**Methods**

**Survey Design**

This study relies on a novel design featuring an online rolling cross-sectional survey of adult social media users in the United States. The survey was administered in 17 sampling frames of 3-4 days each (with Saturday/Sunday counted as one day). This approach affords a unique opportunity to link survey responses and social media content, because it is possible to ‘locate’ responses in temporal proximity to actual content known to be circulating on social media. The research team collected external media data from Facebook’s proprietary platform, CrowdTangle. In each sampling frame we embedded a screenshot of (one of) the most popular news articles in the previous three day with source cues photoshopped out of the image. Because most of the stories came from a single news organization (Fox News), we balanced this dynamic by also including the most popular story from any other news organization (e.g., CNN, *New York Times*, *Washington Post*, etc.), and, thus, in each sampling frame respondents were randomly presented with either the most popular Fox News story or the most popular story from some other organization that was circulating immediately prior to the sampling frame. Respondents were told the story had been “recently circulating on Facebook”—a true statement. This method affords us the ability to develop several cued recall measures, which form key outcome variables.

**Sample and Data**

Data were collected between September 3 and November 1, 2020. Quotas for age, race, gender, and census region were based on the 2018 American Community Survey. The survey has an incidence rate of 100% and a cooperation rate of 70%. It has an overall sample size of *N* = 2,008 with at least *n* = 100 for each sampling frame. The sample reflects the target population on the quota criteria (see Table A1 in the online appendices). The average respondent is between 45 and 54 years old (measured on a 7-point scale where 1 = *18-24* and 7 = 85 or older). Approximately 40% of the sample are persons of color, and 51% are female. Census regions were defined according to the U.S. Census Bureau’s map, and in our sample ~22% of respondents live in the Midwest, ~19% live in the Northeast, ~37% live in the South, and ~23% live in the West. In terms of non-criterion demographics, the average respondent has either some college or a 2-year associate’s degree or trade school diploma (*M* = 4.5 on a 7-point scale where 1 = *Some high school* and 7 = *Post-graduate degree*), and lives in a household that earns between $45,000 and $75,000 per year (*M* = 4.7 on an 8-point scale where 1 = *Less than $15,000* and 8 = *More than $150,000*). The sample underrepresents low-education and low-income individuals, and therefore the data were weighted by education and income (see Table A2 online for the weighting scheme). Missing values were imputed using a chained equations technique (Fully Conditional Specification; see van Buuren & Groothuis-Oudshoorn, 2011).

**Measures**

***Exposure and Engagement***

Prior literature suggests that news exposure has both ‘trait-like’ and ‘state-like’ properties (Weeks & Lane, 2020), and we included both kinds of indicators in our study. On the trait-like side, *total exposure* to political information was measured with six questionnaire items asking respondents how often in the past week they have encountered the following types of information (0 = *Never* and 4 = *Several times a day*): information critical of a candidate they support; information critical of a candidate they oppose; information supportive of a candidate they support; information supportive of a candidate they oppose; information that disagrees with their political views; and information that agrees with their political views (c.f., Weeks et al., 2017). These items were averaged for each respondent, and the resulting scale is reliable (Cronbach’s alpha = .96). The variable has a mean of 1.8 (*SD* = 1.3).

It is important to clarify whether exposure was incidental (Nanz & Matthes, 2020). Therefore, immediately after answering the above battery of questions, respondents were asked a follow-up question: “On social media, some people intentionally search for news or political information, but others come across such information accidentally. What about you?” (0 = *Always intentionally* and 4 = *Always accidentally*). To create a measure of *incidental exposure*, this item was multiplied by the total exposure scale, and then the square root was calculated to maintain the original 5-point metric. The variable has a mean of 1.5 (*SD* = 1.1).

The state-like measures centered on the embedded story stimuli. Once shown the story, respondents were asked whether they had seen it on Facebook. A follow-up question asked whether they had seen it on some other social media platform, and answers to these two items were recoded so that 1 = *Exposed* and 0 = *Not exposed*. Approximately 42% of the sample reported *story exposure*.

Those respondents who reported exposure were asked a series of additional follow-up questions, from which our measures of *incidental exposure* and *engagement* are created. First, this subset of respondents was asked: “When you say the story, were you purposefully seeking information on this topic?” (1 = *Yes* and 0 = *No*; 54% of subset and 23% of full sample said yes). Next, they were asked: “When you saw the story, did you engage in any of the following activities?” (1 = *Yes* and 0 = *No*): click on the story; scan the headline of the story; read the entire story; seek out additional information about the topic; comment on the post; discuss the story; and share the story. Responses were summed for each respondent (Cronbach’s alpha = .79; *Min*. = 0 and *Max*. = 7), and the variable has a mean of 3.5 (*SD* = 2.2). Additionally, a *high-effort engagement* variable was created to isolate those activities that relatively higher amounts of cognitive or behavioral effort, including information seeking, commenting, discussing, and sharing (Cronbach’s alpha = .76; *Min*. = 0, *Max*. = 4, *M* = 1.5, *SD* = 1.5).

***News Involvement***

The study includes four indictors of involvement with news. First, the study measures respondents’ perceptions of *social media as news source* (Weeks & Lane, 2020) by asking them which choice best describes their “reason for accessing [their] social media accounts”: to follow news and public affairs information; for some other reason; do not think very much about the reason. This three-level factor was recoded into a binary variable (1 = *News source* and 0 = *Not news source*). A little more than one-third of the sample (35%) views social media as a news source. Second, the study measures *self-reported interest* with a three-item scale. Questions asked respondents how interested they are (1 = *Not at all interested* and 5 = *Very interested*) in news, politics, and local community. These three items were averaged for each respondent (Cronbach’s alpha = .83) and the variable has a mean of 3.5 (*SD* = 1.0). Third, the measures the extent which respondents *follow accounts for news* with a three-item scale. Items ask how often respondents follow accounts on social media because they are interested in what they post about news or current affairs, politics, and community events (1 = *Never* and 5 = *Very Frequently*). These three items were averaged for each respondent (Cronbach’s alpha = .91), and the scale has a mean of 2.7 (*SD* = 1.2). Finally, we measured *algorithmic categorization* using a technique pioneered by Thorson and colleagues (2021). With the aim of obtaining an observable indicator of Facebook’s classification algorithm, we asked respondents at the end of the survey to open a web browser, navigate to the Settings menu of their Facebook accounts, and locate their Ad Interests section. We then asked them whether they saw the following categories included in their interests: (a) politics; (b) news or specific news organizations (e.g., the *New York Times*); or (c) neither. We coded this item into a binary variable (1 = *News or politics* and 0 = *Neither*). A little less than half (41%) of the sample was categorized as interested in news or politics.

***Covariates***

Prior research has shown that social networks structures are predictors of incidental exposure, and they also related to news involvement (Barnidge & Xenos, 2021). Therefore, it is important to include indicators of social network structures as covariates in the analysis, and we included four such variables. First, network size was measured by asking respondents how many people or accounts they are “friends with,” “follow,” or “subscribe to” on six social media platforms (1 = *None* and 7 = *2,001 or more*). Respondents’ answers to these items were averaged to create a scale (Cronbach’s alpha = .91), which was then unobtrusively logged to correct for skew (*Min*. = 0 and *Max*. = 1.9). The final variable has a mean of 0.7 (*SD* = 0.5). Second, a structural measure of *network diversity* was borrowed directly from prior literature (Hampton et al., 2011). The measure uses a standardized list of 22 occupations and asks respondents whether they are connected someone on social media who belongs to each (1 = *Yes* and 0 = *No*). An averaged scale was created from these items (Cronbach’s alpha = .92), which has mean of 0.3 (*SD* = 0.3). Third, *group activity* on social media was measured with an 8-item scale, where questions asked respondents whether they had discussed news or related topics during the past month in various types of groups. These items were summed for each respondent, and the resulting index was then unobtrusively logged to normalize the distribution. The final variable has a mean of 0.5 (*SD* = 0.6). Finally, social news curation is a concept taken from Thorson and Wells’ (2016) influential work on curated flows. The variable relies on 5 items that ask respondents how much (1 = *None at all* and 5 = *Almost all*) of the content their friends post is about the following topics: the 2020 election; politics or current affairs; social or community issues; racial or social justice issues; and COVID-19. The items were averaged for each respondent (Cronbach’s alpha = .92), and the variable has a mean of 2.9 (*SD* = 1.1).

***Controls***

In addition to *age*, *race*, *gender*, *education*, and *income* (see above for descriptive statistics), the analyses control for *political ideology*, *party identity*, and *frequency of social media use*. Political ideology was measures with a single 11-point L-R scale where -5 = *Very liberal* and 5 = *Very conservative* (*M* = 0.2, *SD* = 3.0). Party identity was measured with three questionnaire items borrowed from the American National Elections Survey. The first asked about party identity (*Democrat*/ *Republican*/ *Other*/ *None*). The second, shown only to those who selected *Democrat* or *Republican* asked about the strength of identity (*Strong*/ *Not that strong*). The third, shown only to those who selected *Other* or *None*, asked about party lean (*Closer to Democrat*/ *Closer to Republican*). These items were then coded to create a 7-point scale where -3 = *Strong Democrat* and 3 = *Strong Republican* (*M* = -0.3, *SD* = 2.0). Finally, frequency of social media use was measured with a single item asking respondents how much time per day they spend actively using social media (1 = *Less than 10 minutes* and 6 = *More than three hours*. The variable has a mean of 3.5 (*SD* = 1.6).

**Analysis and Results**

text here

**References**

Barnidge, M., & Xenos, M. A. (2021). Social media news deserts: Digital inequalities and incidental news exposure on social media platforms. *New Media & Society*. Advance online publication.

Hampton, K. N., Sessions, L. F., & Her, E. J. (2011). Core networks, social isolation, and new media: How Internet and mobile phone use is related to network size and diversity. *Information, Communication & Society, 14*(1), 130-155.

Nanz, A., & Matthes, J. (2020). Learning from incidental exposure to political information in online environments. *Journal of Communication, 70*(6), 769-793.

Thorson, K., Cotter, K., Medeiros, M., & Pak, C. (2021). Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society, 24*(2), 183-200.

Thorson, K., & Wells, C. (2016). Curated flows: A framework for mapping media exposure in the digital age. *Communication Theory, 26*(3), 309-328.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of statistical software, 45*, 1-67.

Weeks, B. E., & Lane, D. S. (2020). The ecology of incidental exposure to news in digital media environments. *Journalism, 21*(8), 1119-1135.

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., & Kwak, N. (2017). Incidental exposure, selective exposure, and political information sharing: Integrating online exposure patterns and expression on social media. *Journal of Computer-Mediated Communication, 22*(6), 363-379.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1  *Group Differences in Incidental Exposure and Total/Story Exposure among the Involvement Groups* | | | | | | | | | | | | |
|  | Incidental Exposure | | | | | | Total Exposure | | | Story Exposure | | |
|  | Trait-Like Variable | | | State-Like Variable | | | Trait-Like Variable | | | State-Like Variable | | |
| **Fixed Effects** | β | | *SE* | β | | *SE* | β | | *SE* | β | | *SE* |
| Intercept | -0.15 | | 0.11 | -2.44\*\*\* | | 0.26 | -0.34\*\* | | 0.12 | -2.06\*\*\* | | 0.21 |
| Involvement (Medium:Low) | -0.09 | | 0.06 | 0.34\*\* | | 0.12 | 0.29\*\*\* | | 0.06 | 0.52\*\*\* | | 0.10 |
| Involvement (High:Low) | -1.07\*\*\* | | 0.09 | 0.22 | | 0.21 | 0.44\*\*\* | | 0.09 | 0.58\*\*\* | | 0.15 |
| Age | -0.07\*\*\* | | 0.02 | 0.09 | | 0.04 | 0.08\*\*\* | | 0.02 | 0.00 | | 0.03 |
| Gender (1 = Female) | 0.03 | | 0.05 | -0.16 | | 0.10 | -0.07 | | 0.04 | -0.23\*\* | | 0.08 |
| Race (1 = Person of Color) | -0.13\*\* | | 0.05 | -0.08 | | 0.10 | -0.20\*\*\* | | 0.04 | -0.01 | | 0.02 |
| Education | 0.05\*\* | | 0.02 | 0.00 | | 0.03 | 0.04\*\* | | 0.01 | 0.02 | | 0.02 |
| Income | 0.00 | | 0.01 | -0.04 | | 0.03 | 0.00 | | 0.01 | -0.03 | | 0.02 |
| Ideology (+ Conservative) | -0.03\*\* | | 0.01 | -0.02 | | 0.02 | -0.02\*\* | | 0.01 | 0.01 | | 0.01 |
| Party Identity (+ Republican) | 0.05\*\*\* | | 0.01 | 0.01 | | 0.03 | 0.04\*\*\* | | 0.01 | -0.00 | | 0.01 |
| Frequency of Social Media Use | 0.06\*\*\* | | 0.01 | 0.06 | | 0.03 | 0.05\*\*\* | | 0.01 | -0.01 | | 0.02 |
| Network Size | -0.19\*\*\* | | 0.07 | -0.14 | | 0.15 | 0.22\*\*\* | | 0.07 | 0.16 | | 0.11 |
| Network Diversity | 0.19 | | 0.10 | -0.03 | | 0.22 | 0.22\* | | 0.10 | 0.33\* | | 0.15 |
| Group Activity | 0.22\*\*\* | | 0.05 | 0.07 | | 0.10 | 0.24\*\*\* | | 0.04 | 0.09 | | 0.07 |
| Social News Curation | 0.41\*\*\* | | 0.03 | 0.07 | | 0.06 | 0.48\*\*\* | | 0.03 | 0.22\*\*\* | | 0.05 |
| Incidental Exposure (Trait-Like) |  | |  | 0.25\*\*\* | | 0.05 |  | |  | 0.02 | | 0.04 |
| **Random Effects** | *Var.* | | *SD* | *Var.* | | *SD* | *Var.* | | *SD* | *Var.* | | *SD* |
| InterceptFrame | 0.01 | | 0.08 | 0.04 | | 0.20 | 0.01 | | 0.10 | 0.04 | | 0.20 |
| Residual | 0.95 | | 0.98 | 1.96 | | 1.40 | 0.86 | | 0.93 | 1.29 | | 1.14 |
| **Fit Statistics** |  |  | |  |  | |  |  | |  |  | |
| ICC | .01 | | | .02 | | | .01 | | | .03 | | |
| LL | -3,076.72 | | | -1,045.05 | | | -2,975.84 | | | -1,340.59 | | |
| Pseudo-*R*2 | .21 | | | .12 | | | .51 | | | .12 | | |
| *Note*: Cell entries are parameter estimates from multilevel models with random intercepts. Linear models are used for trait-like variables, and quasi-binomial models are used for state-like variables. Data are weighted by education and income. *N* = 2,008. Groups = 17. | | | | | | | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 2  *Conditional Effects of Incidental Exposure on Story Engagement* | | | |  |  | |
|  | Engagement | | High-Effort Engagement | | | | |
| **Fixed Effects** | β | *SE* | β | | | *SE* | |
| Intercept | 2.65\*\*\* | 0.42 | 1.21\*\*\* | | | 0.27 | |
| Incidental Exposure (Trait-Like) | -0.14\*\*\* | 0.06 | -0.10\*\* | | | 0.04 | |
| Incidental Exposure (State-Like) | -2.06\*\*\* | 0.31 | -1.16\*\*\* | | | 0.20 | |
| Involvement (Medium:Low) | 0.11\* | 0.72 | -0.10 | | | 0.20 | |
| Involvement (High:Low) | -0.50 | 0.35 | 0.31 | | | 0.23 | |
| Age | 0.01 | 0.05 | -0.03 | | | 0.03 | |
| Gender (1 = Female) | -0.28\* | 0.12 | -0.20\* | | | 0.08 | |
| Race (1 = Person of Color) | 0.25\*\*\* | 0.13 | 0.07 | | | 0.08 | |
| Education | -0.02 | 0.04 | -0.03 | | | 0.03 | |
| Income | 0.02 | 0.03 | 0.01 | | | 0.02 | |
| Ideology (+ Conservative) | 0.03 | 0.02 | 0.03\* | | | 0.01 | |
| Party Identity (+ Republican) | -0.06 | 0.03 | -0.04\* | | | 0.02 | |
| Frequency of Social Media Use | 0.00 | 0.04 | 0.01 | | | 0.03 | |
| Network Size | 0.19 | 0.17 | 0.27\* | | | 0.11 | |
| Network Diversity | 0.68\*\* | 0.24 | 0.38\* | | | 0.16 | |
| Group Activity | 0.24\* | 0.10 | 0.20\*\* | | | 0.07 | |
| Social News Curation | 0.36\*\*\* | 0.08 | 0.19\*\*\* | | | 0.06 | |
| **Interactions** |  |  |  | | |  | |
| Incidental Exposure (State-Like) x Involvement (Medium:Low) | -0.71\* | 0.34 | 0.30 | | | 0.22 | |
| Incidental Exposure (State-Like) x Involvement (High:Low) | -0.38 | 0.41 | -0.12 | | | 0.27 | |
| **Random Effects** | *Var*. | *SD* | *Var*. | | | *SD* | |
| InterceptFrame | 0.02 | 0.21 | 0.02 | | | 0.13 | |
| Residual | 2.67 | 1.51 | 1.15 | | | 1.07 | |
| **Fit Statistics** |  |  |  | | |  | |
| ICC | .01 | | .02 | | | | |
| LL | -1,613.46 | | -1285.40 | | | | |
| Pseudo-*R*2 | .46 | | 0.47 | | | | |
| *Note*: Cell entries are parameter estimates from a multilevel model with random intercepts. Data are weighted by education and income. Analysis uses subset of respondents who report exposure to story. *N* = 842. Groups = 17. | | | | |  | |

Figure 1

*Differences among involvement groups in news exposure.*

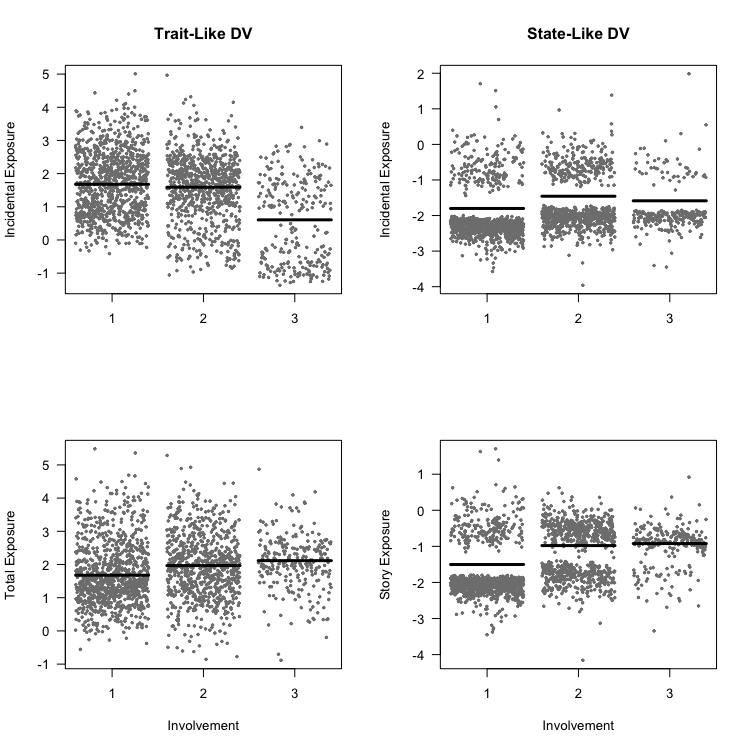
**

Figure 2

*Differences between incidentally and purposefully exposed in news engagement by level of involvement.*

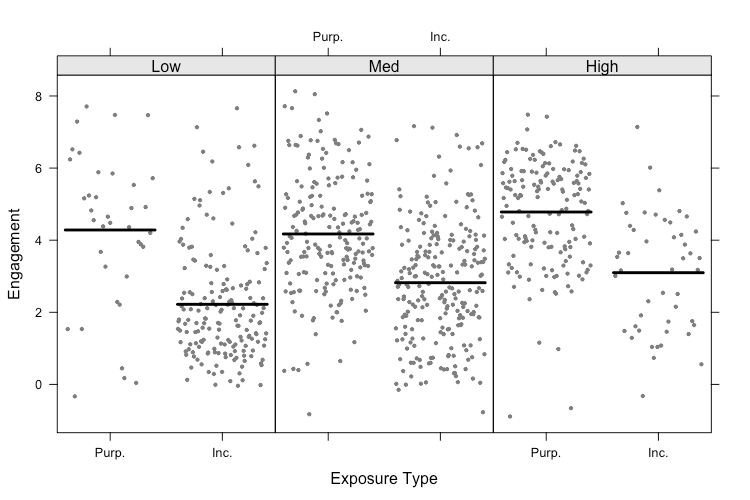
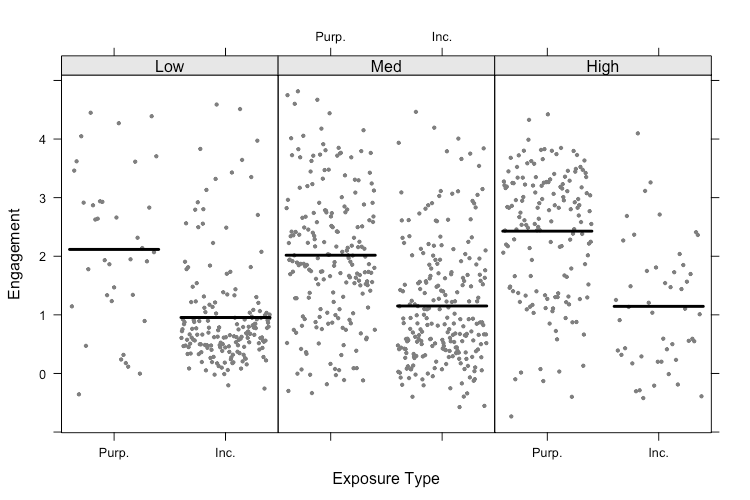


Figure 3

*Differences between incidentally and purposefully exposed in news engagement by level of involvement.*



**Appendix A:**

**Sample Demographics and Weighting Scheme**

|  |  |  |
| --- | --- | --- |
| Table A1  *Demographic Profile of Survey Sample and Target Population* | | |
|  | Current Survey | U.S. Census Bureau:  2016 American Community Survey | |
|  | (%) | (%) | |
| Gender |  |  | |
| Male | 49.0 | 49.2 | |
| Female | 51.0 | 50.8 | |
| Age (median) | 35-44 | 37.7 | |
| Ethnicity/race |  |  | |
| White | 59.6 | 62.0 | |
| Black or African American Native | 15.9 | 12.3 | |
| American Indian and Alaska Native | 1.5 | 0.7 | |
| Asian | 12.9 | 5.2 | |
| Native Hawaiian and other Pacific Islander | 0.2 | 0.2 | |
| Hispanic | 7.6 | 17.3 | |
| Household income (median) | US $60,000–75,000 | US $57,617 | |
| Education |  |  | |
| Less than high school graduate | 2.1 | 13.0 | |
| High school diploma or equivalent | 15.7 | 27.5 | |
| Some college or associate degree | 26.2 | 29.2 | |
| Bachelor’s degree or higher | 56.1 | 30.3 | |
| *Note*: The US Census Bureau 2016 American Community Survey is available online at http://factfinder.census.gov/ | | |

|  |  |
| --- | --- |
| Table A2  *Survey Weights* | |
| Income | |
| Category | Weight |
| Less than $15k | 1.02 |
| $15k to 30k | 1.00 |
| $30k to $45k | 1.00 |
| $45k to 60k | 1.00 |
| $60k to $75k | 1.00 |
| $75k to $100k | 0.86 |
| $100k to $150k | 0.95 |
| More than $150k | 0.95 |
| Education | |
| Category | Weight |
| None, or grades 1-8 | 5.75 |
| High school incomplete (grades 9-11) | 1.77 |
| High school graduate (grade 12 or GED certificate) | 1.33 |
| Some college, no 4-year degree (includes Associate’s Degree) | 0.89 |
| Technical, trade, or vocational school after high school | 0.65 |
| College graduate (Bachelor’s Degree) | 0.42 |
| Post-graduate training/professional school after college | 0.42 |
| *Note*. Income measured as annual household income. Education measured in terms of highest level completed. Final survey weights created by multiplying weights for income and education. | |

**Appendix B:**

**Full Results from Latent Class Analysis**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table B1  *Correlations Among Variables Used in Latent Class Analysis* | | | | |
| Variable | 1. | 2. | 3. | 4. |
| 1. Social Media as News Source (1 = Yes) | 1.00 |  |  |  |
| 2. Self-Reported Interest | .34 | 1.00 |  |  |
| 3. Follow Accounts for News | .52 | .50 | 1.00 |  |
| 4. Algorithmic Categorization (1 = Interested) | .36 | .34 | .46 | 1.00 |
| *Note*: Cell entries are Pearson’s correlation coefficients (*r*). *N* = 2,008 | | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table B2  *Model Fit Statistics for Models with Varying Number of Latent Classes* | | | | |
| Model | AIC | BIC | *G*2 | χ2 |
| 2 Classes | 16,036.27 | 16,153.98 | 398.90 | 409.10 |
| **3 Classes** | **15,810.00** | **15,989.35** | **150.63** | **149.52** |
| 4 Classes | 15,759.35 | 16,000.36 | 77.98 | 82.96 |
| 5 Classes | 15,758.16 | 16,060.83 | 54.79 | 53.95 |
| *Note*: BIC was the primary criterion for model selection. *N* = 2,008. | | | | |

|  |  |  |
| --- | --- | --- |
| Table B3  *Predicted and Observed Latent Class Membership* | | |
| Latent Class | Predicted | Observed | |
| 1: Low Involvement (*n* = 968) | .48 | .48 | |
| 2: Medium Involvement (*n* = 788) | .39 | .38 | |
| 3: High Involvement (*n* = 252) | .13 | .15 | |
| *Note*: Cell entries are predicted probabilities and observed proportions obtained from a latent class analysis (LCA) model. Column totals may not equal 1 due to rounding. *N* = 2,008. | | |

Figure A1

*Group Distributions on Manifest Variables from Latent Class Analysis*

