Internet and Social Media Use in American Adults

Cameron S. Kay[1], Stefania R. Ashby[1], & Ashley L. Miller[1]

[1] University of Oregon

Author Note

Correspondence concerning this article should be addressed to Cameron S. Kay, 1451 Onyx Street, Eugene, OR 97403. E-mail: ckay@uoregon.edu

Abstract

TBD

Keywords: american, social media, internet, pew research center

Internet and Social Media Use in American Adults

Social media sites are commonly defined as an internet-based service allowing for the creation and broadcast of user-generated information (Boyd & Ellison, 2008; Kaplan & Haenlein, 2010; Obar & Wildman, 2015). Obar and Wildman (2015) emphasize the user-generated aspect of this definition, arguing that this content is the lifeblood of social media. Although that may sound hyperbolic, it logically follows that if a site is created with the express purpose of providing user-generated content, it must have user-generated content to function as intended. By way of illustration, without videos created by users, YouTube, a social media site that allows its users to upload and share videos, would fail to serve its primary purpose. Netflix, a site that allows users to only stream videos, does not require user-generated content, as it does not serve user-generated content, and, by extension is not a social media site. Beyond the functional aspects of social media sites, the user-generated focus also highlights the importance of individual differences in the user-service relationship, as users invariably have characteristics that affect how they consume and generate content.

To explore these individual differences in how people consume and generate social media content, as well as how they interact with technology, we conducted secondary data analyses using recent survey data collected by Pew Research Center (2018).

Method

Survey data examining attitudes towards technology and use of technology and social media was collected by Pew Research Center (2018). For the current study, we conducted a secondary data analysis looking at factors that relate to frequency of social media use, as well as factors that relate to overall media consumption.

Participants

Two thousand, two people were surveyed by telephone (75.02% cell phone; 24.98% landline) over a period of 7 days in January of 2018. We excluded any participants who reported that they do not even occasionally use the internet or email ($n = 273$). The resulting sample comprised 1729 people (45.29% female). Ages ranged from 18 to 97 (M age = 48.29; SD age = 17.94)[1]. Concerning race, 68.48% identified as white, 12.78% identified as black, 3.64% identified as Asian, 2.95% identified as mixed race, and 12.15% refused to answer or reported being from some other race.

Materials and Procedure

Research personnel were instructed to read from a script while interviewing each participant. For landline users, the script began as follows:

"Hello, I am _______ calling on behalf of the Pew Research Center. We are conducting a telephone opinion survey about some important issues today and we would like to include your household. May I please speak with the YOUNGEST [RANDOMIZE: (MALE / FEMALE)], age 18 or older, who is now at home? [IF NO MALE/FEMALE, ASK: May I please speak with the YOUNGEST (FEMALE / MALE), age 18 or older, who is now at home?]"

For cell phone users, the script was slightly different:

"Hello, I am _______ calling on behalf of the Pew Research Center. We are conducting a telephone opinion survey about some important issues today. I know I am calling you on a cell phone. If you would like to be reimbursed for your cell phone minutes, we will pay all eligible respondents $5 for participating in this survey. This is NOT a sales call."

Once participants passed initial screening (e.g., 18 years of age or older) and provided verbal consent to the interview, they were asked a series of questions pertaining to internet use (e.g., "How often do you use the internet?"; options ranged from 1 = "Almost Constantly" to 5 = "Less Often"), social media use (e.g., "Do you use any of the following social media sites online or on your cell phone: Twitter, Instagram, Facebook, Snapchat, YouTube, WhatsApp, Interest, and LinkedIn?"), as well as perceptions of social media's

[1]Note that the descriptive statistics for age are slightly lower than reality. Ages 97 and older were recorded as simply 97 in the data.

influence on the self and society (e.g., “When you add up all the advantages and disadvantages of the internet, would you say the internet has mostly been a good thing or a bad thing for society?”; options included: “Good Thing”, “Bad Thing”, “Some of Both”, and “Don’t Know”). Participants were also asked about reading habits (e.g., “During the past 12 months, how many books (print, electronic, and audiobooks) did you read either all or part of the way through?”) and reading format preferences (e.g., “Thinking about all of the books you have read in the past 12 months, were any of those printed books, audiobooks, or E-books?”). The interview ended with a series of demographic questions assessing variables including participant sex, age, race, marrital status, education, current employment status, income, and political affiliation.

Data analysis

We used R (Version 3.5.1; R Core Team, 2018) and the R-packages *bindrcpp* (Version 0.2.2; Müller, 2018), *car* (Version 3.0.0; Fox & Weisberg, 2011; Fox, Weisberg, & Price, 2018), *carData* (Version 3.0.1; Fox et al., 2018), *cowplot* (Version 0.9.2; Wilke, 2018), *dplyr* (Version 0.7.7; Wickham, François, Henry, & Müller, 2018), *emmeans* (Version 1.2.1; Lenth, 2018), *forcats* (Version 0.3.0; Wickham, 2018a), *Formula* (Version 1.2.3; Zeileis & Croissant, 2010), *ggplot2* (Version 3.1.0; Wickham, 2016), *here* (Version 0.1; Müller, 2017), *Hmisc* (Version 4.1.1; Harrell Jr, Charles Dupont, & others., 2018), *jcolors* (Version 0.0.4; Huling, 2018), *lattice* (Version 0.20.35; Sarkar, 2008), *lme4* (Version 1.1.17; Bates, Mächler, Bolker, & Walker, 2015), *lmerTest* (Version 3.0.1; Kuznetsova, Brockhoff, & Christensen, 2017), *lubridate* (Version 1.7.4; Grolemund & Wickham, 2011), *magrittr* (Version 1.5; Bache & Wickham, 2014), *Matrix* (Version 1.2.14; Bates & Maechler, 2018), *pander* (Version 0.6.1; Daróczi & Tsegelskyi, 2018), *papaja* (Version 0.1.0.9842; Aust & Barth, 2018), *plotrix* (Version 3.7.1; J, 2006), *purrr* (Version 0.2.5; Henry & Wickham, 2018), *readr* (Version 1.2.1; Wickham, Hester, & Francois, 2017), *rio* (Version 0.5.10; C.-h. Chan, Chan, Leeper, & Becker, 2018), *sjstats* (Version 0.15.0; Lüdecke, 2018), *stringr* (Version 1.3.1; Wickham, 2018b), *survival* (Version 2.42.3; Terry M. Therneau & Patricia M. Grambsch, 2000), *tibble* (Version 1.4.2; Müller & Wickham, 2018), *tidyr* (Version 0.8.1; Wickham & Henry, 2018), *tidyverse* (Version 1.2.1; Wickham, 2017), and *wesanderson* (Version 0.3.6; Ram & Wickham, 2018) for all our analyses.

Results

First, we sought to examine how frequency of social media use might vary as a function of sex, age, social media site, and the interaction between these variables. Using the *lme4* package by (Bates et al., 2015), we ran a comparison of seven linear mixed-effects models predicting frequency of social media use from age, social media site, sex, and the interactions among these variables. Our first model included only the random effects. Specifically, we suspected that participants may have certain response style in evaluating their frequency of social media use. We entered the subject as a random intercept to account for interdependence in the participants' ratings. Although in many cases the scenario or context (i.e., social media site) would be entered as a random slope, we did not wish to generalize beyond these sites.

Our second model added gender, followed by age in our third model. The fourth model included the social media site in question. The fifth through seventh models added (1) the interaction between sex and age, (2) the interaction between sex and social media site, and (3) the interaction between age and social media site.

As shown in Table 1, we found that the addition of age, social media site, the interaction between sex and social media site, and the interaction between age and social media site all explained a significant proportion of variance in social media site usage frequency. Using the *lmerTest* package from Kuznetsova et al. (2017) to approximate degrees of freedom with Satterthwaite's method, we found that older age is associated with lower social media use, b = -0.01, SE = .00, $p < .003$. Furthermore, Snapchat is used more often than Facebook (b = .73, SE = .18, $p < .001$), and Twitter is used more often than

Facebook (b = -0.74, SE = .18, $p < .001$). However, age interacts with the specific social media site, insofar that increased age results in even lower rates of social media use in the case of Instagram (b = -0.02, SE = .00, $p < .001$), Snapchat (b = -0.04, SE = .00, $p < .001$), Twitter (b = -0.01, SE = .00, $p < .017$), and YouTube (b = -0.02, SE = .00, $p < .001$) when compared to Facebook.

Although the model comparison did not indicate an increase in predictive ability from the introduction of participant gender, we did find that males (in comparison to females) reported less frequency of social media use when using the approximated degrees of freedom, b = -0.40, SE = .14, $p < .005$. When compared to Facebook, however, men exhibitted a greater frequency of using Twitter (b = .49, SE = .12, $p < .001$) and YouTube (b = .61, SE = .09, $p < .001$) than women did.

A comparison of the full model with random effects to the full model without random effects, revealed that the random effects explained a significant proportion of the variance in social media site use, $\chi^2(1) = 176.01$, $p < .001$. In fact, grouping ratings by participant explained 21.65% of the variance in the frequency of using a social media site. A visual inspection of model residuals did not suggest non-normality nor heteroscedasticity. We also quantitively inspected the variables for potential multicollinearity within the fullest model without the interaction effects. No multicollinearity was detected (VIF = 1.01 - 1.04).

We next sought to examine how media consumption may vary as a function of age as well as media format (i.e. individuals exclusively using print media, audiobooks, electronic media or combinations of the three). To explore this data we used the total number of books read in the past year as an indicator of media consumption. Survey respondents who endorsed only using a single type of media format were grouped into their appropriate format bins. Respondents who endorsed using more than one type of media format were grouped together into a "mixed media" format bin. Next, we calculated the average number of books read, grouped by age and media format see Figure 3.

Next, we ran a multiple regression predicting the average number of books read in the past year from age, book format and their interaction. Results showed a significant main effect of age (STATS) with older adults reading more books in 2018 than younger adults.

Discussion

References

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Bache, S. M., & Wickham, H. (2014). *Magrittr: A forward-pipe operator for r.* Retrieved from https://CRAN.R-project.org/package=magrittr

Bates, D., & Maechler, M. (2018). *Matrix: Sparse and dense matrix classes and methods.* Retrieved from https://CRAN.R-project.org/package=Matrix

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Boyd, D. M., & Ellison, N. B. (2008). Social networking sites: Definitions, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*, 210–230.

Chan, C.-h., Chan, G. C., Leeper, T. J., & Becker, J. (2018). *Rio: A swiss-army knife for data file i/o.*

Daróczi, G., & Tsegelskyi, R. (2018). *Pander: An r 'pandoc' writer.* Retrieved from https://CRAN.R-project.org/package=pander

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (Second.). Thousand Oaks CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Fox, J., Weisberg, S., & Price, B. (2018). *CarData: Companion to applied regression data sets.* Retrieved from https://CRAN.R-project.org/package=carData

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, *40*(3), 1–25. Retrieved from http://www.jstatsoft.org/v40/i03/

Harrell Jr, F. E., Charles Dupont, & others. (2018). *Hmisc: Harrell miscellaneous.* Retrieved from https://CRAN.R-project.org/package=Hmisc

Henry, L., & Wickham, H. (2018). *Purrr: Functional programming tools.* Retrieved from https://CRAN.R-project.org/package=purrr

Huling, J. (2018). *Jcolors: Colors palettes for r and 'ggplot2', additional themes for 'ggplot2'.*

Retrieved from https://jaredhuling.github.io/jcolors/

J, L. (2006). Plotrix: A package in the red light district of r. *R-News*, *6*(4), 8–12.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*, 59–68. doi:10.1016/j.bushor.2009.09.003

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. doi:10.18637/jss.v082.i13

Lenth, R. (2018). *Emmeans: Estimated marginal means, aka least-squares means*. Retrieved from https://CRAN.R-project.org/package=emmeans

Lüdecke, D. (2018). *Sjstats: Statistical functions for regression models (version 0.17.2)*. doi:10.5281/zenodo.1284472

Müller, K. (2017). *Here: A simpler way to find your files*. Retrieved from https://CRAN.R-project.org/package=here

Müller, K. (2018). *Bindrcpp: An 'rcpp' interface to active bindings*. Retrieved from https://CRAN.R-project.org/package=bindrcpp

Müller, K., & Wickham, H. (2018). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

Obar, J. A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy*, *39*(9), 745–750.

Pew Research Center. (2018). Core trends survey. Retrieved from http://www.pewinternet.org/dataset/jan-3-10-2018-core-trends-survey/

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ram, K., & Wickham, H. (2018). *Wesanderson: A wes anderson palette generator*.

Retrieved from https://CRAN.R-project.org/package=wesanderson

Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer. Retrieved from http://lmdvr.r-forge.r-project.org

Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival data: Extending the Cox model*. New York: Springer.

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from http://ggplot2.org

Wickham, H. (2017). *Tidyverse: Easily install and load the 'tidyverse'*. Retrieved from https://CRAN.R-project.org/package=tidyverse

Wickham, H. (2018a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2018b). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H., & Henry, L. (2018). *Tidyr: Easily tidy data with 'spread()' and 'gather()' functions*. Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., François, R., Henry, L., & Müller, K. (2018). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., Hester, J., & Francois, R. (2017). *Readr: Read rectangular text data*. Retrieved from https://CRAN.R-project.org/package=readr

Wilke, C. O. (2018). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=cowplot

Zeileis, A., & Croissant, Y. (2010). Extended model formulas in R: Multiple parts and multiple responses. *Journal of Statistical Software*, *34*(1), 1–13. doi:10.18637/jss.v034.i01

Table 1

Log likelihood comparison of linear mixed-effects models predicting social network site use from sex, age, and specific social network site.

	K	χ^2 df	Loglik.	χ^2	p
Null	3		-6857.08		
Sex	4	1	-6856.28	1.61	.205
Age	5	1	-6775.08	162.39	<.001
SNS	9	4	-6620.63	308.91	<.001
Sex X Age	10	1	-6620.27	0.72	.396
Sex X SNS	14	4	-6591.10	58.32	<.001
Age X SNS	18	4	-6532.52	117.18	<.001

Note. Participant $N = 1541$; Social Network Site $N = 5$. Null model includes only the random intercept. SNS = social network site. X indicates an interaction.

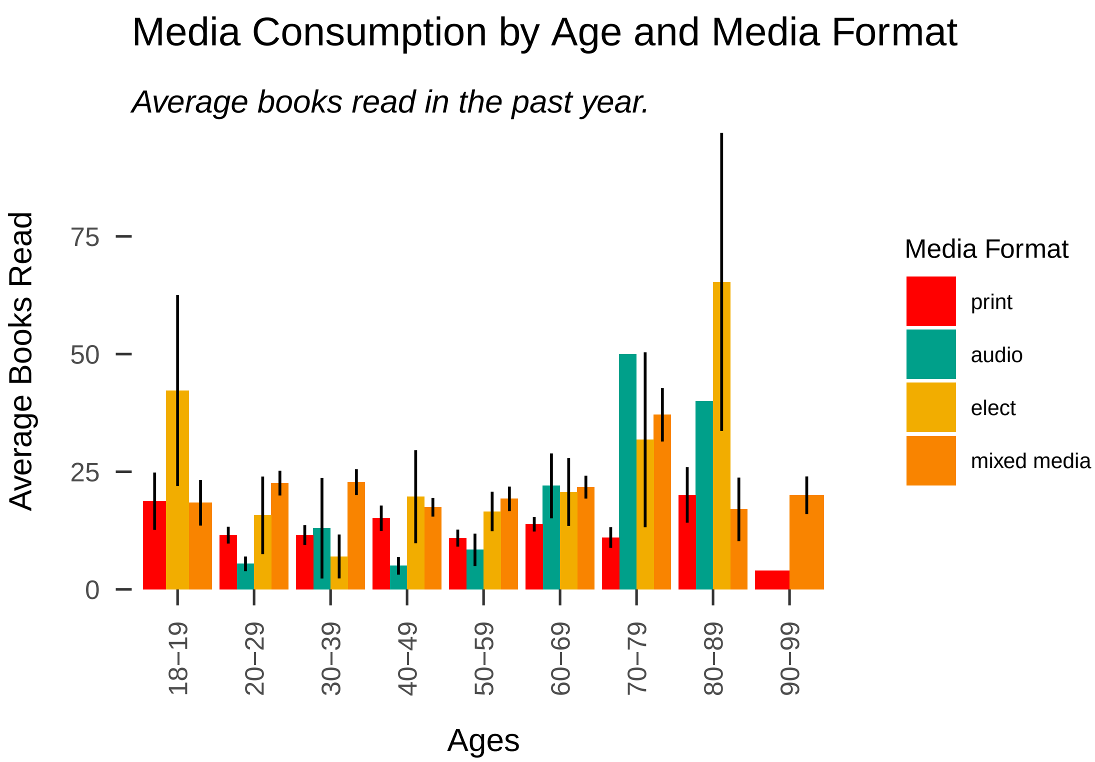

Figure 1. Average books read grouped by age and media format