dnr24.com
dnr24.su
donbasstoday.ru
doneck-news.com
dontimes.ru
e-gorlovka.com.ua (No RSS Feed)
evening-crimea.com
free-news.su (No RSS Feed
fromdonetsk.net
luga1news.ru
lugansk1.info
miaistok.su
newc.info
news-front.info
novorosinform.org
novorossiy.info
novosti.icu
patriot-donetsk.ru
politnavigator.net
pravdanews.info
ruinformer.com
rusdnepr.ru
rusnext.ru
sevastopol.su
sevnews.info
time-news.net
voenkor.info (No RSS Feed)
voskhodinfo.su
vsednr.ru
xvesti.ru

Table 4: Anti-Russia-Low Credibility News Domains

Domain
agrimpasa.com
antikor.com.ua
finoboz.net
inforesist.org
informator.news (No RSS Feed)
kyiv.press
spektrnews.in.ua
ukranews.com

Table 5: Unclear Low Credibility News Domains

Domain
glavred.info
jizn.info
kompromat1.news
korr.com.ua
lifedon.com.ua
news247.com.ua (No RSS Feed)
podrobnosti.ua (No RSS Feed)
kafanews.com
meridian.in.ua

B Article Exclusion Protocol:

Our target content is online written news. Consequently, content that is primarily video, audio, or image based will be excluded. Additionally, content that is explicitly labelled opinion, op-ed, or analysis will be excluded.

Content that primarily concerns sports, satire, celebrity gossip, or other niche content such as video games will be excluded from all sources. Any content that is not english based or covering niche non-U.S. content will also be excluded. Any content that is excluded will be recorded, along with the reason(s) for its exclusion, and the link to the article made available.

C Feedly (RSS feed)

RSS feeds are a means by which users can pull articles from many websites into a single feed, viewing article titles and times of publication from many different publications in one place. Feedly is a popular RSS feed that has an additional feature that allows users to sort articles based on their popularity. In our study, we have created three separate RSS feeds on Feedly, one for each of the five categories. We will source a daily article, which has been published in the last 24 hours, from each RSS feed for each category based on its popularity according to Feedly's popularity metric. This will ensure that we are supplied with false news that has a variety of ideological perspectives, as well as being both timely and popular.

D Coding Instructions for False News Sources:

We are asking you to classify news websites by their political and/or partisan orientation. In particular, we are asking you to go through a list of known news websites and classify each one by its political ideology. We expect this to take about five minutes per website.

For each website, please do the following:

- Visit each website.
- Record if the website still appears to be an active news website by confirming that there is at least one new post in the past 30 days .
- Read the headlines, articles of interest and the about page of the website.
- **Classify the website in terms of their ideological leaning:**
  - Anti-Russia
  - Not Anti-Russia
  - Right-wing Ultra Nationalism
  - Unclear

You can use the following to make a determination about the website's ideological lean:

- Name of the website
- About page
- Slogan of the website
- Article headlines
- Articles and their content

E Inter-Rater Reliability Statistics For Coding the Ideological Lean of News Sources

Table 6: Inter-Rater Reliability Statistics For Coding the Ideological Lean of News Sources

Coders	Agreement	Cohen Kappa Score	Categories
01 and 02	0.809	0.730	3
02 and 03	0.865	0.810	3
01 and 03	0.775	0.695	3

F Dependent Variables of Interest

Misinformation Rated as True (Y_true_{ia})

The Y_true_{ia} variable is assigned a 1 if an article rated as false/misleading by fact-checkers is rated as true by the respondent. If the article is not rated as false/misleading by fact-checkers this variable is assigned the value of "NA."

Table 7: Veracity Question (Categorical Scale)

Question	Possible Answers
What is your assessment of the central claim in the article?	**(A) True:** The central claim you are evaluating is factually accurate. **(B) Misleading and/or False:** Misleading: The central claim takes out of context, misrepresents, or omits evidence. False: The central claim is factually inaccurate. **(C) Could Not Determine:** You do not feel you can judge whether the central claim is true, false, or misleading.

Perceived Veracity as a Four-Point Ordinal Scale ($Y\_ordinal\_4_{ia}$)

The variable, $Y\_ordinal\_4_{ia}$, is assigned the value of the response chosen from the question below by the respondent (1 through 4).

Table 8: Veracity Question (Ordinal Scale)

Question	Possible Answers
To the best of your knowledge, how accurate is the claim in the above article?	(A) Very accurate (4) (B) Somewhat accurate (3) (C) Not very accurate (2) (D) Not at all accurate (1)

Perceived Veracity as a Seven-Point Ordinal Scale ($Y\_ordinal\_7_{ia}$)

The variable, $Y\_ordinal\_7_{ia}$, is assigned the value of the response chosen from the question below by the respondent (1 through 7).

Table 9: Veracity Question (Ordinal Scale)

Question	Possible Answers
Now that you have evaluated the article, we are interested in the strength of your opinion. Please rank the article on the following scale	**(A) 7 - Definitely True** (7) **(B) 6** (6) **(C) 5** (5) **(D) 4** (4) **(E) 3** (3) **(F) 2** (2) **(G) 1 - Definitely Not True** (1)

G Independent Variables of Interest

Language Preference

We ask the respondent at the beginning of the survey what language they prefer to take the survey in.

We would like you to complete this survey in the language you are most comfortable reading in. Which language is this?
(A) Ukrainian
(B) Russian
(C) Other

The variable, $Pref\_Lang\_UKR_i$, is assigned 1 if the respondent chooses "Ukrainian." The variable, $Pref\_Lang\_RUS_i$, is assigned 1 if the respondent chooses "Russian." The Treatment is $Lang_Treatment$ if the respondent is randomly assigned a news article to evaluate in their non-dominant language.

Concern about Russian disinformation

To determine how concerned a respondent is for Russian disinformation.

Now that you have evaluated the article, we are interested in the strength of your opinion. Please rank the article on the following scale:
(A) 1 - I am not worried about it all
(B) 2
(C) 3
(D) 4
(E) 5
(F) 6
(G) 7 - I worry about it a lot

The variable, $Concern\_Disinfo_i$, is assigned the value of the answer chosen by the respondent (1 through 7).

Age

The Age_i variable is assigned the age of the respondent

Education

The $Education_i$ variable is assigned the education level of the respondent. Possible education levels are as follows: Basic secondary education (up to 9th grade) (1) ; Complete secondary education (grades 10-11) (2) ; Vocational education (vocational school) (3) ; Higher education (bachelor, specialist, master) (4) ; Scientific degree (candidate of sciences and above) (5)

Gender

The $Female_i$ dummy variable is assigned 1 if the respondent self-identifies as female.

Ethnicity

The $Ethnicity_i$ variable. A dummy is assigned 1 if the respondent self-identifies as Ukrainian (multiple allowed).

Ideology of respondent:

We ask individuals three questions to determine their ideology. The ideology score they receive on the ideological scale is an average of their answers in parentheses next to the answer they give.

Was the dissolution of the USSR a mistake?
(A) Yes (1)
(B) Unsure (0)
(C) No (-1)

What is the best political system of the three listed below?
(A) Political system in Russia (1)
(B) Political system in Ukraine (0)
(C) Political systems in the West (-1)

Which party is closest to you?
(A) Opposition Platform - For Life (1)
(B) Opposition Bloc (1)
(C) For the Future (1)
(D) VO Batkivshchyna (-1)
(E) Holos (-1)
(F) European Solidarity (-1)
(G) Samopomich (-1)
(H) Servant of the People (-1)
(I) Svoboda (0)
(J) Trust (0)
(K) United Centre (0)
(L) Bila Tserkva Together (0)
(M) Other: *textbox* (make the determination based on party)

The variable $UKR\_Ideology_i$ is assigned the average values that corresponds to their response. A value of -3 denotes a very anti-Russia ideological perspective and a value 3 denotes a very Pro-Russia ideological perspective. If a respondent receives an ideological score between -3 and -1 they are considered to have an anti-Russia ideological perspective. If a respondent receives an ideological score between 1 and 3 they are considered to have an pro-Russia ideological perspective.

Proficiency and Usage in Ukrainian relative to Russian:

Language ability can be divided into two related concepts: proficiency and usage Andreenkova 2019.

In our mainly analysis we proxy for proficiency by categorizing those who choose to take the Study in Ukrainian as $Pref\_Lang\_UKR_i$ and those that choose to take the study in Russian and $Pref\_Lang\_RUS_i$

We also, however, will measure both self assesses usage and proficiency and analyze both of these variables and their relationship to belief in misinformation as well. We will also examine whether the extent of differences in proficiency/usage matter in exploratory analysis. [1]

We will create an additive score of the following three questions to measure usage.

What language do you speak in daily life/at home?
(A) Only Russian (1)
(B) Mostly Russian with a few Ukrainian words interspersed (.5)
(C) Equal measure Russian and Ukrainian (0)
(D) Mostly Ukrainian with a few Russian words interspersed (-.5)
(E) Exclusively Ukrainian. (-1)
(F) Mostly some other language (0)

Which language or languages do you usually speak outside home – at work, in the street, and in public places?[2]
(A) Only Russian (1)
(B) Mostly Russian (.5)
(C) Equal measure Russian and Ukrainian (0)
(D) Mostly Ukrainian with a few Russian words interspersed (-.5)
(E) Exclusively Ukrainian. (-1)
(F) Mostly some other language (0)

Which language or languages do you think of as your native? (multiple choice is allowed).[3] We score the question 0 if both Russian and Ukrainian are chosen, -1 if Ukrainian but NOT Russian is chosen, and 1 if Russian but NOT Ukrainian is chosen.
(A) Ukrainian
(B) Russian
(C) Crimean Tatar
(C) Hungarian
(D) Bulgarian
(E) Moldavian
(F) Romanian
(G) Other

We measure proficiency as the normalized sum of differences of the following three questions where we ask respondents about both Ukrainian and Russian.[4]

In your opinion, how well can you SPEAK ... (UKRAINIAN/RUSSIAN)?
(A) I SPEAK fluently (3)
(B) I can SPEAK but with minor difficult (2)
(C) I can SPEAK but with major difficulties (1)
(D) I CANNOT SPEAK his language at all (0)

[1] Since all subjects are unconditionally randomized, we should be balanced on these dimensions as well.
[2] If respondents do not work or study, we will omit this question from the calculation.
[3] We take all languages with more that 100,000 estimated speakers.
[4] We adapted these from the **LangUSSR** study cited in Andreenkova 2019 to take into consideration more fine-grained distinctions between Russian and Ukrainian speakers.

And how well can you read texts in ... (UKRAINIAN/RUSSIAN)
(A) I READ fluently (3)
(B) I can READ but with minor difficult (2)
(C) I can READ but with major difficulties (1)
(D) I CANNOT READ this language at all? (0)

And how well can you WRITE texts in ... (UKRAINIAN/RUSSIAN)
(A) I WRITE fluently
(3) (B) I can WRITE but with minor difficult (2)
(C) I can WRITE but with major difficulties (1)
(D) I CANNOT WRITE this language at all? (0)

We calculate the proficiency of the Ukrainian language ($Prof\_UKR_i$) and Russian language ($Prof\_RUS_i$) by taking the average response to these questions about each language. If only two of these questions is answered we still calculate this ideological score, but normalize it, but if only one or more of these questions are answered we omit this response. To measure the relative proficiency of Ukrainian versus Russian ($Relative\_Proficiency\_UKR_i$) we subtract our measure of their proficiency in the Russian language ($Prof\_RUS_i$) from our measure of their proficiency in the Ukrainian language ($Prof\_UKR_i$). To measure the relative proficiency of their dominant language versus their less proficient language ($Relative\_Proficiency_i$) we subtract our measure of their proficiency in their dominant language from our measure of their proficiency in their non dominant language.

Treatment: Tips To Spot Fake News

Tips to spot fake news some respondents are shown to help them identify misinformation.

Be skeptical of headlines. False news stories often have catchy headlines in all caps with exclamation points. If shocking claims in the headline sound unbelievable, they probably are.

Look closely at the URL. A phony or look-alike URL may be a warning sign of false news. Many false news sites mimic authentic news sources by making small changes to the URL. You can go to the site to compare the URL to established sources.

Investigate the source. Ensure that the story is written by a source that you trust with a reputation for accuracy. If the story comes from an unfamiliar organization, check their "About" section to learn more.

Watch for unusual formatting. Many false news sites have misspellings or awkward layouts. Read carefully if you see these signs.

Consider the photos. False news stories often contain manipulated images or videos. Sometimes the photo may be authentic, but taken out of context. You can search for the photo or image to verify where it came from.

Inspect the dates. False news stories may contain timelines that make no sense, or event dates that have been altered.

Check the evidence. Check the author's sources to confirm that they are accurate. Lack of evidence or reliance on unnamed experts may indicate a false news story.

Look at other reports. If no other news source is reporting the same story, it may indicate that the story

is false. If the story is reported by multiple sources you trust, it's more likely to be true.

Is the story a joke? Sometimes false news stories can be hard to distinguish from humor or satire. Check whether the source is known for parody, and whether the story's details and tone suggest it may be just for fun.

Some stories are intentionally false. Think critically about the stories you read, and only share news that you know to be credible.

H Hypotheses and Models

Hypothesis 1: Respondents who are dominant Russian speakers are less likely to rate a false/misleading article written in their less proficient language (Ukrainian) when they are more proficient in their less proficient language

This hypothesis is only tested using evaluations from respondents whose preferred language is Russian on articles rated as false/misleading by professional fact-checkers.

Main:

$$Y\_ordinal\_4_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

Secondary:

$$Y\_true_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

$$Y\_ordinal\_7_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

Hypothesis 2: Respondents who are dominant Ukrainian speakers are less likely to rate a false/misleading article written in their less proficient language (Russian).

This hypothesis is only tested using evaluations from respondents whose preferred language is Ukrainian on articles rated as false/misleading by professional fact-checkers.

Main:

$$Y\_ordinal\_4_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

Secondary:

$$Y\_true_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

$$Y\_ordinal\_7_{ia} = \alpha_a + \gamma_1 * Lang\_Treatment_{ia} * East\_South\_Region_i + \gamma_2 * Lang\_Treatment_{ia} + \gamma_3 * East\_South\_Region_i + \Gamma\mathbf{Z}$$

Hypothesis 3: Ukrainian-dominant speakers in Southern and Eastern Ukraine are less likely to rate a false/misleading article as true written in their less preferred language (Russian) then in their more preferred language (Russian)

This hypothesis is only tested using evaluations from respondents whose preferred language is Ukrainian and are located in an Oblast in the South or East region (as specified in Section XX of this Supplementary Materials) on articles rated as false/misleading by professional fact-checkers.

Main:

$$Y\_ordinal\_4_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

Secondary:

$$Y\_true_{ia} = \alpha_a \ + \ \gamma * Lang\_Treatment_{ia} \ + \ \Gamma\mathbf{Z} \ + \ \epsilon$$

$$Y\_ordinal\_7_{ia} = \alpha_a + \gamma_1 * Lang\_Treatment_{ia} * East\_South\_Region_i + \gamma_2 * Lang\_Treatment_{ia} + \gamma_3 * East\_South\_Region_i + \Gamma\mathbf{Z}$$

Hypothesis 4: Dominant Ukrainian speakers in Southern and Eastern Ukraine and Dominant Ukrainian speakers in the rest of Ukraine are differentially affected reading misinformation in their less proficient language when evaluating the veracity of misinformation.

Each model is run first . Once where Y_true_{ia} represents the categorical measure of veracity and once where $Y\_ordinal_{ia}$ represents the ordinal measure of veracity.

Main:

$$Y\_ordinal\_4_{ia} = \alpha_a + \gamma_1 * Lang\_Treatment_{ia} * East\_South\_Region_i + \gamma_2 * Lang\_Treatment_{ia} + \gamma_3 * East\_South\_Region_i + \Gamma\mathbf{Z}$$

Secondary:

$$Y\_true_{ia} = \alpha_a + \gamma_1 * Lang\_Treatment_{ia} * East\_South\_Region_i + \gamma_2 * Lang\_Treatment_{ia} + \gamma_3 * East\_South\_Region_i + \Gamma\mathbf{Z} + \epsilon$$

$$Y\_ordinal\_7_{ia} = \alpha_a + \gamma_1 * Lang\_Treatment_{ia} * East\_South\_Region_i + \gamma_2 * Lang\_Treatment_{ia} + \gamma_3 * East\_South\_Region_i + \Gamma\mathbf{Z}$$

I Recruitment Information

Every week, we sample 200 respondents evenly from four regions used by Barrington and Herron (2004) (West, Central, South, and East), which can be found in the next section (Section J) of the Supplementary Materials (We exclude rebel-controlled Donetsk, Luhansk, and Crimea). These groups will also be balanced by gender (1/2 who self-identify as male ; 1/2 who self-identify as female), and age (18-35: 1/3 ; 35-55: 1/3 ; 55 plus: 1/3).

This survey firm recruits individuals through various means. They are paid for their participation in either airline miles or direct transfers of money. Not all respondents are paid the same amount as it is up to both the participant and the vendor (*Qualtrics*) to negotiate terms. The handling of this by *Qualtrics* maintains the anonymity of the respondents in our survey. Survey respondents must answer every question on the survey. This sample size maximizes power to test hypotheses outlined in this plan, but is constrained by funds we have available.

J Regional Divides in Ukraine

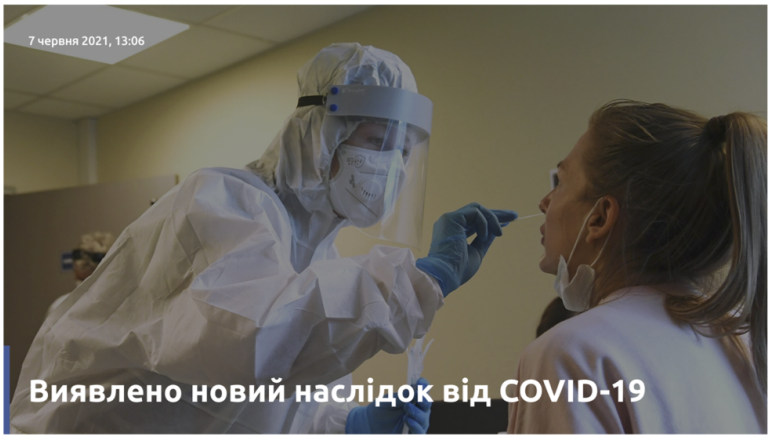
Barrington and Herron (2004) aggregate the following oblasts into four regions:

1. **West:** Volyn, Rivne, Ivano-Frankivsk, Lviv, Ternopil, Chernivetsi, Zakarpatia
2. **Center/North:** Chernihiv, Cherkasy, Poltava, Sumy, Kirovohrad, Kyivska Oblast, Kyiv City, Khmelnyskyi, Vinnytsia, and Zhytomyr.
3. **South:** Odesa, Mykolaiv, and Kherson.
4. **East:** Kharkiv, Dnipropetrovsk, Zaporizhzhia, Donetsk and Luhansk (under Ukrainian control)

Excluded: Krym (Crimea) and Sevastopol

K Article Format (Ukrainian and Russian

Figure 1: Article from Low-Quality Unclear List (written in Ukrainian)



Дослідники з Бірмінгемського університету стверджують, що захворювання COVID-19 може привести до розвитку синдрому Гієна-Барре.

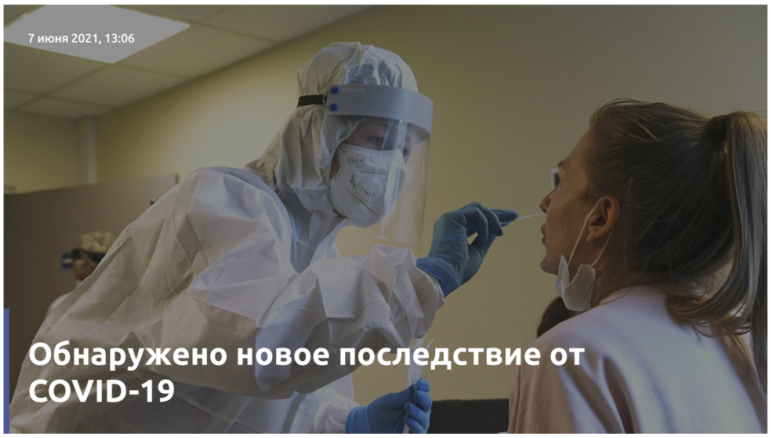
Про це повідомляє кореспондент.

Автор дослідження Алекс Ріхтнер заявив, що антитіла, які організм виробляє після перенесеного коронавируса, схожі з антитілами, що викликають захворювання серця і шкіри.

За результатами дослідження, першими симптомами захворювання Гієна-Барре вважаються оніміння ступень і кистей рук, м'язова слабкість і головний біль, проблеми з координацією. Вони також виявлені в учасників дослідження, які перенесли коронавірус в легкій і важкій формах.

Синдром Гієна-Барре - це аутоімунна запальна полирадикулоневропатия, що протікає в гострій формі і що виявляється млявими парезами, порушеннями чутливості, вегетативними розладами.

Figure 2: Article from Low-Quality Unclear List (written in Russian)



Исследователи из Бирмингемского университета утверждают, что заболевание COVID-19 может привести к развитию синдрома Гийена-Барре.

Об этом сообщает Корреспондент.

Автор исследования Алекс Рихтнер заявил, что антитела, которые организм вырабатывает после перенесенного коронавируса, похожие с антителами, вызывающими заболевание сердца и кожи.

По результатам исследования, первыми симптомами заболевания Гийена-Барре считаются онемение ступней и кистей рук, мышечная слабость и боли, проблемы с координацией. Они также обнаружены у участников исследования, перенесших коронавирус в легкой и тяжелой формах.

Синдром Гийена-Барре — это аутоиммунная воспалительная полирадикулоневропатия, протекающая в острой форме и проявляющаяся вялыми парезами, нарушениями чувствительности, вегетативными расстройствами.

L Balance Table for Pilot Study

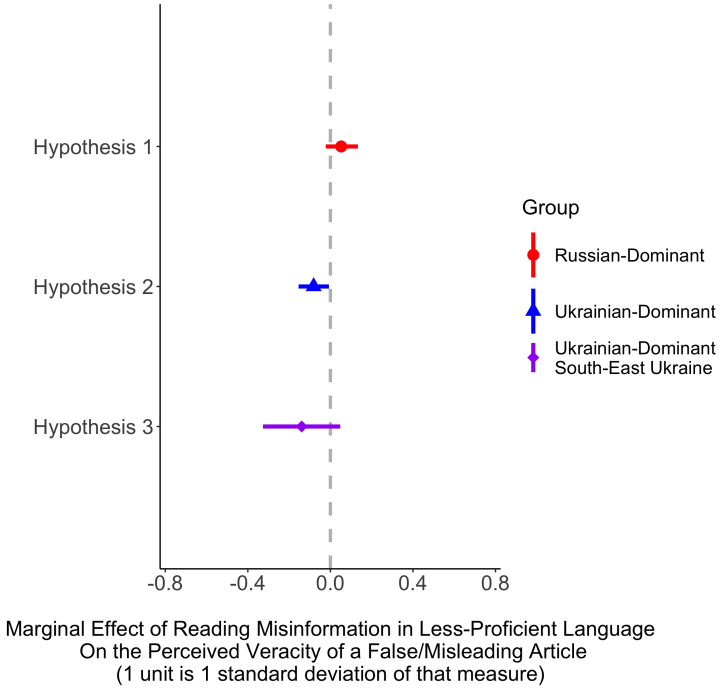
Table 10: Balance Table

Group	Observ.	Age	Income	Gender (Prop. Female)	Education	Ideology	Ethnicity (Ukrainian)	Ethnicity (Russian)
All respondents	422	43.68	3.27	0.50	3.65	-0.02	0.86	0.12
Russian Language Chosen	194	44.49	3.23	0.46	3.61	-0.12	0.73	0.25
Ukrainian Language Chosen	228	42.97	3.31	0.54	3.69	0.05	0.97	0.01
Russian Language Chosen (Control)	106	45.09	3.25	0.46	3.61	-0.17	0.72	0.25
Russian Language Chosen (Treatment)	88	43.77	3.20	0.47	3.60	-0.06	0.75	0.25
Ukrainian Language Chosen (Control)	109	43.29	3.43	0.53	3.80	0.07	0.98	0.01
Ukrainian Language Chosen (Treatment)	119	42.68	3.20	0.55	3.59	0.04	0.97	0.01

M Figures for Exploratory Analysis and Robustness Checks

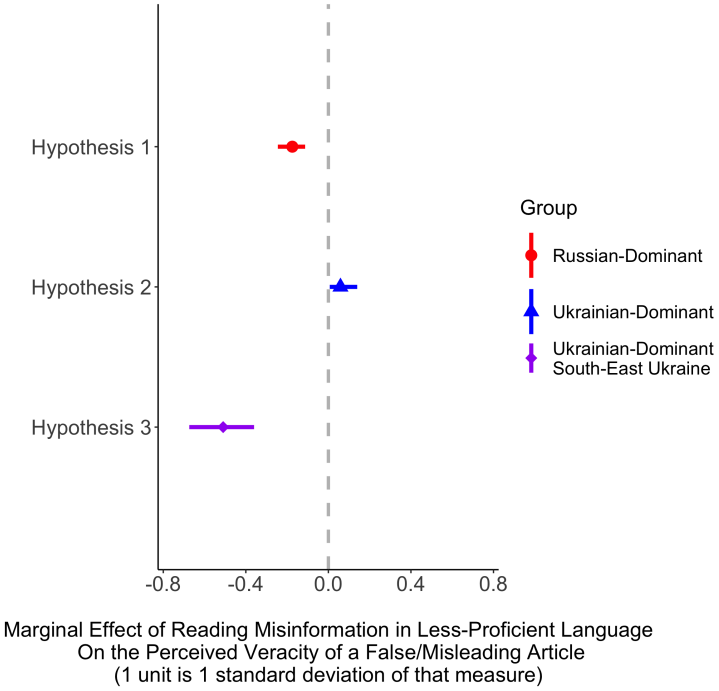
M.1 Exploratory Analysis 2: Effect of Language on Belief in True Information

Figure 3: Predicted Pooled Marginal Effect of Evaluating *True* News Articles in One's Less Preferred Language



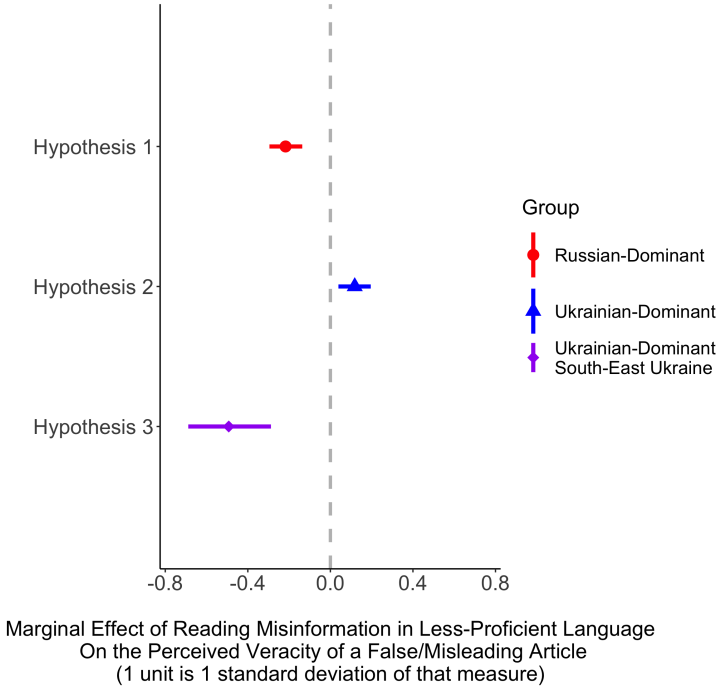
M.2 Robustness Check 1 (Substituting Proficiency)

Figure 4: Predicted Pooled Marginal Effect of Evaluating *True* News Articles in One's Less Preferred Language



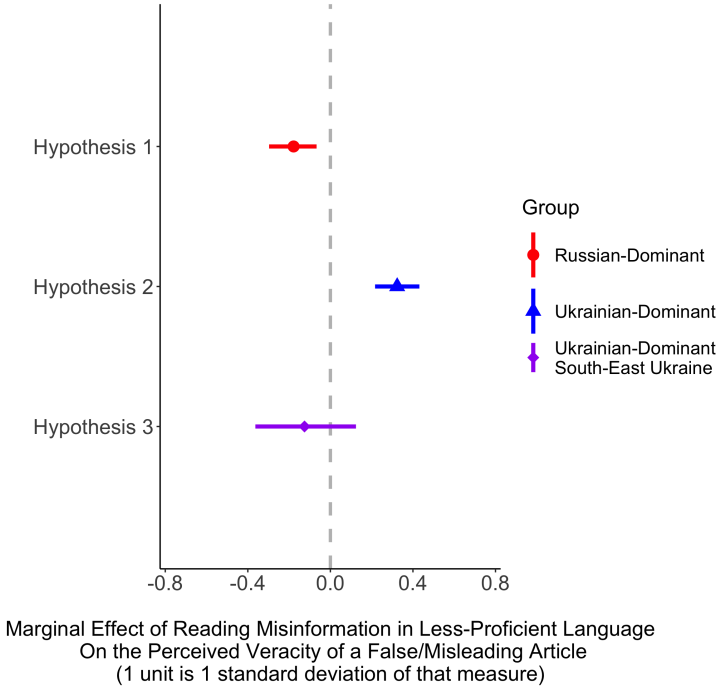
N Robustness Check 2 - Control for Demographics and Digital Media Literacy Intervention

Figure 5: Predicted Pooled Marginal Effect of Evaluating *True* News Articles in One's Less Preferred Language



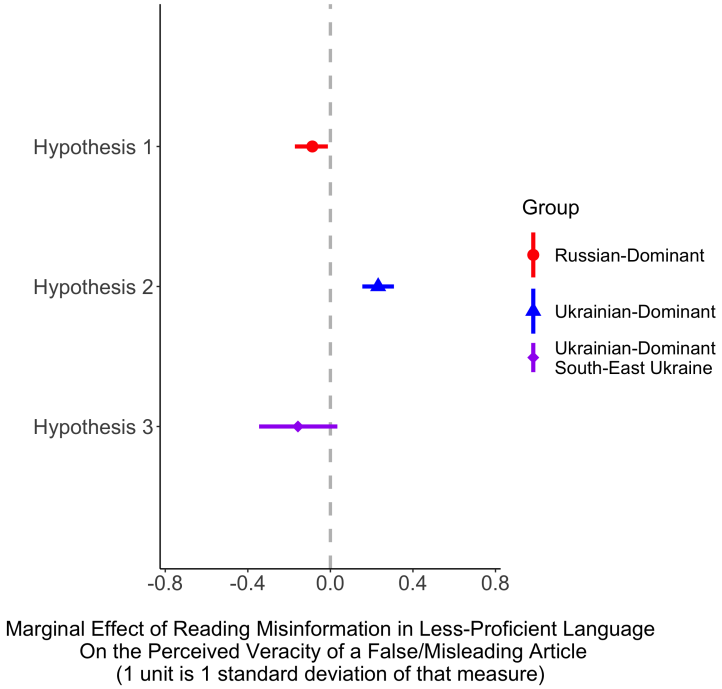
O Robustness Check 3 - Only Those Who Were Not Treated By Digital Media Literacy Intervention

Figure 6: Predicted Pooled Marginal Effect of Evaluating *True* News Articles in One's Less Preferred Language



P Robustness Check 4 - Replace Dependent Variable with Categorical Measure

Figure 7: Predicted Pooled Marginal Effect of Evaluating *True* News Articles in One's Less Preferred Language



Q Articles Evaluated during Pilot Study:

Table 11: Headlines for Articles Chosen from the Low Quality Liberal News Stream in Study I

News Stream	Date	Headline	Modal Fact-Checker Rating	Topic	Lean of Article
Not Anti-Russia Low-Quality	06/07/21	Zelensky will impose a tax on workers - Kiev experts	False/Misleading	Economics	Neutral
Anti-Russia Low-Quality	06/07/21	The Russian Federation is “unaware” of the shape of Ukraine – Says the Ukrainian Defense Ministry	True	Politics	Anti-Russia
Unclear Low-Quality	06/07/21	A new side-effect of COVID-19 has been discovered	Could Not Determine	Science	Neutral
Russian Mainstream	06/07/21	The Ukrainian national football team will participate at the European Championship with the symbol of a nationalist slogan	False/Misleading	Politics	Not Anti-Russia
Ukrainian Mainstream	06/07/21	He is not worthy! KNU has deprived Lukashenko of the title of honorary doctor	True	Politics	Anti-Russia

R Article Selection Process for Experiment

To select popular true and false/misleading articles from a diverse set of ideological perspectives we send out the most popular article from five different lists of news sources: (1) Not Anti-Russia low credibility; (2) Anti-Russia low-credibility news sources; (3) General low credibility news sources with an ambiguous or unclear ideological affiliation; (4) Ukraine-based mainstream sources; and (5) Russia-based mainstream sources.

All low-credibility news sources are selected from Texty, but we additionally use trained coders to classify each web-based news source into one of three ideological perspective categories ("anti-Russia", "not anti-Russia", and "unclear") that are based on a longstanding distinction in Ukrainian politics (Erlich and Garner 2021). Although we do not know where all of these news sources are based, we do know that the majority of "anti-Russia" low-credibility news sources are based in Ukraine and the majority of the "not anti-Russia" low-credibility news sources are based in Russia. Coders were asked to use the headlines, the content of its articles, as well as the websites domain and 'about' page to classify websites. [5] Websites were only classified as "not anti-Russia" or "anti-Russia" when at least two of the three coders agreed. If a website did not receive at least two "not anti-Russia' ratings or two "anti-Russia" ratings it was deemed to have an "unclear" ideological lean. In a pilot study, all three coders were in complete agreement with each-other 67.3 % with 0.488 Fleiss' Kappa score. Although this kappa score is low in terms of reliability, it is somewhat expected as some news sources often publish stories with different ideological perspectives. That is why we place any source that is not rated as " anti-Russia" or "not anti-Russia" by at least two of the three research assistants as "unclear". In total, during the pre-test, we categorize eight domains as "anti-Russia–low-credibility" news streams, thirty-eight domains as "not anti-Russia– low-credibility" news streams, and nine domains as "unclear–low-credibility" news streams."[6] The prevalence of "not anti-Russia low-credibility" news streams comports with previous research that reports high levels of production of pro-Kremlin misinformation in Ukraine (Snegovaya 2015; StratCom 2015).

Our two mainstream news articles are sourced from a list of popular Ukraine-based mainstream news sites and a list of Russia-based mainstream news sites consumed in Ukraine, both derived from an independent data journalism platform- Texty.[7]

Using these five lists of online news websites we place each list into its own distinct RSS feed (Feedly) and source the most popular articles from each list every week our survey runs. 8 displays the weekly process of selecting articles and distributing these articles to both professional fact-checkers and respondents. Every week, on Tuesday morning we will source the most popular articles published on Monday from each of the five lists of sources: (1) Not Anti-Russia low credibility news sources (that may publish misinformation) ; (2) anti-Russia low-credibility news sources (that may publish misinformation) ; (3) General low credibility news sources (that may publish misinformation) with an ambiguous or unclear ideological affiliation ; (4) Ukraine-based mainstream sources (that likely publishes true information) ; and (5) Russia-based mainstream sources (that likely publishes true information) consumed in Ukraine. The most popular articles 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 query respondens with the exact same question wording. 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'.[8] 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).[9] Using this process, respondents evaluate

[5] Section D, Supplementary Materials contains the complete coding instructions given to these coders.

[6] Section A, Supplementary Materials lists each domain and their ideological perspective.

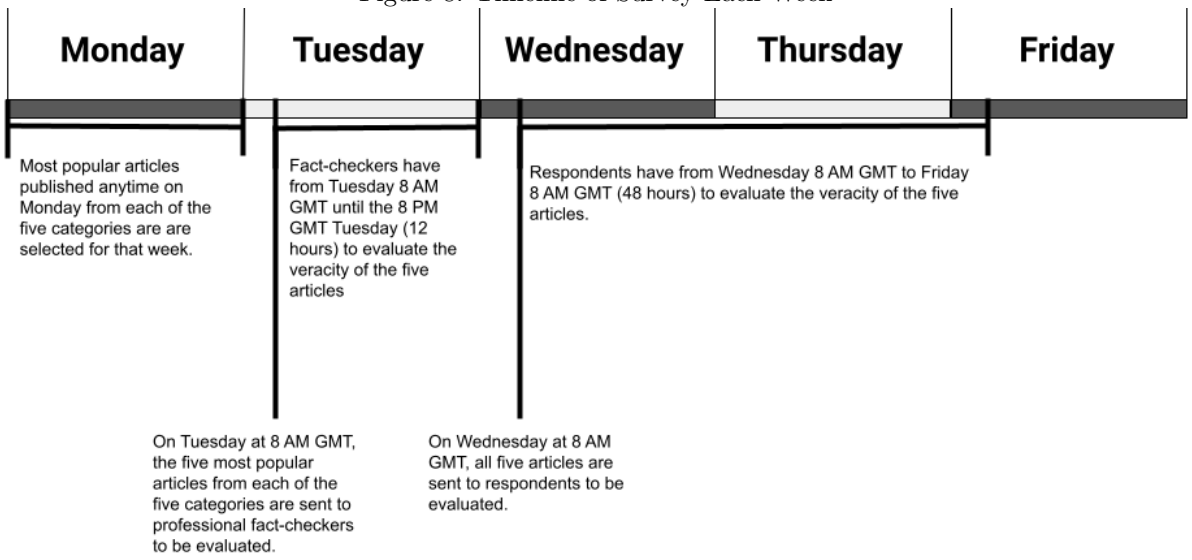
[7] `https://topic-radar.texty.org/#/`

[8] Professional 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.

[9] Articles are translated to Ukrainian or Russian if the article is only available in one language on the website. They are

articles within 48-96 hours of publication.

Figure 8: Timeline of Survey Each Week



translated by one translator and then checked by another translator.