Political Science Research and Methods (2023), **11**, 110–128
doi:10.1017/psrm.2022.17

ORIGINAL ARTICLE

Digital literacy and online political behavior

Andrew M. Guess[1]* and Kevin Munger[2]

[1]Department of Politics and School of Public and International Affairs, Princeton University, Princeton, NJ, USA and [2]Department of Political Science and Social Data Analytics, Penn State University, State College, PA, USA
*Corresponding author. Email: aguess@princeton.edu

(Received 18 March 2021; revised 29 October 2021; accepted 24 January 2022; first published online 22 April 2022)

Abstract

Digital literacy is receiving increased scholarly attention as a potential explanatory factor in the spread of misinformation and other online pathologies. As a concept, however, it remains surprisingly elusive, with little consensus on definitions or measures. We provide a digital literacy framework for political scientists and test survey items to measure it with an application to online information retrieval tasks. There exists substantial variation in levels of digital literacy in the population, which we show is correlated with age and could confound observed relationships. However, this is obscured by researchers' reliance on online convenience samples that select for people with computer and internet skills. We discuss the implications of these measurement and sample selection considerations for effect heterogeneity in studies of online political behavior. We argue that there is no universally applicable formula for selecting a given non-probability sample or operationalization of the concept of digital literacy; instead, we conclude, researchers should make theoretically informed arguments about how they select both sample and measure.

Keywords: Data collection; digital literacy; information retrieval; survey measurement

1. Introduction

Mark Zuckerberg had plenty to worry about before his blockbuster appearance at the US Senate in April 2018. His company was again the target of intense public backlash after explosive details about Cambridge Analytica's access to user data came to light. More than other controversies before it, this particular scandal brought together all of the public's concerns about internet platforms that had been developing over the previous several years: privacy, manipulation by algorithms, internet companies' scale, subliminal persuasion, and now an added layer of dirty political tricks. Congress had targeted Facebook's CEO with the threat of potential regulation at its disposal and the public's outrage at its back. In the many hours Zuckerberg reportedly spent preparing for his two days of testimony, he likely received coaching on how to explain the logic of advertising auctions and the company's efforts to maintain the "integrity" of its products in the face of disinformation, among other delicate and complex issues not well understood by the public.

Yet for all his diligent preparation, Zuckerberg found himself focusing on the basics. Some of his questioners appeared to confuse rudimentary concepts, such as email and the WhatsApp messaging service. Others invoked conspiracy theories about Facebook secretly listening to its users. In one of his most-discussed answers, Zuckerberg was compelled to explain Facebook's business model to Sen. Orrin Hatch: "Senator, we run ads."[1] Two notable features of these exchanges quickly gained traction on Twitter (a platform with very different features than

[1]See: https://www.theverge.com/2018/4/11/17224184/facebook-mark-zuckerberg-congress-senators and https://www.cnet.com/news/some-senators-in-congress-capitol-hill-just-dont-get-facebook-and-mark-zuckerberg/

© The Author(s), 2022. Published by Cambridge University Press on behalf of the European Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

Facebook, Zuckerberg reassured Sen. Lindsey Graham): They reflected low fluency with basic technological concepts central to politics, and much of the confusion was emanating from the chairs and ranking members of the Senate committees, whose median age was reportedly nearly 80.[2] It would be difficult to conjure a more vivid demonstration of generational gaps in technological savvy and their impact on public policy.

Ironically, research by Facebook itself had already begun turning up findings suggesting that age-based differences in online engagement were relevant for political outcomes. A prominent digital voting experiment, conducted in concert with the platform in 2010, uncovered striking evidence of heterogeneity in the effect of sharing "I Voted" stickers to one's friends (Bond et al., 2012, 2017): "The [Facebook GOTV experiment] effect size for those 50 years of age and older versus that of those ages 18 to 24 is nearly 4 times as large for self-reported voting and nearly 8 times as large for information seeking." Turning to a very different form of networked online political activity, the 2016 US presidential election triggered intense scholarly interest in online misinformation. Using distinct data sets and statistical models, several research teams separately reached two main conclusions. First, the overall consumption of low-quality, untrustworthy, or "fake" news was small. Second, the consumption and sharing of this fake news was very unequally distributed—and much higher among older internet users (Barberá, 2018; Guess et al., 2019; Grinberg et al., 2019; Osmundsen et al., 2020).

The magnitude and consistency of age-based heterogeneities in online political behavior implies some qualitative differences in how people experience digital media. This is suggested by descriptive evidence on both older and younger adults' internet habits. For instance, a review of older adults' internet use patterns finds differences in overall internet skill as well as the kinds of activities they partake in online when compared to other age groups (Hunsaker and Hargittai, 2018). Much more research has focused on younger adults, however. For example, panel surveys have shown significant age-based differences in social media use for political news, with significantly stronger associations with political participation among younger citizens (Holt et al., 2013). Of course, age itself is not likely the driver of these differences. Rather, we posit that age serves as a proxy for some other characteristic that covaries with generational cohort and that plausibly relates to people's experiences with digital media. Literature in sociology and communication has documented the advantages of being a "digital native" as well as the pathologies of the "second digital divide" (Hargittai, 2001, 2010).

Motivated by these concerns, we explore the construct of *digital literacy* and argue for its importance to political science. First, we sift through a diffuse multi-disciplinary literature and propose a conceptualization of digital literacy as *online information discernment combined with the basic digital skills necessary to attain it*. Given the concept's multifaceted nature, we focus primarily on this "skills" dimension, which captures the kind of basic technological fluency commonly associated with effective use of online tools for political purposes. We then make two key methodological contributions for scholars interested in incorporating the concept into their research. First, adapting and expanding on scholarship in other disciplines, we assess prominent survey batteries proposed to measure skills. Second, we draw attention to a subtle but consequential form of selection bias that has implications for researchers' sampling decisions. We illustrate the importance of these factors in explaining a behavior likely to be related to digital skills: online information retrieval.

We argue that the choice of subject pool for surveys is critically important for questions related to digital literacy; specifically, fielding studies on Mechanical Turk (MTurk) is likely to drastically undersample low-skill respondents, which could skew inferences in unexpected ways. Returning to one of the motivations for our investigation, we document with a large national sample that skills do indeed covary with age, with older Americans having lower levels on average. This hints at the possibility that unmeasured variation in digital literacy has driven prior findings

[2]See: https://twitter.com/iankoski/status/983780337829449728

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

on the apparent relationship between age and online behaviors such as sharing low-quality (or "fake") content on social media. Finally, we illustrate these issues with an analysis of online information retrieval accuracy. We conclude with a discussion of how the concept of digital literacy can usefully contribute to several other literatures likely impacted by recent shifts in the online information environment.

2. Digital literacy and media modality: a brief history

The rapid adoption of the internet and social media in the United States has changed the media landscape. But by most ways of counting, the spread of first radio and then television through society happened more rapidly, and each had a larger impact on the daily media diets of the average consumer (Prior, 2007). This was only possible because radio and television are *easy* to consume: everyone is trained to listen to human speech from birth, or to listen to speech paired with a visual stimulus. Other than language barriers or the rare niche program, anyone could turn on their radio set or television and share an experience with anyone else listening.[3]

The internet is different in every respect. It took many years to develop and has been constantly mutating. During its formative years, the only consumers tended to also be producers. Only people with specialized skills could even access online media, which was at first only written text. This early online media was thus produced and consumed by a small number of professionals in academia and tech companies, gradually expanding to geeky hobbyists with enough free time and disposable income to purchase and use the necessary, complicated hardware (Abbate, 2000; Burke et al., 2005).

For the past 50 years, television has been the dominant modality for political media. The total supply of television content is constrained by the cost of production and distribution, and televisual content requires minimal skill to appreciate. So, with a small number of notable exceptions,[4] the dominant theories of media effects were focused on *homogeneity*.

The internet inverts this. Central tendencies are simply not that informative. Individuals' experiences are so distinct that heterogeneity should arguably be the baseline expectation.

The technological affordances of online media allow for a much greater variety of content, expanding the scope of "politics" beyond the evening news or cable talk show. Furthermore, this variety has increased over time, as increasingly different types of people and organizations produce that content. But this is only one source of increased heterogeneity of media effects due to the expanded diversity of the audience for online media. Unlike a radio or television broadcast, where the range of the experiences among adults exposed to a given piece of content is limited, the range of the experiences among adults exposed to online media is *extremely wide*, at least for the internet audience of the late 2010s.

These *experiences* are created at the intersection of a media consumer and a piece of online media. The classic model of political sophistication in Luskin (1990) conceives of the acquisition of political information as a process with three inputs: access to that information, the motivation to acquire it, and the ability to process it. Access has clearly increased with the web, and Sood and Lelkes (2018) claim that rising education levels suggest increased ability, at least in the aggregate. The key complication, we argue, is that the *difficulty* of acquiring (true) political information on social media is much higher than in other modalities—and that *ability* is currently highly heterogeneous.

[3]This audience-capacity line of thinking is inspired by Bennett and Iyengar (2008)'s magisterial retrospective on the history of communication scholarship, which argues that the analytic perspectives that scholars find compelling at a given time are driven by their technological and social framework. Importantly, communication theorists of the early 20th century were grappling with the novelty of mass society made possible by these broadcast technologies.

[4]In addition to the widespread theory that media effects are heterogeneous in partisanship, Prior (2007) and Arceneaux and Johnson (2013) establish the importance of audience preferences for entertainment limiting the total reach of political media. Mutz (2015) makes a similar case for studying the heterogeneity in audiences' conflict avoidance in understanding the reach of uncivil cable talk shows.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

Our aim is in part to translate a large literature on this internet *ability* from other disciplines into the realm of political science. This literature begins with Hargittai (2001), who describes the "Second-Level Digital Divide." To that point, concern about the "Digital Divide" was over differential access to the internet. Hargittai's method was participant observation: she sat behind people and watched them use the internet to search for information.

This simple act was revelatory—and is a tool that could still be fruitfully used by researchers to study novel internet technologies. People used the internet *very differently*. They varied considerably in their capacity to find factual information, and in their speed in doing so. Hargittai conceptualized this as "online skills"—a concept that refers explicitly to the actual, validated ability "to efficiently and effectively find information on the Web," or, put another way, to be digitally *literate*.

3. Digital literacy: theoretical building blocks

The term *digital literacy* (or *digital media literacy*) is frequently invoked but rarely defined consistently.[5] Possibly for that reason, its boundary is porous: discussions of digital literacy tend to overlap with related concepts such as the aforementioned internet skill (Hargittai, 2005), media literacy (Vraga and Tully, 2015; Vraga et al., 2015), and digital inequality. Unsurprisingly, then, research on the topic is dispersed across multiple disciplines ranging from sociology and communication (e.g., Koltay, 2011) to library and information sciences (e.g., Eshet, 2004). Building on existing work, a contribution of this paper is to theoretically situate the concept of digital literacy and to embrace its multidimensional character in a way that translates to straightforward measures suitable for use in political science.

We begin by postulating that being digitally literate means being able to reliably assess the credibility of information encountered online (e.g., Flanagin and Metzger, 2007). This in turn depends on the ability to verify claims and look up answers to questions using a variety of strategies.[6] While this may seem similar to traditional notions of information fluency (Sharkey, 2013), it is situated within people's digital environments and the social context they experience online. In other words, we do not conceive of digital literacy as something specifically related to formal information sources (news organizations) or topics of continued interest to scholars (politics, health). A key feature of multifaceted environments is the constant need to assess the credibility of claims and requests, both formal *and* informal—not only while perusing social feeds and news headlines, but also in everyday encounters like phishing attempts and spam in one's email inbox. Whether routine or connected to specific informational needs, all of these tasks require judging what's on the screen in front of them: separating the trustworthy (genuine email communication, professional-quality news, credible claims by friends and acquaintances) from everything else.

Becoming proficient at these judgments requires experience as well as the fundamental skills needed to enable it. In the realm of routine tasks, recognizing a phishing attempt requires familiarity with the mechanics of email. Checking the credibility of a source might require opening a new browser tab (on a desktop or laptop computer) or switching apps (on a mobile or tablet device)—building blocks so essential that they are often bundled alongside higher-level competencies when taught in libraries and schools.

At its core, then, digital literacy consists of a *skills* component and an *information literacy* component. This is implied in some existing scholarship. In a foundational book on the topic, Gilster (1997) asserts that "digital literacy is about mastering ideas, not keystrokes" (15), suggesting the primacy of a more general kind of information literacy. But in the same book, Gilster

[5]"The indistinct use of the term causes ambiguity, and leads to misunderstandings, misconceptions, and poor communication among researchers" (Eshet, 2004); Bawden (2008) describes it more bluntly as "a topic whose terminology is very confused" (17).

[6]Such strategies for critically evaluating online information include "lateral reading" (Wineburg and McGrew, 2018) and the "SIFT" technique (https://hapgood.us/2019/06/19/sift-the-four-moves/).

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

defines core competencies of digital literacy which include the ability to search the internet and navigate hypertext. Taking the link between skills and information literacy to its logical conclusion, one could argue that literacy is itself a kind of skill. Hargittai and Micheli (2019) identify ten dimensions of "internet skills," such as interpersonal communication and managing privacy, which collectively amount to a digital citizenship test for the 21st century. One of Hargittai and Micheli's skills dimensions is the "ability to find and evaluate information," which corresponds most directly to notions of information literacy or discernment.

This twin conception of digital literacy—information discernment combined with the basic digital skills necessary to attain it—has two advantages. It unifies disparate strands of scholarship while establishing boundaries that separate it from related ideas such as digital savvy and news literacy. At the same time, it allows for tractable measurement: Instead of attempting to detect information literacy, we focus on the skills component as a useful proxy for the whole.

The question of how to define and measure digital literacy returns us again to political sophistication (Luskin, 1990), which similarly bundles several components (i.e., knowledge, interest, and abstract reasoning ability). Reviewing the many existing approaches to measuring the concept, Luskin (1987) proposed a type of "recognition and understanding" measure (Converse, 1964) that grades respondents on a series of factual questions about the ideological leanings of the parties. This measure intentionally taps the knowledge element of sophistication ("what") rather than assessing directly people's belief systems ("how") or the extent to which they think about politics at all ("how much"). As Luskin argues, "Empirically ... there should be a strong relationship between the 'how much' and 'how' and some factual aspects of the 'what'" (881)—in other words, while it may be theoretically possible to have a sophisticated political belief system built on basic factual misconceptions, it is difficult to envision in practice. Likewise, someone who lacks the digital acumen to download and send a file may be well versed at searching for reliable information online, but the combination would be unusual.[7]

As political sophistication came to be an important moderator in studies of retrospective voting and preference formation, we argue that digital literacy is a crucial factor in online political behavior whose role has to date been obscured by disciplinary practices designed for an earlier media-technological environment.

But how should digital literacy—in particular, the skills component we take as a tractable proxy for the whole—be measured? We argue that there is not a "one-size-fits-all" measurement strategy for digital literacy. There are, instead, digital *literacies*, defined with respect to the specific technological context of the user and the task they wish to complete. Reading English text has been a generally stable task for at least a century, and literacy has been similarly stable and simple to conceptualize and measure. Digital literacy, by contrast, can be understood by reference to specific subdomains, with a relevant knowledge base that is not necessarily fixed over time. We contend that the conception of digital literacy outlined here is flexible enough to encompass multiple skills and competencies (Hargittai and Micheli, 2019), and as such, the optimal survey battery may depend on the nature of the research question. For example, does the question concern people at the extremes of digital literacy? Is temporal validity a concern (Munger, 2019)? Are researchers interested in populations that primarily use mobile devices?

4. Data and measures

Our investigation focuses on three measures capturing distinct dimensions of the construct, in addition to an information retrieval exercise that we consider both a validation exercise and a demonstration of a behavioral approach to measuring digital literacy. The analyses we present span five different samples. We introduce our measures and data sources in turn below.[8]

[7]An interesting corollary is that the skills component lends itself to direct behavioral observation without the need for self-reported survey responses. However, as we discuss below, translating that to usable measures is a nontrivial challenge.

[8]Versions of all scales used for this study can be found in online Appendix A.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

4.1 Survey instruments

Internet skills scale. The first measure that we examine is Hargittai's 21-question "internet skills" battery asking respondents to self-report their familiarity with a series of computer- and internet-related terms, on a five-point scale (Hargittai, 2009). This is an updated version of a survey instrument validated against Hargittai's original qualitative skills assessments as a more scalable proxy measure. The items are designed to reflect terms central to everyday internet use, such as "app," "hashtag," "wiki," and "selfie," alongside common but somewhat more technical jargon such as "spyware" and "phishing" that are useful for recognizing online threats. This approach makes sense as a broad conception of the skills component of digital literacy and has the additional benefit of flexibility, since the items can be adapted to include recent coinages as well as more domain-specific terms.

Power user scale. The second is the "power user" scale developed by Sundar and Marathe (2010). This scale consists of 12 questions asking respondents how they interact with technology, on a nine-point agree-disagree scale. Examples include "I make good use of most of the features available in any technological device" and "Using information technology makes it easier to do my work." We speculate that the lack of proper nouns in this scale could improve "temporal validity" (Munger, 2019) over alternatives. This scale is most useful for research questions with specific predictions about individuals with especially high technological competencies. Given the self-reported nature of these questions, they also offer a subjective dimension that could be useful for theories related to occupational identity or (over)confidence.

Low end scale. We designed an additional scale: the "low end" scale. This scale adapts the intuition behind the "power user" scale, which was designed to separate high-skill users from everyone else, while ours is designed to separate low-skill users from everyone else. Some research questions may be less concerned with overall variation in digital literacy than with individuals with the lowest levels of digital literacy; the scale was developed with these in mind. Examples include the literature on online misinformation consumption and sharing, and related concepts such as susceptibility to online scams (Mathur et al., 2020).

Information retrieval tasks. We include in our survey instrument a series of three information retrieval tasks of varying degrees of difficulty. These are online information-related tasks of interest to political scientists; to our knowledge, this is the first internet-based survey implementation of these types of tasks. Internet-based surveys can be a drawback for scholars studying political knowledge because they enable the possibility that respondents will "cheat" and look up the answers online (Clifford and Jerit, 2016; Smith et al., 2020). We use this feature to our advantage, explicitly instructing respondents to look up the answers and intentionally asking questions which are sufficiently obscure to US-based respondents that few could be expected to know the answer.

The three information retrieval questions were:

- Who is the Prime Minister of Croatia? (Andrej Plenković; the presence of the concluding diacritical means they copy+pasted the answer)
- What is the capital city of Malawi? (Lilongwe)
- What is the only US National Park that begins with the letter "T"? (Theodore Roosevelt National Park)

4.2 Data

To ensure that our analysis was not skewed by the particularities of any given sample and to ensure face validity, we conducted the survey on five distinct samples. Our selection of samples

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

is consistent with the principle of "fit for purpose," which has been used in pharmaceutical sciences as a criterion guiding validation efforts (Lee et al., 2006).[9] Put simply, this approach advocates that "the scientific rationale ... should guide the validation requirements" (Ray, 2016). A fit-for-purpose approach in our case suggests choices for both sampling and measurement. For the former, we are interested in sampling from populations whose online behavior is of substantive interest: sharing and otherwise engaging with politics on social media is an increasingly important form of political participation (e.g., Bode, 2014; Bond et al., 2017; Forestal, 2017; Settle, 2018). Moreover, the ways in which these behaviors covary with digital literacies necessitate samples with sufficient variation in this characteristic in order to gauge how accurately measures capture it. Thus, more traditional (probability) samples may either miss or vastly undersample populations of interest to the researcher. In terms of measurement, fit for purpose implies validation for specific applications—such as information retrieval—as a competency to be predicted or explained. In the terminology of Kruizinga (2020), our investigation covers several stages of the fit-for-purpose validation cycle: For example, we conduct *technical validation* by benchmarking our several measures against a "gold standard": purposive, non-probability samples with *a priori* assumed skill levels. We also perform *clinical validation* by testing correlations with other known measures: concrete tasks with known correct answers. This indicates that our candidate measures "show promise" in specific populations of interest.

Mechanical Turk (MTurk). We recruited a sample from Mechanical Turk ($N = 503$), a now-ubiquitous source of subjects for many types of social science experiments (Berinsky et al., 2012). The MTurk interface is somewhat difficult to navigate, and qualitative evidence from Brewer et al. (2016) suggests that this barrier prevents low-skill people from using the platform. As a result, we expect this sample to skew to the high end of digital literacy and to contain a hard floor, below which there will be virtually no respondents.

Facebook ads. We follow Zhang et al. (2018) and Munger et al. (2020) and use a Facebook ad campaign to recruit survey participants ($N = 451$). Facebook is generally very accessible, and many users have accounts created for them by younger, more tech-savvy relatives to make it easier to keep in touch. Redmiles et al. (2018) report that certain groups of Facebook users are much more likely to click on spam, in particular women and lower-skill users. Munger et al. (2020) report that similar groups are more likely to click on these Facebook recruitment ads. As a result, we expect this sample to skew toward the lower end of digital literacy relative to other samples.

High-skill targeted sample. Our "high-skill" sample ($N = 83$) was constructed by asking friends and colleagues who work at technology companies to distribute the survey among their workforce. Navigating computer technology is central to these people's careers, and our assumption is thus that they are highly skilled.

Low-skill targeted sample. Our "low-skill" sample ($N = 18$) is comprised of people who were taking introductory computer skills classes at the Princeton Public Library and the Brooklyn Public Library. Our assumption here—corroborated by conversations with librarians at these institutions—is that people who were taking the time to learn the basics of computing are lacking in those skills.[10] At a minimum, this sample illustrates that populations of low-digital-literacy

[9]We thank a reviewer for suggesting this point.

[10]Without careening into an epistemic abyss, we consider the way this population is treated by researchers. At any point over the past three decades, the population of internet users is non-representative of the US population; researchers studying online behavior plausibly defend the former population as worthy of study on its own terms. But growth in access to the internet during that time period also means that the population of internet users at time t is not representative of the population of internet users at time $t+1$. This is all to say that there is no defensible infimum of computer skills we can use to define the population of internet users today.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

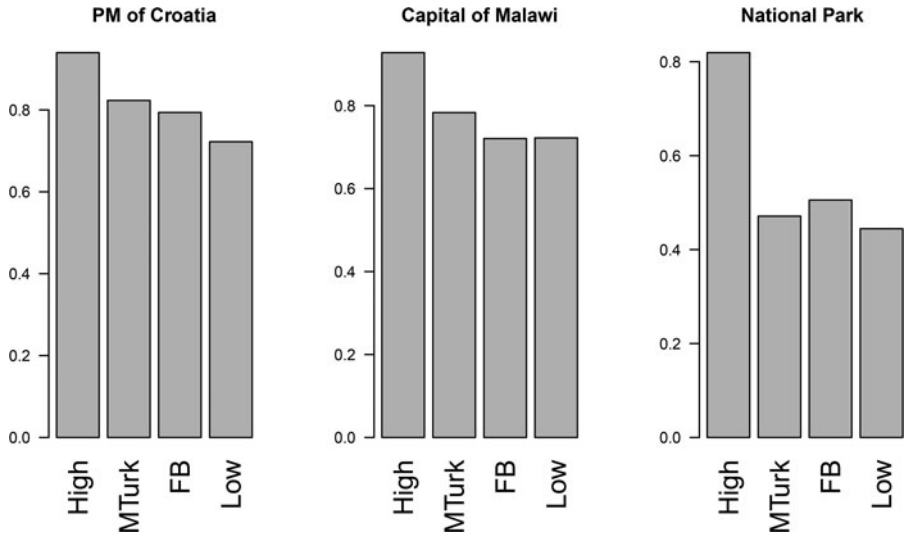

Figure 1. Information retrieval accuracy across four samples, percentage of respondents from each sample (Facebook sample, $N = 451$; high DL sample, $N = 83$; low DL sample, $N = 18$; MTurk sample, $N = 503$) who correctly answer each of three information retrieval questions.

respondents do exist, that they have intentions to go online, and that their characteristics differ in important ways from typical samples used by political-science researchers. Given our extremely limited sample of these respondents, we emphasize that our analysis of their responses serves merely as an existence proof; we encourage more research on these difficult-to-sample populations.

Lucid. Finally, we recruited a national sample from Lucid ($N = 2, 146$), an online survey platform recently attracting social scientists' attention (Coppock and McClellan, 2019). We expected to find individuals in this sample that span a wide range of digital literacies, although it may still somewhat under-represent people at the lowest end.[11]

5. Distributions of digital literacy

We begin by investigating the distribution of the three distinct measures, each theoretically related to an ideal-type task in the domain of digital media: retrieving factual information. The distributions of the correct answers to the three questions in the retrieval exercise across our first four samples comports with our expectations. On average, the tech-worker "high" sample does the best, and the computer-class "low" sample does the worst, as can be seen in Figure 1.

Focusing on each task separately, it is evident that the first two questions are both similarly straightforward. A single Google search of the exact question will produce one of Google's auto-populated "answer boxes" with the correct answer. On these questions, our expectations of the relative performance were correct: the MTurk sample did better than the Facebook sample, and they were between the high and low samples.

The third question was selected specifically to require slightly more skill at crafting a search term and navigating the options provided. Here, the only gap is between the "high" sample

[11]We collected the data in late 2019, before a documented decline in attentiveness began to be evident in Lucid samples (Aronow et al., 2020).

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

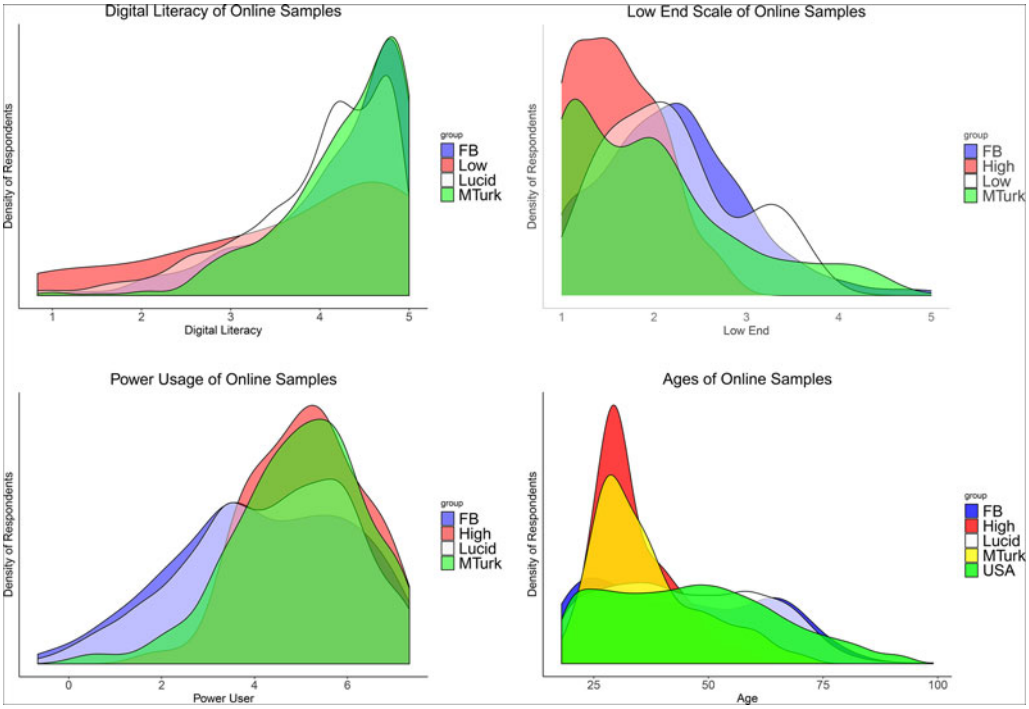

Figure 2. Distributions of relevant measures across the samples, each plot represents the distribution of a given variable across four of five different samples (Facebook sample, $N = 451$; high DL sample, $N = 83$; low DL sample, $N = 18$; MTurk sample, $N = 503$; Lucid sample, $N = 2146$). The first plot is the distribution of skills of respondents, as measured by the 21 Hargittai identification questions (shortened seven-question scale used for Lucid); the high DL sample is excluded because it is so far skewed to the right that the graph is unreadable. The second plot is the distribution of the "low end" measure according to our novel five-question scale, which we did not ask the Lucid respondents. The third plot is the distribution of the "power user" measure according to the 12-question power user scale; the low DL sample is excluded because this measure is not designed to distinguish between people on the lower end. The final plot is the distribution of ages of respondents; the low DL sample is excluded because very few of those respondents answered the open-ended prompt with a number; "USA" refers to the population distribution from the 2010 Census.

and everyone else. This suggests that lower-skill individuals, of which there should have been some across the other samples, had difficulty with this question.

This exercise demonstrates the existence of large gaps in the apparent ability of individuals across different populations to find the answers to straightforward questions using standard online tools. We have outlined reasons why digital literacy would not be evenly distributed across these samples, and how this moderator evidently captures a set of skills that are highly relevant for the particular task. Next, we delve further into this result by exploring how the survey batteries that we have introduced can help us measure variation in digital literacy *within* samples.

Armed with our suite of measures, we describe the distribution of digital literac(ies) within our main samples. Figure 2 plots these distributions, along with either the low- or high-DL samples discussed above, depending on the measure.

The bottom right panel of Figure 2 presents the distribution of the demographic that first motivated this investigation into digital literacy: age. The results are stark. The age distributions for the MTurk and high-DL samples are similar to each other and skewed far in the direction of the young, relative to the data from the 2010 US Census (in green, labeled "USA"). In contrast, the age distribution of the Facebook sample is actually skewed slightly older, with people in their 60s

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

over-represented. The quota sample from Lucid matches the Census numbers closely, suggesting that this sampling was well-conducted.

The top left panel of Figure 2 displays the results for the internet skills measure. The full 21-question scale was used for most of the samples, while a shortened seven-question scale was used for Lucid. As expected, the low-DL sample has a much flatter distribution, with some respondents who report not being familiar with nearly any of the internet terms. The MTurk sample is the least diverse in terms of digital literacy, with close to a hard floor at 2.5 on the five-point scale. Only 1.4 percent of MTurk respondents were below this level, compared to 16.7 percent of the low-DL sample, 5.1 percent of Facebook respondents, 5.8 percent of Lucid respondents, and 0 out of 83 respondents from the high-DL sample.[12]

The top right panel shows our novel low end scale, designed to detect more variation among people with lower levels of digital literacy. The modal high-DL and MTurk respondents are at the bottom edge of this scale, although the support for the MTurk distribution does span the entire range. Facebook respondents are more normally distributed, while (as expected) the low-DL sample provides the highest proportion of respondents who score high on the low end scale.

Finally, the bottom left displays the corresponding distributions for the 12-question power user scale. This is in many ways the most troubling result for Mechanical Turk: the distribution of this measure for our high-DL sample of tech company employees is nearly identical to the distribution for MTurkers. In contrast, the measure is much more broadly distributed across the Facebook and Lucid samples, with a bump on the higher end in the latter.

5.1 Within-sample correlations

Figure 2 suggests a potentially serious methodological problem. The nationally representative studies of misinformation on social media during the 2016 election (Guess et al., 2019; Grinberg et al., 2019) demonstrated a correlation between age and sharing "fake news." We have suggested that this finding is due to the fact that the age effect is largely—but not entirely—capturing the effect of digital literacy. We demonstrate this relationship in Figure 3, which provides some logic for why the power user scale performed so well, as well as evidence for the crucial weakness of using MTurk samples for studying online behaviors potentially moderated by digital literacy.

The top left panel of the first subfigure is evidence of the inverse relationship between age and internet skills in the Facebook sample. The loess line of best fit demonstrates the non-linearity of this relationship; the line is flat (even sloping slightly upward at the low end) throughout the 18–50 age range, when it becomes significantly negative.

A similar relationship is found in the Lucid sample, portrayed in the bottom right corner of the first subfigure. In this sample, there is a negative, linear relationship between age and internet skills throughout the sample. This represents the best evidence to date of a negative relationship between age and internet skills in the United States today. The sample ($N = 2146$) matches national distributions in age, gender, ethnicity, and region demographics.

However, the top right plot of the first subfigure, using the MTurk sample, demonstrates *no correlation* between age and internet skills. The average skill level of the 20-year-olds and 70-year-olds recruited via this platform is the same. A similar non-relationship is observed in the high-DL sample in the bottom left plot, where (in accordance with our purposive sampling) a majority of respondents scored at the maximum of the scale.[13]

[12]The point of insufficient variation at the low end of digital literacy among MTurk workers was recently noted as well by Hargittai and Shaw (2020).

[13]We cannot perform the same analysis on the low-DL scale because too few respondents provided us with their age. Only seven provided a numeric response in the open-ended text field; several of the other respondents provided responses such as (paraphrased for respondent privacy) "Too old for nosy questions" and "Old enough."

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

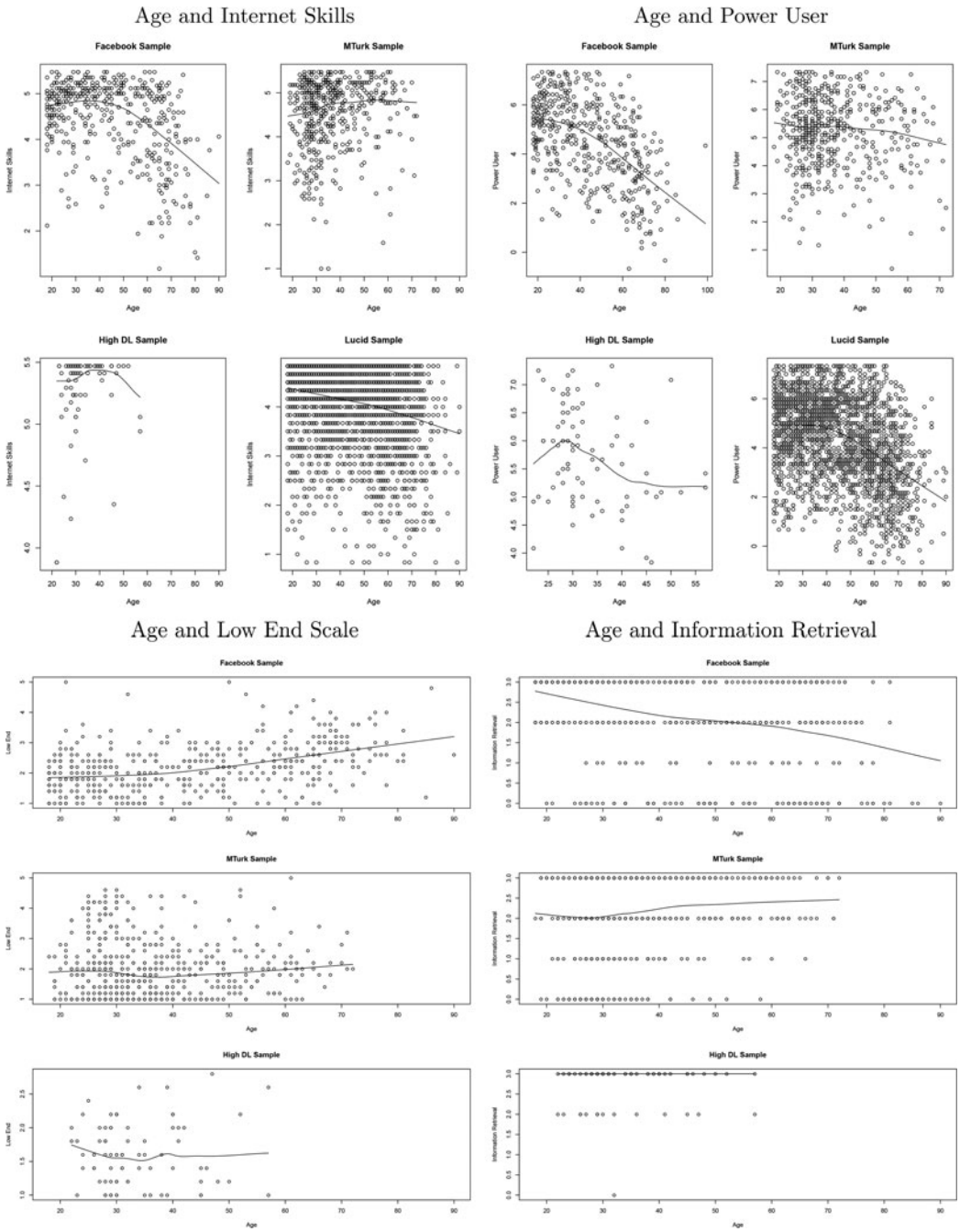

Figure 3. Each scatterplot and LOESS curve represents the relationship between age and either internet skills (top left); power user (top right); low end (bottom left); or information retrieval (from 0 to 3, the number of successful internet searches for information, bottom right) in a given sample (Facebook sample, $N = 443$; high-DL sample, $N = 83$; MTurk sample, $N = 503$; Lucid sample, $N = 2,\ 146$). The low-DL sample is not present because too few respondents offered a numerical age. Age and internet skills, age and power user, age and low end scale, age and information retrieval.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

Table 1. Information retrieval and age/digital literacy across samples

	Facebook sample				MTurk sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skills	0.548***	0.492***	0.484***	0.441***	0.462***	0.360***	0.439***	0.396***
	(0.073)	(0.080)	(0.087)	(0.086)	(0.068)	(0.079)	(0.083)	(0.085)
Low end		−0.118*	−0.113	−0.081		−0.133**	−0.151***	−0.155***
		(0.072)	(0.075)	(0.075)		(0.053)	(0.054)	(0.053)
Power user			0.008	−0.037			−0.092***	−0.072**
			(0.033)	(0.034)			(0.035)	(0.035)
Age				−0.013***				0.009**
				(0.003)				(0.004)
Constant	0.033	0.498	0.483	1.349***	0.377	1.016***	1.219***	0.940**
	(0.267)	(0.388)	(0.393)	(0.431)	(0.253)	(0.359)	(0.365)	(0.379)
N	451	451	451	442	504	504	504	504
Adj. R^2	0.111	0.114	0.113	0.155	0.083	0.093	0.104	0.113

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

The trends are similar in the second subfigure, age and power user. Again, there is a negative relationship between age and power user in the Facebook and Lucid samples, and while the relationship is negative rather than flat in the MTurk sample, the relationship is weak. The big difference is in the high-DL sample, in the bottom left panel. Unlike for internet skills, there is no ceiling effect for power user, and we see a non-linearity in this sample, peaking at around age 30. This has some face validity, but we note that the sample size is not large enough to allow for high confidence in this trend.

The third subfigure plots the relationship between age and low end. The trends are the same as for internet skills, but inverted: there is a positive relationship between age and low end on Facebook, but no such relationship in the high-DL or MTurk samples.

The fourth subfigure of Figure 3 presents the bivariate relationship between age and information retrieval in the various samples. There is a strongly negative relationship between age and information retrieval in the Facebook sample, although this time the slope is close to constant and negative across the age range. By contrast, there is no evidence of a negative relationship between age and information retrieval in the MTurk sample—in fact, there is a weak but *positive* relationship among people 30 and older.

Presenting this another way, Table I displays a series of regressions with our index of three information retrieval questions as the dependent variable. The first four columns use the Facebook sample; the next four, the MTurk sample. In column 4 (Facebook), we see the predicted negative relationship between age and information retrieval; we also see the expected negative relationship with the low end scale and the positive relationship with the internet skills scale.[14] However, the eighth column replicates the analysis on the MTurk sample and indicates strong evidence of the *opposite* relationship: age here is positively correlated with information retrieval, with nearly identical magnitude and standard errors. The adjusted R^2 also differs across the two samples. Because there is more variation in the digital literacy measures and age in the Facebook sample than in the MTurk sample, these variables explain more of the variance in the former than in the latter.

This result can be described as an example of conditioning on a collider, a form of sample selection bias (Pearl, 2009). Equivalently, this is a problem of "selecting on the dependent variable" of successfully navigating relevant internet spaces. The issue parallels the canonical

[14]An important nuance to consider beyond these general relationships is the extent to which "overconfidence" and "underconfidence"—i.e., mismatches between self-reported skill or familiarity and actual information retrieval performance—are predicted by individual or sample characteristics. We investigate this, and related concerns about differential motivation, in online Appendix B.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

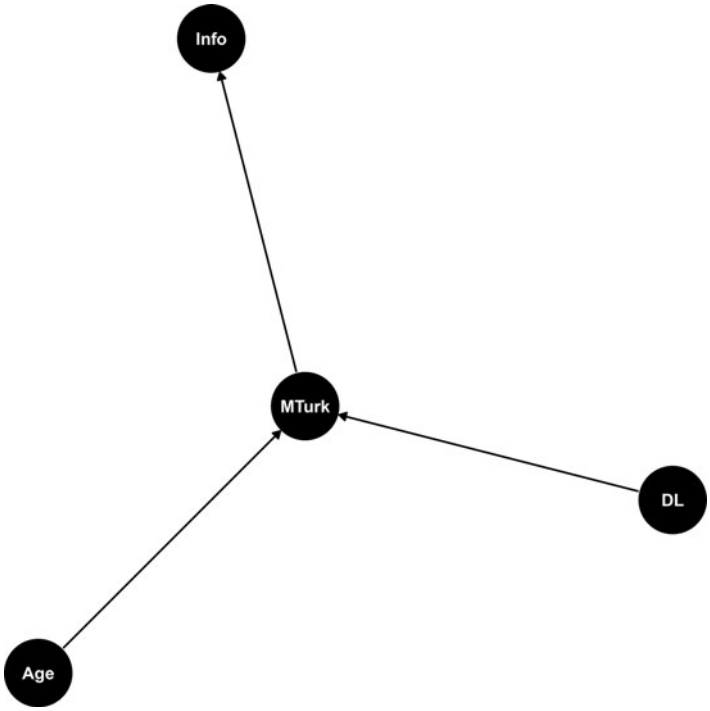

Figure 4. Directed acyclic graph of selection process.

examples of a non-relationship between SAT scores and GPA in selective colleges, or the non-relationship between height and performance among NBA players. In each case, individuals who score high on the former measure are more likely to be included in the sample; individuals who are included in the sample despite scoring low on the former measure are likely to possess other traits that contribute to their performance within the sample.

We would like to be able to estimate the effect of age and digital literacy on information retrieval. However, to do so we need to sample from the general population. In the present case, age and digital literacy are causally related to being an MTurk worker. Lower-digital-literacy internet users are less likely to be aware of the platform, and less likely to be able to navigate its non-intuitive interface.[15] Figure 4 visualizes the problem using Pearl's graphical framework. This selection process means that within the MTurk sample (where the "collision" between age and digital literacy can be observed), the causal relationship between age and information retrieval is broken.

As a result, experimental studies of digital media effects or online misinformation conducted on samples recruited from MTurk are unlikely to accurately estimate the role of age as a treatment effect moderator. Furthermore, as the first panel of Figure 2 demonstrates, there are close to zero MTurkers who are below a threshold level of the internet skills or power user scales. Given the highly unequal distribution of engagement and sharing of "fake news" during the 2016 election, concentrated among those age 65 or older, it is likely that the most important respondents (who would be influential observations in any statistical analysis) are precisely those who are structurally excluded from MTurk.

[15]Brewer et al. (2016) perform a qualitative analysis on a small sample of older adults who have not used MTurk, signing them up for the platform and interviewing them about the experience. Although these subjects generally reported being comfortable using computers, they could not complete basic tasks on MTurk.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

6. Literacies and heterogeneity in political science

In this paper we argue that taking digital media seriously requires moving away from the assumption of media effect homogeneity developed in the broadcast era. Specifically, important heterogeneity in an exemplary digital media task can be explained with multiple measures related to the construct of digital literacy. However, we find significant heterogeneity left to explain—and that age remains a powerful factor.

We demonstrate that digital literacy varies considerably among populations frequently used by political scientists. In particular, we show that samples from MTurk contain vanishingly few low-digital-literacy respondents—precisely the population among whom we might expect to find the largest persuasive effects of digital media messages (including misinformation). We recommend that MTurk not be used for studies of digital media effects in the absence of strong theoretical reasons to expect effect homogeneity.

However, our results should not be construed to speak only to the topic of MTurk and digital literacy. We emphasize our previous discussion about digital *literacies*; there is a parallel logic regarding the role that theoretical expectations play in evaluating the choice of both an appropriate sample and an appropriate measure of digital literacy. None of our measures of digital literacy vary sufficiently in our "high digital literacy" sample, for example, and none of them can explain any of the heterogeneity in the capacity of that sample to perform a simple information retrieval task. However, consider a much more advanced task related to digital media effects: the large-scale detection of "bot" accounts on Twitter. Say that one-third of the high-DL sample could perform this task, but less than 1 percent of any other sample could; MTurk would obviously be a poor choice of subjects for this experiment. None of our existing measures of digital literacy could explain the variation in performance among the high-DL sample, but perhaps a single survey question about Python fluency could.

We thus see our contribution as relevant for broader conversations about how political science should handle novel online platforms for subject recruitment and how novel theoretical constructs should be evaluated. The case of MTurk is illustrative. Berinsky et al. (2012) "introduced" the platform to the discipline, and while their validation work is careful and nuanced, it has frequently been cited to imply that MTurk is generally "okay to use."[16] The article made a large impact and is overwhelmingly the most-cited article published in *Political Analysis* in the past decade. Numerous papers have argued about the merits of MTurk as a respondent pool for experimental studies. Coppock (2019), the most recent and comprehensive, establishes that a wide variety of experimental findings generated from nationally representative surveys can be replicated using MTurk samples. The reasoning is that since treatment effects are similar for many different types of subjects, the composition of online convenience samples along commonly measured demographic and political characteristics is relatively unimportant for recovering experimental treatment effects. Crucially, however, results about the generalizability of MTurk samples are not themselves generalizable to all possible experiments.

They are also not necessarily generalizable to alternative online convenience sample marketplaces. For example, Coppock and McClellan (2019) explore and offer justification for the use of Lucid, a source of research subjects used in this study. Aside from the validation exercise, the paper offers a useful discussion of the scientific value of convenience samples for testing theories. Coppock and McClellan accurately observe that concerns about external validity are often poorly motivated. Whether experiments conducted on a convenience sample and a nationally representative sample produce identical treatment effects is immaterial; what matters is whether a given sample is *theoretically relevant*, an argument with which we strongly agree. They propose,

[16]The sociological processes by which political methodology deems certain elements of research design generally acceptable are central to the practice of political science but have largely escaped systematic study within the discipline. For example, it is possible to think of MTurk as a "trading zone" facilitating exchanges at the intersection of multiple fields and intellectual approaches (Collins et al., 2007).

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

for example, that a sample of French speakers would serve as an inappropriate sample to test a theory about the effect of reading an English-language newspaper. *Literacy* is self-evidently a crucial moderator of textual media effects.

This example works because social scientists intuitively understand how literacy operates in the world. A paper purporting to find evidence of treatment effect heterogeneity in the French literacy of respondents would be bizarre; *literacy* is a bedrock of our shared model of media effects. We feel nearly as confident about *digital literacy*, though we are well aware that no consensus yet exists on this point. By conducting research demonstrating the importance of digital literacy in mediating digital media effects, we hope to change the intuitions of social scientists and make the choice of online convenience samples contingent on the theoretical argument being tested, in the same way that the choice of constructs as potential treatment effect moderators are selected by theoretical logic.

In both cases, of course, theories can be bolstered or weakened through empirical tests. The intersection of highly heterogeneous and rapidly shifting digital media experiences with the proliferation of companies and organizations offering online convenience samples makes it impractical to empirically verify the validity of every possible research question involving digital media effects on each of these novel samples. Munger (2021) asks whether "a given class of research designs should be 'generalizable from MTurk samples until proven otherwise,'" and we think that the best way to answer this question is to seriously engage with the theory in question.

We would therefore qualify the claim in Coppock and McClellan (2019) that "Whether or not future experiments will also exhibit low treatment effect heterogeneity is, of course, only a matter of speculation." We argue that digital literacy is likely to be a key moderator of digital media treatment effects, that it varies widely in the current US population (and that it does so even among the current population of internet users), and that this is a significant problem for opt-in crowdworker platforms like MTurk for which digital literacy affects selection into the sample.

Although social science has not yet accumulated the dozens of separate studies necessary to firmly establish this claim, the trend of political life increasingly taking place online is not, we argue, "only a matter of speculation." Media technology has radically increased media heterogeneity, a development that shows no signs of abating. This is perhaps a problem for the ecological validity of a given media survey experiment, as a greater variety of stimuli are thus necessary to ensure a representative sample of even a specific type of media (Hovland et al., 1949). More immediately relevant is the continuing existence of the "second digital divide" in computer skills identified by Hargittai 20 years ago. In the absence of an unprecedented increase in the level of digital literacy among those at the bottom of the distribution, we have strong reason to expect that future experiments will exhibit greater effect heterogeneity along this dimension. But only, of course, if those experiments are conducted using samples with sufficient variation in the moderator, and valid measures of it.

7. Conclusion

As awareness increases of the potential for endogenous selection of digital literacy in commonly used samples, scholars will need to grapple with its substantive importance as a moderator in literatures not directly related to the study of online misinformation. Below we discuss three additional areas for exploration of effect heterogeneity by digital literacy. In addition to sampling considerations, these areas also raise questions about the specific measures to be used, which will reflect overlapping but distinct dimensions of the multifaceted construct of digital literacy.

Political knowledge. While political knowledge has long been established as a strong predictor of issue-position stability and ideological constraint, Kleinberg and Lau (2019) find that this relationship is now *reversed* for people belonging to the "internet generation." People for whom searching for information online is second nature do not need as much political information

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

stored in their biological memories because they are able to access that information in their digital memories. These findings are consistent with psychological research that conceptualizes people's use of the internet as a kind of external or "transactive" memory (Sparrow et al., 2011). Accordingly, people are better at remembering *where* to find information than the information itself if they expect to be able to find it later—even if this makes them think that they are more knowledgeable than they actually are (Fisher et al., 2015; Ward, 2021).

Future research on the determinants and consequences of political knowledge will need to explicitly take into account variation in individuals' ability to seek out and process online information. Not only might the effect of knowledge vary across levels of digital literacy, but knowledge itself might function differently for people aware that answers will always be at their fingertips. Smith et al. (2020) suggest that "Scholars who are interested in measuring political knowledge should take efforts to minimize and diagnose search behavior," but this assumes a static, crystallized view of knowledge that may no longer hold for people who effectively externalize their store of factual information. This novel, transactive view of political knowledge is distinct from the traditional, static view—but overall, given the evidence in Kleinberg and Lau (2019), the former may now be more relevant for understanding how citizens engage in the process of transforming political knowledge into political preferences and actions. We speculate that none of our survey measures may be particularly well-suited to operationalize digital literacy for this domain of inquiry, but that a direct measure of age cohort should be used until a more theoretically relevant measure is developed.

Nudges and defaults. Behavioral scientists have demonstrated that classes of relatively low-cost interventions can have outsized effects. These "nudges" have been promoted as a way to guide human decisions in a prosocial direction while preserving liberty of choice, as in the case of default options (Thaler and Sunstein,`2009). Scholarship on the topic has focused on domains with a clear public welfare dimension, such as education and public health. But the insights of behavioral economics can be fruitfully applied to internet media, as contemporary accounts overlook the hidden obstacles and defaults that structure people's behavior (Lorenz-Spreen et al., 2020). To take a simple example, modern web browsers come pre-loaded with bookmarks for large news and entertainment sites. Many people still use portals for email and other services which link to headlines, weather and other information. Sometimes, such sites automatically load on startup or with a new tab. It is not hard to customize one's settings, but the perceived cost of doing so may be too high for people lacking in digital literacy, with observable implications for political news diets (e.g., Guess, 2021).

The effects of online choice architecture are not always innocuous, as more recent writing on "sludge" (Thaler, 2018) and "dark patterns" (Mathur et al., 2019) has documented. As worries about online privacy and surveillance become more acute, the ability to protect oneself from insidious efforts to influence consumer and political choices will depend on what Hargittai and Micheli (2019) call "awareness of what is possible"—a dimension of internet skills comprising knowledge about default settings that can be changed. This arguably also covers defensive measures that can be taken against malicious or unscrupulous actors, such as installing blockers for illicit tracking. At a more basic level, people lacking in digital literacy may be unable to defend themselves against clickbait, misleading advertisements, and spam (perhaps political in nature; see Perlstein 2012).

Microtargeting and campaign persuasion. The possibility that digital literacy captures heretofore unobserved heterogeneity in people's responses to online "sludge" raises more general questions about the effects of modern campaign techniques. Evidence already suggests that online advertisements can generate substantial heterogeneity; Kaptein and Eckles (2012) find that heterogeneity in social influence strategies is large relative to the average effect, for example. Moreover, since ads follow the same rules of "engagement" as other types of online content, campaign targeting of favored demographic and political subgroups is subject to feedback loops in which users

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

most likely to respond to appeals will become even more likely to be targeted in the future via platforms' ad optimization algorithms (e.g., Eckles et al., 2018; Zarouali et al., 2020). Given generally low clickthrough rates, the kinds of people who are most likely to engage with commercial or political appeals on social media has important implications for democratic participation, representation, and campaign strategy. Future research should explore whether a digital literacy measure that captures overall awareness of how the social media environment operates—similar to the concept of "algorithmic skills" proposed by Klawitter and Hargittai (2018)—explains variation in this important but understudied online political behavior.

Acknowledgments. This research was supported by an unrestricted Facebook Integrity Foundational Research Award and approved by the Princeton University Institutional Review Board (Protocol 11745). The authors are grateful to Eszter Hargittai for discussions and feedback on this project as it developed. We are indebted to Melissa Morrone at the Brooklyn Public Library and Kelsey Ockert and Morgan Taylor at the Princeton Public Library for their advice and help recruiting respondents. Thanks to Drew Dimmery, Jennifer Oser, Jason Reifler, the editor, and two anonymous reviewers for providing helpful comments. Special thanks to Nejla Asimovic and Jacob Sweetow for outstanding research assistance.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2022.17.

References

Abbate J (2000) *Inventing the Internet. Inside technology.* M I T PRESS, MA.

Arceneaux K and Johnson M (2013) *Changing minds or changing channels?: Partisan news in an age of choice.* Chicago: University of Chicago Press.

Aronow PM, Kalla J, Orr L and Ternovski J (2020) Evidence of rising rates of inattentiveness on Lucid in 2020.

Barberá P (2018). Explaining the spread of misinformation on social media: evidence from the 2016 U.S. presidential election. In: Symposium: fake news and the politics of misinformation. APSA; 2018.

Bawden D (2008). Origins and concepts of digital literacy. In Lankshear C and Knobel M (eds), Digital Literacies: Concepts, Policies and Practices, pp. 17–32.

Bennett WL and Iyengar S (2008) A new era of minimal effects? The changing foundations of political communication. *Journal of communication* **58**, 707–731.

Berinsky AJ, Huber GA and Lenz GS (2012) Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis* **20**, 351–368.

Bode L, Vraga EK, Borah P and Shah DV (2014) A new space for political behavior: Political social networking and its democratic consequences. *Journal of Computer-Mediated Communication* **19**, 414–429.

Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE and Fowler JH (2012) A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295.

Bond RM, Settle JE, Fariss CJ, Jones JJ and Fowler JH (2017) Social endorsement cues and political participation. *Political Communication* **34**, 261–281.

Brewer R, Morris MR and Piper AM (2016). Why would anybody do this?: Understanding older adults' motivations and challenges in crowd work. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems ACM. pp. 2246–2257.

Burke P, Briggs A, Burke P and Burke UP (2005) *A Social History of the Media: From Gutenberg to the Internet.* New York: Wiley.

Clifford S and Jerit J (2016) Cheating on political knowledge questions in online surveys: an assessment of the problem and solutions. *Public Opinion Quarterly* **80**, 858–887.

Collins H, Evans R and Gorman M (2007) Trading zones and interactional expertise. *Studies in History and Philosophy of Science Part A* **38**, 657–666.

Converse P (1964). The nature of belief systems in mass publics. In David Apter (ed.), Ideology and Discontent. New York: Free Press.

Coppock A (2019) Generalizing from survey experiments conducted on Mechanical Turk: a replication approach. *Political Science Research and Methods* **7**, 613–628.

Coppock A and McClellan OA (2019) Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics* **6**, 2053168018822174.

Eckles D, Gordon BR and Johnson GA (2018) Field studies of psychologically targeted ads face threats to internal validity. *Proceedings of the National Academy of Sciences* **115**, E5254–E5255.

Eshet Y (2004) Digital literacy: a conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia* **13**, 93–106.

Fisher M, Goddu MK and Keil FC (2015) How the Internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General* **144**, 674.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

Flanagin AJ and Metzger MJ (2007) The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society* **9**, 319–342.

Forestal J (2017) The architecture of political spaces: trolls, digital media, and Deweyan democracy. *American Political Science Review* **111**, 149–161.

Gilster P (1997) *Digital Literacy*. Wiley Computer Pub., New York.

Grinberg N, Joseph K, Friedland L, Swire-Thompson B and Lazer D (2019) Fake news on Twitter during the 2016 US presidential election. *Science* **363**, 374–378.

Guess AM (2021) (Almost) everything in moderation: new evidence on Americans' online media diets. *American Journal of Political Science* **65**, 1007–1022.

Guess A, Nagler J and Tucker J (2019) Less than you think: prevalence and predictors of fake news dissemination on Facebook. *Science Advances* **5**, eaau4586.

Hargittai E (2001). Second-level digital divide: mapping differences in people's online skills. arXiv preprint cs/0109068.

Hargittai E (2005) Survey measures of web-oriented digital literacy. *Social Science Computer Review* **23**, 371–379.

Hargittai E (2009) An update on survey measures of web-oriented digital literacy. *Social Science Computer Review* **27**, 130–137.

Hargittai E (2010) Digital na (t) ives? Variation in internet skills and uses among members of the "net generatio". *Sociological Inquiry* **80**, 92–113.

Hargittai E and Micheli M (2019) Internet skills and why they matter. Society and the Internet: how networks of information and communication are changing our lives. **2019**, 109.

Hargittai E and Shaw A (2020) Comparing Internet experiences and prosociality in Amazon Mechanical Turk and population-based survey samples. *Socius* **6**.

Holt K, Shehata A, Strömbäck J and Ljungberg E (2013) Age and the effects of news media attention and social media use on political interest and participation: do social media function as leveller?. *European Journal of Communication* **28**, 19–34.

Hovland CI, Lumsdaine AA and Sheffield FD (1949) Experiments on mass communication. (Studies in social psychology in World War II), Vol. 3..

Hunsaker A and Hargittai E (2018) A review of Internet use among older adults. *New Media & Society* **20**, 3937–3954.

Kaptein M and Eckles D (2012) Heterogeneity in the effects of online persuasion. *Journal of Interactive Marketing* **26**, 176–188.

Klawitter E and Hargittai E (2018) "It's like learning a whole other language": the role of algorithmic skills in the curation of creative goods. *International Journal of Communication* **12**, 3490–3510.

Kleinberg MS and Lau RR (2019) The importance of political knowledge for effective citizenship: differences between the broadcast and Internet generations. *Public Opinion Quarterly* **83**, 338–362.

Koltay T (2011) The media and the literacies: media literacy, information literacy, digital literacy. *Media, Culture & Society* **33**, 211–221.

Kruizinga MD, Stuurman FE, Exadaktylos V, Doll RJ, Stephenson DT, Groeneveld GJ, Driessen GJA and Cohen AF (2020) Development of novel, value-based, digital endpoints for clinical trials: a structured approach toward fit-for-purpose validation. *Pharmacological Reviews* **72**, 899–909. Available from: https://pharmrev.aspetjournals.org/content/72/4/899.

Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, Keller S, Weinryb I, Green M and Duan L (2006) Fit-for-purpose method development and validation for successful biomarker measurement. *Pharmaceutical Research* **23**, 312–328.

Lorenz-Spreen P, Lewandowsky S, Sunstein CR and Hertwig R (2020) How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nature Human Behaviour* **4**, 1102–1109.

Luskin RC (1987) Measuring political sophistication. *American Journal of Political Science* **31**, 856–899.

Luskin RC (1990) Explaining political sophistication. *Political Behavior* **12**, 331–361.

Mathur A, Acar G, Friedman MJ, Lucherini E, Mayer J, Chetty M and Narayanan A (2019) Dark patterns at scale: findings from a crawl of 11 K shopping websites. *Proceedings of the ACM on Human-Computer Interaction* **3**, 1–32.

Mathur A, Wang A, Schwemmer C, Hamin M, Stewart BM and Narayanan A (2020). Manipulative tactics are the norm in political emails: evidence from 100 K emails from the 2020 US election cycle.

Munger K (2019). Temporal validity in online social science.

Munger K, Luca M, Nagler J and Tucker J (2020) The (null) effects of clickbait headlines on polarization, trust, and learning. *Public Opinion Quarterly* **84**, 49–73.

Munger K, Gopal I, Nagler J and Tucker JA (2021) Accessibility and generalizability: are social media effects moderated by age or digital literacy? *Research & Politics* **8**, 20531680211016968.

Mutz DC (2015) *In-Your-Face Politics: The Consequences of Uncivil Media*. New York: Princeton University Press.

Osmundsen M, Bor A, Vahlstrup PB, Bechmann A and Petersen MB (2020). Partisan polarization is the primary psychological motivation behind "fake news sharing on Twitter".

Pearl J (2009) *Causality*. London: Cambridge University Press.

Perlstein R (2012) The long con: mail-order conservatism. *The Baffler* **21**, 22–32.

Prior M (2007) *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. London: Cambridge University Press.

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press

Ray C (2016). Fit-for-purpose validation. In Translating Molecular Biomarkers into Clinical Assays. Springer, pp. 1–15.

Redmiles EM, Chachra N and Waismeyer B (2018). Examining the demand for spam: who clicks? In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, p. 212.

Settle JE (2018) *Frenemies How Social Media Polarizes America*. New York: Cambridge University Press.

Sharkey J (2013) Establishing twenty-first-century information fluency. *Reference & User Services Quarterly* **53**, 33–39.

Smith B, Clifford S and Jerit J (2020) TRENDS: how Internet search undermines the validity of political knowledge measures. *Political Research Quarterly* **73**, 141–155.

Sood G and Lelkes Y (2018). Don't expose yourself: discretionary exposure to political information. In Oxford Research Encyclopedia of Politics.

Sparrow B, Liu J and Wegner DM (2011) Google effects on memory: cognitive consequences of having information at our fingertips. *Science* **333**, 776–778.

Sundar SS and Marathe SS (2010) Personalization versus customization: the importance of agency, privacy, and power usage. *Human Communication Research* **36**, 298–322.

Thaler RH (2018) Nudge, not sludge. Science.

Thaler RH and Sunstein CR (2009) *Nudge: improving decisions about health, wealth, and happiness*. New York: Penguin.

Vraga EK and Tully M (2015) Media literacy messages and hostile media perceptions: processing of nonpartisan versus partisan political information. *Mass Communication and Society* **18**, 422–448.

Vraga EK, Tully M, Kotcher JE, Smithson AB and Broeckelman-Post M (2015) A multi-dimensional approach to measuring news media literacy. *Journal of Media Literacy Education* **7**, 41–53.

Ward AF (2021) People mistake the internet's knowledge for their own. *Proceedings of the National Academy of Sciences* **118**, 2105061118.

Wineburg SA and McGrew S (2018) Lateral reading and the nature of expertise: reading less and learning more when evaluating digital information. *Sociology* **2015**, 12–15.

Zarouali B, Dobber T, De Pauw G and de Vreese C (2020) Using a personality-profiling algorithm to investigate political microtargeting: assessing the persuasion effects of personality-tailored ads on social media. *Communication Research* **2020**, 0093650220961965.

Zhang B, Mildenberger M, Howe PD, Marlon J, Rosenthal S and Leiserowitz A (2018) Quota sampling using Facebook advertisements. *Political Science Research and Methods* **8**, 1–16.

Cite this article: Guess AM, Munger K (2023). Digital literacy and online political behavior. *Political Science Research and Methods* **11**, 110–128. https://doi.org/10.1017/psrm.2022.17

https://doi.org/10.1017/psrm.2022.17 Published online by Cambridge University Press