The relationship between internet usage and trust in news media

Anna Lamb

BDS 522

Final Project

Introduction: Research Question

Hypothesis 1: The more time that an individual spends online, the lower their trust in news media

Why Does This Matter?

Misinformation has proliferated online, as fake news spreads faster and farther than true stories.[1] Moreover, research suggests that exposure to online fake news is associated with a decrease in trust of mainstream media.[2] Fake news exposure has been shown to influence a person's beliefs, which can affect their subsequent behavior.[3] The anti-mask movement and vaccine hesitancy we've witnessed during the COVID-19 pandemic might be two consequences of this link between misinformation and the growing mistrust in media. Understanding the association between time spent online and trust in news media and possible confounding variables can help inform policy efforts to curb the spread of fake news.

[1]Dizikes, P. (2018, March 8). *Study: On Twitter, false news travels faster than true stories*. MIT News. Retrieved December 23, 2021, from https://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308

[2]Ognyanova, K., Lazer, D., Robertson, R. E., & Wilson, C. (2020). Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harvard Kennedy School (HKS) Misinformation Review*. https://doi.org/10.37016/mr-2020-024

[3]Weeks, B. E., & Garrett, R. K. (2014). Electoral Consequences of Political Rumors: Motivated Reasoning, Candidate Rumors, and Vote Choice during the 2008 U.S. Presidential Election. *International Journal of Public Opinion Research*, *26*(4), 401–422. doi: 10.1093/ijpor/edu005

General Society Survey (GSS) Data

About the GSS

- A nationally representative survey of US adults conducted since 1972

- Designed to monitor and explain changes in both social characteristics and attitudes currently being conducted in the US

- Includes a standard core of demographic, behavioral, and attitudinal questions, plus topics of special interest (e.g., civil liberties, psychological wellbeing, and national spending priorities)

Data Collection

- The data comes from the 2021 GSS Survey, collected between December 2020 and May 2021 as a cross-sectional survey

- The sample population includes adults (18+) in the US who live in noninstitutional housing

- Data was collected online instead of through in-person interviews due to the COVID-19 pandemic

- The original sample included 4,032 respondents, but after excluding missing data and outliers, the final sample size was 1676 respondents.

Variable Measurements

Dependent & Independent Variables

- **Dependent Variable—Trust in News Media** (*TRMEDIA*): How much do you personally trust the following institution? The news media
  - Categorical on a scale 0-10; 0 means you do not trust an institution at all, and 10 means you trust it completely
  - Converted to binary for logistic regression, where 0-4 = 1 (do not trust), and 5-10 = 0 (everything else)

- **Independent Variable—Time Spent Online** (*WWWHR*): Not counting e-mail, about how many hours per week do you use the Web? (Include time you spend visiting regular web sites and time spent using interactive Internet services like social media, streaming services, chat rooms, online conferencing services, discussion boards or forums, and the like.)
  - Continuous variable, measured in whole hours

Confounding Variables

- **PARTYID:** Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?
  - Categorical, Scale 0-7, where 0 = strong democrat and 6 = strong republican, with 7 = other
  - Converted to a nominal variable, where 0-1 = democrat, 2-4 = independent, 5-6 = republican, and 7 = other.

- **NEWS:** How often do you read the newspaper--every day, a few times a week, once a week, less than once a week, or never?
  - Categorical, Scale 1-5 (1 = every day, 5 = never)

- **SEX: 1 = male, 2 = female**
  - Recoded as dummy *male*, where 1 = male and 0 = female

Methods: Variable Assumptions

- GSS 2021 data is representative of US population

- Exploratory analysis showed that *trust in media* and *time spent online* followed a non-normal distribution, but **normality is not required for logistic regression**.
    - Converting *TRMEDIA* to a binary variable, I assumed scores less than or equal to 4 represented "mistrust" and scores above this captured everyone else (either they neither trusted nor mistrusted, or they trusted media)

- **Missingness**
    - Non-response rate for the dependent variable was over half (55%), and 39% for the independent variable
    - Non-response rates for the confounding variables were: 1% for political party ID, 33% for news consumption, and 2% for sex

- **Outliers:** A boxplot of the independent variable, WWWHR, revealed the presence of outliers. I assumed these were made in error (e.g., the maximum value = 168, which meant the respondent was online 24 hours/day/week). I used the IQR to calculate outlier fences, which identified 71 as the outer cutoff. 30 observations were excluded.

- The variable PARTYID was self-reported on a scale of 0-7, which was converted into a nominal variable to represent political affiliation. I assumed those who scored themselves as: 0 or 1 were Democrats, 2-4 were independents, 5-6 were Republicans, and 7 identified with a third party.

- Alpha level was set to 0.05 for analyses

Methods: Data Analysis

Spearman Rank Correlation

- Since *TRMEDIA* is an **ordinal** variable and *WWWHR* is **continuous**, the nonparametric Spearman Rank Correlation test was conducted to understand **the strength and direction of association** between the dependent and independent variables.

- **Assumptions:**
  - Two variables are ordinal and/or continuous
  - Two variables represent paired observations
  - There is a monotonic relationship
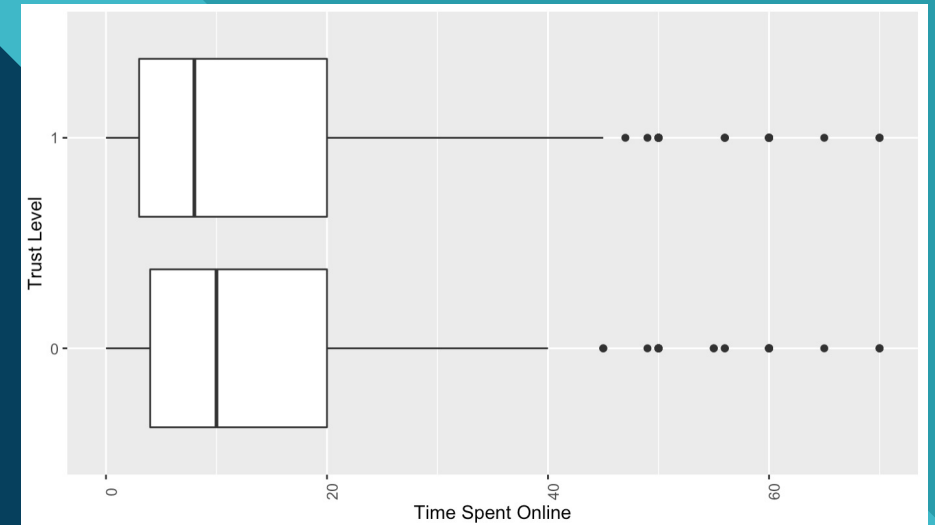
Logistic Regression

- Two models were run as the primary analyses to further study the association
  - The first model tested if **time spent online impacted the likelihood a respondent mistrusted the media**, as TRMEDIA was converted into a binary variable
  - The second model added **confounding variables** (political party ID, news consumption, and sex)

- **Assumptions:**
  - Sample size is sufficient
  - Dependent variable is binary
  - Observations are independent
  - There is little-to-no multicollinearity among explanatory variables
 - No extreme outliers

Initial Results

Exploratory Analysis: A boxplot of the two groups of the binary variable *media (1=mistrusting, 0 = everyone else)* shows that the median time spent online of those who do not trust the media (*median = 8*) is lower than the second group (*median = 10*). This suggests the opposite direction of the hypothesis.

Spearman Rank Correlation: There is a statistically significant relationship between the time spent online and a respondent's trust in news media (*rho = 0.067, p = 0.006*). Though this represents a positive association—trust increases as time spent online increases (opposite direction of *H1*)—the coefficient value suggests an <u>*extremely*</u> small, negligible association.

Time Spent Online – Mistrustful Group vs. Everyone Else

Correlation Table
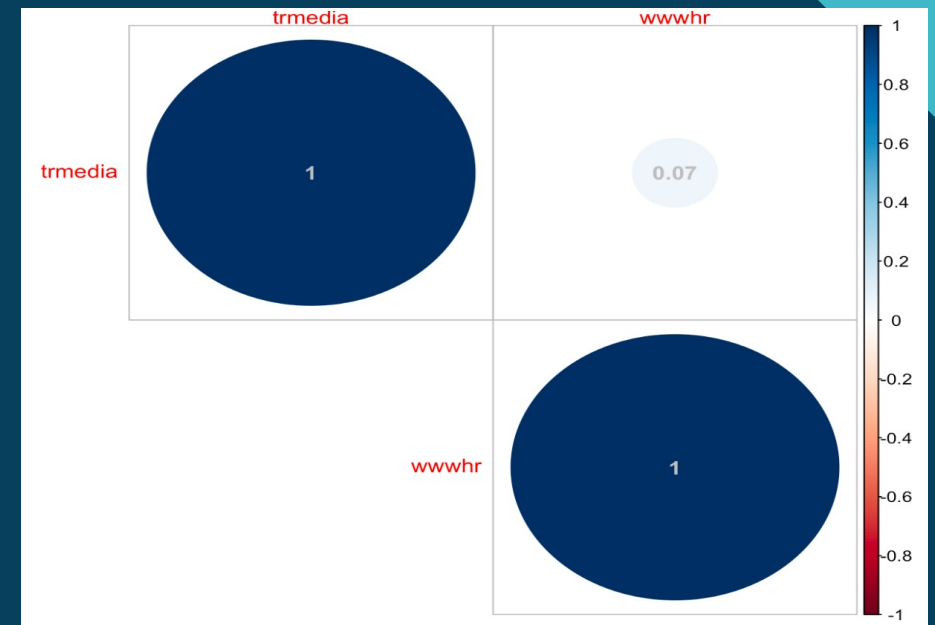
Logistic Regression Results

- The first model suggests that the odds of mistrusting the media decreases as time spent online increases. That is, each additional hour is associated with a .8% lower odds of not trusting the media. However, **this association was insignificant and not substantive** (*B = -0.007, p = 0.05*).

- The results of a multiple logistic regression with 3 confounders confirmed that there is **no *association* between time spent online and trust in media**.

- However, there were **significant associations** between the dependent variable and **political party affiliation, newspaper consumption, and sex**.

- The odds that Republicans, independents, and third party respondents mistrust the media is higher compared to Democrats, all other variables held constant. For example, the odds a GOP respondent mistrusts the media is 1462% higher than the odds for a Democrat respondent (*B = 2.749, p < 0.001).*

- The less frequent that people read the newspaper, the higher likelihood they are distrustful. For example, the odds that someone who never reads the newspaper (news=5) mistrusts the media is 183% higher than the odds of someone who reads it everyday (*B = 1.042, p < 0.001*).

- Males are less likely to mistrust the media than females (*B = -0.257, p = 0.03*).

- The AIC value has decreased between Model 1 and 2, suggesting the latter is better fitted relative to the former.

	Dependent variable:	
	media	
	(1)	(2)
wwwhr	-0.007	-0.004
	(0.004)	(0.004)
factor(demrep)gop		2.749***
		(0.183)
factor(demrep)ind		1.208***
		(0.123)
factor(demrep)other		2.191***
		(0.377)
factor(news)2		0.349
		(0.184)
factor(news)3		0.651**
		(0.229)
factor(news)4		0.822***
		(0.178)
factor(news)5		1.042***
		(0.145)
male		-0.257*
		(0.116)
Constant	0.423***	-1.154***
	(0.070)	(0.151)
Observations	1,676	1,676
Log Likelihood	-1,137.699	-934.740
Akaike Inf. Crit.	2,279.397	1,889.479
Note:	*p<0.05; **p<0.01; ***p<0.001	

Conclusions

- When investigating the association between time spent online per week and mistrust of the media, both models imply that each additional hour spent online makes it less probable a respondent is mistrustful. However, this difference is neither substantively nor statistically significant. Therefore, there appears to be no association between time spent online and media trust.

- Controlling for confounders makes the coefficient of time spent online nearly 0, and seem to more accurately predict a person's likelihood of trusting media.

- Democrats are more trusting relative to other political affiliations. This confirms previous findings.[1]

- Newspaper consumption is linked to media trust, which intuitively makes sense—someone who doesn't trust the media is unlikely to spend their time reading a newspaper.

- Additional research is needed to understand how someone's exposure to misinformation is shaped by the time spent online, and if this is linked to their perceptions of news media.

[1]Gottfried, J., & Liedke, J. (2021, August 30). *Partisan divides in Media Trust Widen, driven by a decline among Republicans*. Pew Research Center. Retrieved from https://www.pewresearch.org/fact-tank/2021/08/30/partisan-divides-in-media-trust-widen-driven-by-a-decline-among-republicans/

Limitations

- The independent variable, *WWWHR,* did not separate the time spent online per week for leisure versus work/study. A future study should segment *WWWHR* into two questions: (1) how much time a respondent spends online for work/study purposes, versus (2) leisure, as the latter might more accurately investigate if there is an association with trust in news media.

- Only 1 demographic variable (sex) used in analysis. Future research should investigate if additional demographics (e.g., race and geographic location) impact the relationship between time spent online and trust in news media.

- Dependent variable limited by the binary grouping generated for logistic regression.

- Since most people have spent the past 2 years at home and limited in their ability to do in-person events, it's likely they have spent more time online than prior to the pandemic, regardless of how they perceive the media. To understand if 2021 is an outlier year for the GSS, a time-series analysis with pre-pandemic data should be conducted to investigate if the association changed over time.

Summary

- **Hypothesis:** The more time an individual spends online, the lower their trust in news media

- **Sample:** 1676 US respondents (18+) from the 2021 GSS

- **Methods:** A Spearman Rank Correlation test to investigate the correlation between variables, followed by logistic regression including a basic model and one with confounders.

- **Findings:** There was a significant, positive relationship between time spent online and trust in media in the correlation test, but it was not substantive. Once analyzed with the logistic regression, the relationship between the two variables was insignificant.

- **Conclusion**: Time spent online is not associated with the level of trust in media. Partisanship, gender, and newspaper consumption more accurately predict the likelihood a person is trustful or not.

- **Limitations**: Assumed binary grouping of dependent variable; 2021 might be an outlier year; time spent online *for leisure* might more accurately measure an association (if any).