[Published in *American Behavioral Scientist*]

Representativeness of social media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+ and Instagram

Grant Blank, Oxford Internet Institute, University of Oxford

Christoph Lutz, Handelshoyskolen BI, Oslo, Norway

DOI: 10.1177/0002764217717559

Abstract

Sociological studies show that Internet access, skills, uses and outcomes vary between different population segments. However, we lack differentiated statistical evidence of the social characteristics of users of distinct social media platforms. We address this issue using a representative survey of Great Britain and investigate the social characteristics of six major social media platforms. We find that age and socio-economic status are driving forces of several – but not all – of these platforms. The findings suggest that no social media platform is representative of the general population. The unrepresentativeness has major implications for research that uses social media as a data source. Social media data cannot be used to generalize to any population other than themselves.

Keywords: digital divide; representativeness; social media; Facebook; Twitter; LinkedIn; Instagram; Pinterest; Google+

Representativeness of Social Media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+ and Instagram

Introduction

Social media are ubiquitous. As of mid-2015, Facebook, for example, boasts 1.44 billion monthly active users. More than half of all Americans have a profile on that platform (Duggan, Ellison, Lampe, Lenhart, & Madden, 2015). The fastest growing social media site is Instagram with more than 600 million users – a staggering growth of more than 100 million within a year (Statista, 2017). Given the massive adoption of social media, it does not surprise that researchers are increasingly interested (Wilson, Gosling, & Graham, 2012). In fact, we see an explosion of studies on Facebook and Twitter after 2006, the year Twitter was founded. Figure 1 shows number of published articles in the Web of Science with "Facebook", "LinkedIn", "Twitter", "Pinterest", "Google+" or "Instagram" in the title. It reveals that most studies are published about Facebook and Twitter. The other social media platforms have received much less attention. LinkedIn, for example, has seen 60 entries in the same period and even fewer are published on Instagram, Pinterest and Google+. Since less studied social media platforms disappear at the bottom of Figure 1, Table 1 shows their frequencies.

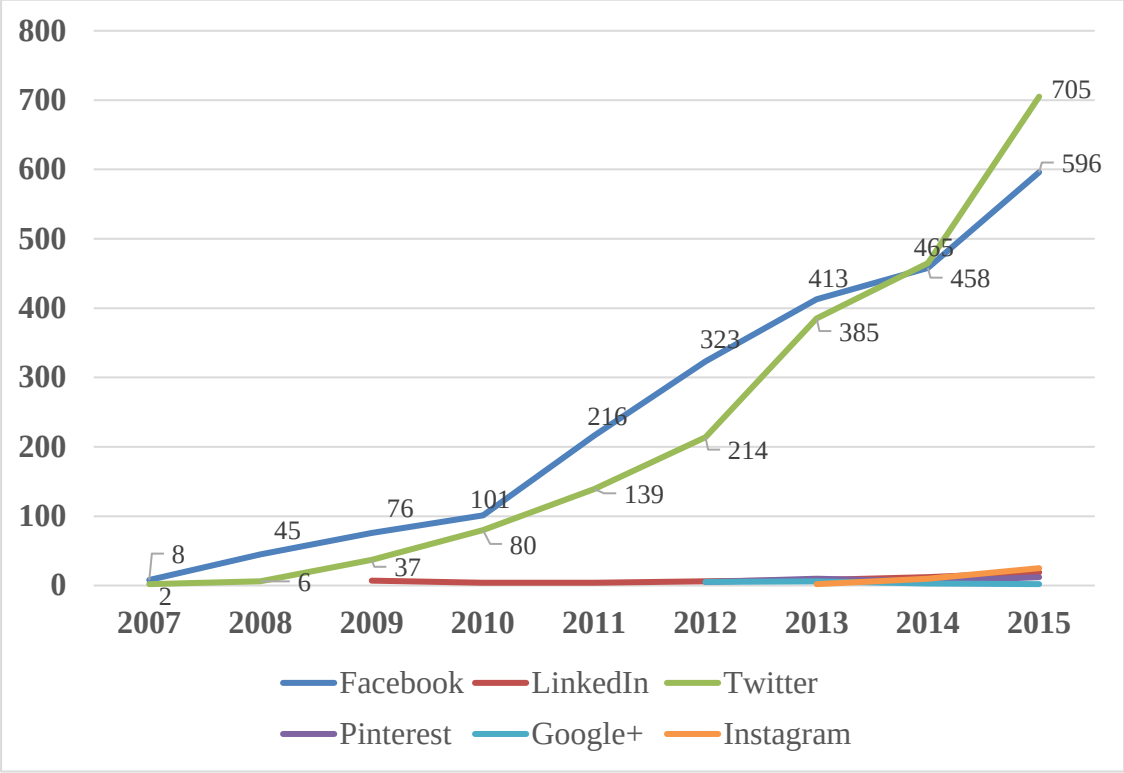


Figure 1: Articles with SNS name in the title in Web of Science

Table 1: Number of articles with the SNS name in the title in the Web of Science

Year	2009	2010	2011	2012	2013	2014	2015	Total
LinkedIn	7	4	4	6	8	12	19	60
Pinterest				5	10	6	12	33
Google+				5	6	3	2	16
Instagram					2	10	25	37

In this article, we provide a multivariate analysis of the digital divide among users of six social network sites (SNS). Despite numerous studies of the digital divide, relatively little is known about the characteristics of people who participate in social media platforms. The few studies of social media user characteristics have often used non-representative samples, such as students (Hargittai, 2007) or employees of a tech company (Archambault & Grudin, 2012). This study joins the relatively small number of studies that describe the populations that use social media (Wells & Link, 2014; Blank, 2016).

At the same time, much social media research focuses on a single platform, often generalizing from that platform but neglecting the biases that come with its specific audiences. This is especially serious when studies rely on user-generated (trace) data or recruit participants through a particular social media platform. In this paper, we address this problem with a large survey data set representative of the British population. The data set includes questions on a wide range of Internet attitudes and uses and covers most major social media platforms. Thus, we can tackle the question of social media use on the platform-level and investigate antecedent conditions commonly not considered in descriptive studies (e.g., Duggan et al., 2015). In the tradition of digital inequality research, we show that users of each platform have unique social characteristics. This has major implications for social media research in terms of sampling and biases. The two central research questions of the paper are: *What are the user characteristics for six major social media platforms in Great Britain? Which potential biases for platform-specific social media research arise from these user characteristics?* Before the empirical analyses, we briefly summarize previous research on the topic of divides in social media use. After the results and their discussion, we conclude with a summary of the findings, implications for research using social media data and limitations of our approach.

Literature Review

Social media adoption: What do we know?

One of the first and most cited studies to investigate the use of different social media platforms is Hargittai's (2007) work on college students. She showed how four social media platforms – Facebook, MySpace, Xanga and Friendster – differ in their social profile. Parental education and race/ethnicity differences are noteworthy: students from educated backgrounds

disproportionally use Facebook, whereas the opposite is true for MySpace. In terms of race, the percentage of white and Asian-American Facebook users is significantly higher than that of Hispanic users. By contrast, the latter adopt MySpace in significantly higher proportions. Both Xanga and Friendster cater particularly to Asian-American students and both are not strongly differentiated in terms of education and gender.

The most recent overview of several social media platforms comes from a Pew study (Duggan et al., 2015). They find that income matters more for some platforms (Facebook, Pinterest) than for others (Twitter, Instagram, Tumblr), implying that it is important to differentiate individual platforms when looking at social media adoption. Or in other words, "[…] disaggregating which specific site one is researching is important, because people do not randomly select into their uses, and aggregate analyses of SNS use may make it difficult to identify important trends" (Hargittai, 2007, p. 277).

Table 2 shows the proportions of American and British adults who use the 6 platforms in 2013. The proportion who use Facebook is identical in the US and the UK within the margin of sampling error. A higher proportion of American's use Instagram, LinkedIn and Pinterest, but more British use Twitter.

Table 2: Percent of Internet Users who use each Social Media Platform

	Facebook	Google+	Instagram	LinkedIn	Pinterest	Twitter
USA	71	NA	17	22	21	18
UK	69	37	11	9	3	30

Notes: 2013 data. Total Internet users: USA 80%; UK 78%. Total N: USA 2,003; UK 2,057. Google+ data is not available for the USA.
Sources: USA Duggan et al. 2015; UK OxIS 2013.

Hargittai (2015) uses 2013 Pew survey data to investigate the digital divide in Facebook, LinkedIn and Twitter. Using logistic regression, she finds that age is the strongest demographic predictor and that all three sites are most used by younger individuals. Gender matters for Facebook, where women are more likely to use the platform, and LinkedIn, where

men are more likely to use it. However, it does not matter for Twitter. The race/ethnicity, education and income differences are more complex. Black users are significantly more likely to use Twitter and LinkedIn than white users. Highly educated and high-income individuals are most likely to adopt LinkedIn. Overall Facebook adoption is not stratified along socio-economic lines.

Blank (2016) compares US and UK Twitter use in terms of demographics, attitudes and Internet use. Various comparisons show that Twitter users are younger and better educated than Internet users who don't use Twitter, and Internet users are, in turn, younger and better educated than non-users. US and UK Twitter users are not identical. Among several differences, UK Twitter users are better educated than non-users; education is not significant in the US data. We return to the implications of these studies in the discussion

The Pew studies are themselves largely descriptive and do not look at a range of antecedent conditions. Especially, they neglect some antecedents deemed important in previous research, such as Internet skills (Hargittai, 2002, 2010; Van Deursen & Van Dijk, 2015), self-efficacy (Hoffmann, Lutz, & Meckel, 2015) and attitudes (Correa, 2010). In the next section, we proceed to enrich the basic coverage of demographic antecedents with a set of cognitive and attitudinal predictors.

Research models

To investigate the user base of social media platforms more holistically, we use twelve predictors: age, gender, income, education, marital status, number of children in the household, self-efficacy, skills, mobile vs. non-mobile Internet use, experience, privacy concerns, and trust in Internet companies. We briefly describe the rationale for including each variable.

Age, gender, income and education are the "classical" variables used in digital inequalities research. *Age* is typically a strong predictor across a wide range of Internet uses, especially content-related ones via social media. In general, younger users are more likely to use participatory media, especially for non-political purposes – a finding that is robust across a range of countries (Blank, 2013; Correa, 2010; Hargittai & Walejko, 2008; Hoffmann et al., 2015; Van Deursen & van Dijk, 2014). Thus, we expect age to be a strong and significant predictor across all platforms.

Gender has a less obvious influence. While some platforms are "gendered" – such as Pinterest and Instagram – others are used almost equally by men and women, e.g., Twitter and LinkedIn (Duggan et al., 2015). Among individuals living in the US and Canada, women adopt social media to a higher degree than men (Haight, Quan-Haase, & Corbett, 2014; Hampton, Goulet, Rainie, & Purcell, 2011). Differences in the gendered adoption of platforms show the need to disaggregate.

Income and *education* can be interpreted as markers of socio-economic status (*SES*). SES is an important predictor for general Internet use but to a lesser degree for social media use and content production (Blank, 2013, for the discussion). However, studies of participation and social media divides have mostly looked at overall social media/SNS adoption (Haight et al., 2014), aggregated indices (such as political, skilled, and social & entertainment content; Blank, 2013), or individual activities (Blank, 2016). Research has not systematically distinguished and disaggregated platforms, so that we have little knowledge how they differ in terms of the users' SES. Based on previous research, we expect that SES has a positive influence on social media platform adoption, especially when it comes to newer platforms, such as Pinterest and Instagram.

Although *race* might be an important predictor of social media use, we did not include it in the analysis because non-whites are not present in the UK in sufficient numbers to obtain meaningful results (e.g., there are only 61 Black respondents in the dataset). We included race as an independent variable in the regression but it is never significant.

Individuals' *life circumstances* might play a role in the adoption of social media services. Previous research on general Internet access has shown that the household status can matter, so that, for example, households with a child/children aged 12-24 in Germany are more likely to use the Internet than those without children (Korupp & Szydlik, 2005). We included marital status and number of children in the household as two such variables. None of the platforms considered in this study is targeted at specific groups in terms of family situation and marital status (e.g., dating platforms would be platforms where marital status matters a lot). Previous research on social media has largely excluded such information so that we cannot make reliable predictions.

The *device* primarily used to access the Internet (mobile vs. non-mobile) might also affect which platforms users adopt (Pearce & Rice, 2013). Some social media are particularly easy for mobile use (Instagram, Twitter), whereas others (Facebook, LinkedIn) – due to their range and complexity of functionalities – are suited for both desktop and mobile use. We propose that Instagram, Pinterest and Twitter are especially mobile-friendly, while for Facebook, LinkedIn and Google+ the device should not have an influence.

We included *experience* mainly as a control variable and think that *self-efficacy* and (self-rated) *skills* should be more powerful explanatory constructs. Self-efficacy is understood here as "people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (Bandura, 1994, p. 71). Previous

research has shown that both skills and self-efficacy have a strong and positive impact on social media use (e.g., Hargittai & Litt, 2011; Hoffmann et al., 2015).

Finally, we include two attitudinal variables: *privacy concerns* and *trust*. Both have been important as research fields of their own (Beldad, De Jong, & Steehouder, 2010; Smith, Dinev, & Xu, 2011). However, empirical results have been ambiguous. In many cases privacy concerns were not significantly related to user behavior (Tufekci, 2008), leading to a so-called privacy paradox (Barnes, 2006). Trust has been investigated in the context of online shopping and e-business, where financial transactions are involved, but not so much in more hedonic contexts like social media (Blank & Dutton, 2012; Van der Heijden, 2004). If there are effects, they should be weak and positive in the case of trust but weak and negative in the case of privacy concerns.

Methodology

Data

The Oxford Internet Surveys (OxIS) collect data on British Internet users and non-users. Conducted biennially since 2003, the surveys are nationally representative random samples of more than 2,000 individuals aged 14 and older in England, Scotland, and Wales. Interviews are conducted face-to-face by an independent survey research company. The response rate for 2013, the latest available data set, was 51%. The analyses below are based on 1,611 Internet users out of the full sample of 2,053 respondents.

The full sample includes people who do not use the Internet — non-users. If we were to include non-users in the logistic regressions below there would have been two groups of respondents who had zeros on the dependent variable: either those who are non-users of the Internet or Internet users who do not use social media. The causal processes that generate

non-use of the Internet are very different from those that generate non-use of particular capabilities on the Internet. This means that many of the predictive variables would be different. To include both groups in the same model is theoretically problematic. Consequently, all of the models reported here are based only on Internet users, N = 1,611.

In our sub-population, 51% are male and 49% female, the average age is 45.5 years. The modal category of education is “secondary school or equivalent” with 40% of respondents and the next most frequent category “university or equivalent” with 30% of respondents. Income shows a right-skewed distribution, with the average value of 2.6, falling in the category “£12,500-£20,000”. The numbers show social media users are somewhat younger, better educated, with higher incomes than the British population.

Measures

The questionnaire addresses the participants’ Internet use and includes a broad range of attitudinal variables. The dependent variables for the regressions consist of six items that assess the adoption of the six social media services in question. The question wording was: “Do you use any of the following?” The individual platforms were then read out individually and the respondents could answer with *“Yes”* or *“No”*. We measured income as yearly household income before taxes in eight categories, ranging from “up to £12,500” to “£70,000-£80,0000”. Marital status was measured with five categories: “single”, “living with partner”, “married”, “divorced” and “widowed”. The number of children in the household was counted directly. We measured the Internet device with a categorical variable with the following values: “mostly mobile”, “mostly something else”, “both equally” and “do not use mobile Internet”. Skills were assessed with the item: “How would you rate your ability to use the Internet“, with response choices ranging from “bad” (1) to “excellent” (5). We used a

composite index of six items for self-efficacy (see Appendix). The scale has a Cronbach's Alpha of 0.890. Privacy concerns and trust were both assessed with one item each: "People should be concerned about protection of credit card details when they are using new technologies" (privacy concerns) and "How much trust do you have in the people providing Internet services?" (trust). Finally, experience was measured with one item: "About how long have you been going online?" The range is from "one month" to "15 years or more".

Methods

We used logistic regression to address the research questions. The statistical analysis was carried out with Stata 13.1, using the "svy" prefix and post-stratification weights to handle the cluster sample.

Results

According to the results shown in Table 3, each platform has different demographic and social characteristics.

For Facebook, age and gender are decisive, with younger and female users being more likely to adopt it. Moreover, the significant effect of mobile use – compared to non-mobile and mostly non-mobile – shows that the device matters. In other words, Facebook seems to be especially attractive to mobile users. This finding coincides with the strategy of the company, which is increasingly tailoring its services to mobile devices. Finally, self-efficacy increases the likelihood of Facebook adoption. Scoring low on self-efficacy presents a barrier to using Facebook. Interestingly (self-reported) skills turn out to not be significant, which implies that the barriers to using Facebook are matters of choice rather than a lack of actual Internet skills.

These results are mostly consistent with Hargittai's (2015) analysis of the Pew data. She also finds significant age and gender differences but no income or education differences. Our models control for additional variables beyond demographic characteristics so any substantive differences may be due to differences in model specification.

The picture is different for LinkedIn, which is targeted at professional users. In contrast to Facebook, age and gender do not matter but income does. High income citizens are significantly more likely to use LinkedIn than their lower income counterparts. In addition, we find that divorced people are significantly more likely to use LinkedIn than single people. This finding could be due to life circumstances and life course patterns. Divorce is most likely to occur when individuals are between 30 and 50 years old (ONS, 2015), with the peak at the age of 40-44. In this phase, citizens are also in the middle of their professional career and hence more likely to adopt LinkedIn.[1] Similarly to Facebook, self-efficacy strongly and positively influences the use of the platform but (self-reported) skills do not.

Table 3. Logistic regression odds ratios for demographic and attitudinal predictors on use of Facebook, LinkedIn, Twitter, Pinterest, Google+ and Instagram

Predictor	Facebook	LinkedIn	Twitter	Pinterest	Google+	Instagram
Age	0.957***	0.982	0.961***	0.949*	0.996	0.978
Gender	1.892***	0.697	1.176	0.971	1.047	0.846
Income	1.130	1.177*	1.239**	1.281*	1.108	1.177
Education (reference: no qualification)						
Secondary	0.913	0.585	1.217	0.361	1.008	2.753
Further	1.590	0.835	0.884	2.568	0.920	1.538
Higher	0.465*	1.561	0.918	0.931	1.050	2.919
Marital status (reference: single)						
Married	1.513	1.163	0.989	2.752	0.674	0.663

[1] There are relatively few respondents in the "divorced" row, so the large odds ratio may also be a methodological artifact of the small N.

Living with partner	1.578	1.292	0.940	1.138	1.074	0.750
Divorced	2.116	4.078**	1.515	19.175***	0.754	2.198
Widowed	1.252	1.264	1.387	7.640	0.034*	1.189
Children in household	0.855	0.708	0.848	1.441	0.792	1.032
Skills	0.981	1.233	1.398	3.488**	1.217	1.371**
Location: mobile vs. non-mobile (reference: mostly mobile phone)						
Mostly use something else	0.341**	0.966	0.536*	0.899	0.787	0.400*
Both equally	0.898	1.131	1.151	0.635	1.067	0.792
Don't use mobile Internet	0.194***	0.215*	0.136***	-	0.620	0.061**
Experience	0.967	1.109*	1.040	0.952	0.959*	0.961
Privacy concerns	0.854	1.083	1.035	0.926	0.989	0.771
Trust in Internet providers	1.185	0.863	1.261*	0.575**	0.854	1.063
Self-efficacy	1.086***	1.053**	1.063**	0.941	0.990	1.092**
N	1052	1051	1052	860	1049	1048

Notes: ***: $p < 0.001$ **: $p < 0.01$ *: $p < 0.05$; “don’t use mobile Internet” was not included because it perfectly predicts non-use of Pinterest; no R^2

In terms of scholarship, surprisingly few studies look at LinkedIn (see Table 1). Given its popularity, we think that LinkedIn deserves more attention, especially when research questions related to users’ professional life are the center of attention. Our results for LinkedIn differ from Hargittai’s (2015) results in some ways. Unlike her, we do not find significant age and gender differences. Given that the age effect for LinkedIn is only significant at the 5-percent level in Hargittai’s (2015) study and that she does not control for marital status, we can conclude that LinkedIn is less socially structured in terms of age than other major social media platforms. Other research has found that LinkedIn usage does not differ sharply by age (Blank & Groselj, 2016b) so absence of an age effect is not surprising. However, the difference in gender between Hargittai’s (2015) American sample and our British sample might indicate differences in use between the two countries.

Twitter adopters, again, have different characteristics. Here, age and income are significant but not education and gender. Like Facebook, self-efficacy and the use of mobile devices have a positive impact on the adoption. Again, skills do not matter. Many older, low-income Internet users might not see the purpose of using Twitter and might find its jargon hard to penetrate (hashtags, retweets, abbreviations etc.). This perception and its usefulness only for certain interest groups (entertainment and celebrity news, career advice, marketing, news) could account for the digital divide in Twitter use and its specific demographic profile (Duggan et al., 2015; Hargittai & Litt, 2011). Our results for Twitter match those of Hargittai (2015). We also find a strong age effect, a significant income effect and no significant gender difference. However, in contrast to her, our data do not reveal a significant education variable. Again different model specifications may account for these differences.

An overwhelming majority does not use Pinterest in Great Britain (see Table 2), indicating that it is the most limited of the six services considered. Few variables significantly predict Pinterest adoption. Like Twitter, high income and younger users are substantially more likely to have adopted it. The status of Pinterest as an emerging and new platform means that only early adopters are using it, a group that is traditionally well equipped and tech savvy (Rogers, 2010). The absence of a gender effect surprises somewhat. However, we do find a skills effect with a very high odds ratio and an interesting effect for “divorced”. The large odds ratio for “divorced” most likely a methodological artefact: there are relatively few respondents in the “divorced” row and a large proportion are Pinterest users. Pinterest, as a new, relatively unknown platform, might mostly cater to skilled Internet users. The effect of trust is also remarkable. The more users trust Internet service providers, the less likely they are to adopt Pinterest.

As far as we know, Duggan et al.'s (2015) Pew study is the only investigation where the profile of Pinterest users is assessed with a high quality sample. The most striking finding is the gendered uptake of the platform in the US: In 2014, 42 percent of female Internet users have adopted Pinterest but only 13 percent of males, making it the most gendered platform. Moreover, white users predominant: 32 percent of white Internet users have adopted it but only 12 percent of black users. The significant age and income effects in our sample, however, are in line with the American data.

For Google+, we find few significant demographic differences. Together with LinkedIn, Google+ is the only platform where experience has a significant – and negative – effect. Users with little Internet experience are most likely to adopt Google+. The service's role as the SNS of an Internet giant that many novice Internet users may use as one of their first applications might explain this tendency.

To our knowledge, Hargittai's (2015) research on college students is the only other study to look at the social profile of Google+ users. The only significant effect is that individuals who use the Internet more frequently are also more likely to have adopted Google+. Thus, in line with our results, Google+ seems to have a very loosely defined user base. The research that we found on Google+ (see Table 1) is mostly about using the site in pedagogical contexts.

Finally, Instagram is not influenced by income, age, gender and education. Despite having markedly higher adoption rates than Pinterest, Instagram had not reached mainstream status in Great Britain at the time of the survey. The high odds ratio for skills shows that Internet skills increase the likelihood of Instagram adoption. Like Pinterest, Twitter and Google+, the newness of Instagram may mean that many users have not had the time to become familiar with the rules and codes of the platform.

However, a brief look at the published studies about Instagram (see Table 1) as well as the rise in user numbers in recent years indicates that Instagram might become a more interesting platform for general social media research. Moreover, Instagram is more accessible for research than LinkedIn, Google+ and Facebook because many users have a public profile. Instagram seems to cater to mobile and skilled users with high levels of self-efficacy. Similar to Facebook, we may see a trickledown effect in the coming years, so that broader population groups will adopt Instagram. Despite this, the image-based nature of most data and selective self-presentation patterns make it less fit for analysing certain phenomena on social media such as more spontaneous and negative expressions.

Implications for Representativeness of Social Media data

Digital divide research has usually been concerned with population groups that are not online, and do not receive the benefits of the Internet. The digital divide on social media is important for a different reason. It is important because of what it implies for research using datasets collected from social media users. Can researchers generalize from social media users (on any platform) to any important population? To answer this question we will summarize the results from specific platforms and then draw some general conclusions.

For Facebook the fact that age and gender differences remain significant after accounting for a broad range of attitudinal and behavioural variables shows that Facebook users are more likely to be younger, female and less well educated. Analyses that rely only on Facebook data or that recruit participants solely via Facebook (for example recruiting a convenience sample approach via a survey link on Facebook or another website), should be aware of these demographic biases. Even a random sample of Facebook users will not be

representative of the population, either the US population (see Hargittai 2015) or the UK population.

Studies that rely on LinkedIn as a data source should be aware of its skewed adoption patterns in terms of income. LinkedIn tends to be used frequently by people in knowledge-intensive sectors such as management, marketing, higher education and consulting (Van Dijck, 2013). Thus, studies that want to include low-income individuals and those outside the workforce might not rely on data collected on or through LinkedIn too strongly. LinkedIn is not representative of the population.

Twitter data are widely used. Twitter users tend to be disproportionately younger and wealthier. The primary appeal of Twitter is that data are extraordinarily easy to collect and sample sizes of millions of tweets are not unusual. Of all the platforms studied, Twitter is the least restrictive when it comes to using the data for research purposes. However, Twitter data also have major biases. Blank (2016) looks at Twitter use in the US and UK in detail, showing that Twitter users (or tweets) are not representative of any population and are unsuitable for any research where representativeness is important.

Given the low uptake and specific profile of Pinterest users (like Twitter users: young and wealthy), empirical studies relying on data from the platform should be cautious to generalize to broader population samples.

We are not the first to notice the problems of using social media data to generalize to populations that include non-users of social media or non-users of the Internet. Hargittai (2015) addresses the biases of big data studies when they rely on social media platforms for data when they wish to generalize to a population. If the selection of respondents is biased then the results will also be biased. This is not a problem that can be overcome by collecting more data. If the selection process is biased, then the results will be biased regardless of

sample size. A big sample is just as biased as a small sample. These demographic biases will be reflected in any results. Since most social media tend to be used by younger, wealthier respondents, the results will give information about young, wealthy elites. It is important to note that these problems apply not just to quantitative results, but also to attempts to infer sentiments or attitudes from social media like sentiment analysis. The results will not give information about the general population. They are representative only of themselves.

This point applies to four of the six social media platforms in this study. Only Google+ and Instagram escape this problem, because demographics do not significantly predict their user characteristics. However, Google+ and Instagram data are relatively hard to collect (especially in comparison to Twitter) and there have been few studies that use their data. Furthermore, there is the question of whether Google+ and Instagram users are distinguished by other characteristics beyond demographics, such as psychological variables (e.g. extroversion), that we do not use here.

This suggests that the four biased social media platforms are not suitable research sites when being representative of a population is important. This includes studies of elections or voting intentions, of social attitudes, sentiments or activities of populations.[2]

---

[2] These points do not exhaust the problems of using social media data for research. For a summary, see Blank (2016) and Huberly (2015). Both argue that a major problem with social media data is that the researcher is dependent on the information users decide to supply in the form that users choose to make it available. Researchers cannot ask for information when users have not provided it and the vocabulary is not standardized. Consequently, many influential variables are unavailable, typically including age, gender, income, education, marital status, children, occupation, and others. The researcher has no

Conclusion

In this article, we analyzed the user characteristics of six social media platforms in Great Britain. In addition to demographic variables, we included a range of antecedent conditions deemed important in the literature, such as self-efficacy, skills, privacy concerns and the device of accessing the Internet (mobile vs. non-mobile).

Users are different on each platform. Facebook use is influenced by age and gender, but not income and education. LinkedIn adoption is affected by income, but not age, gender and education. For Twitter, age and income but not gender and education matter. Pinterest depends on age and income (not education and gender), while no demographic characteristics significantly predict Google+ and Instagram use. Next to demographic characteristics, a set of other predictors affect the adoption of individual platforms. Especially skills and self-efficacy foster the uptake of social media services. Whereas self-efficacy matters for established platforms – Facebook, LinkedIn and Twitter – (self-reported) skills are important for the newer players, namely Pinterest and Instagram. Instagram relies both on (self-reported) skills and self-efficacy.

The findings have implications for future research using social media data. First, they indicate the need to consider fine-grained uses (Pearce, 2015). The consideration of individual platforms helped to see a more complete picture of the social structuration of social media than aggregate indices. Second, they call for the inclusion of antecedents beyond demographics. This has proved helpful and substantially extended the explanatory scope of

---

control over the data generating process, which is not designed to produce valid, reliable research data.

the regressions. Third, the findings show that digital inequality is still large and constantly shifting to the newest platform, service and use.

Limitations and suggestions for future research

Despite the contributions, this research has some limitations which restrict its scope and provide points of departure for future research. First, research on digital inequality and sampling biases in social media is only in its infancy and often lacks a strong theoretical grounding. Our study is no exception. The research could be stronger embedded in existing theories of social inequality, such as Bourdieu's capital theory (Robinson, 2009; Sims, 2014), social milieu theory (Lutz, 2016), Weber's theory of stratification (Blank & Groselj, 2015) or the knowledge gap hypothesis (Bonfadelli, 2002). We encourage future research to use such theories to systematically study social media inequality.

Second, our data only covers one point in time. Thus, inferences across time are not possible and the issue of isolating different causal effects remains. Future research on social media divides could use panel designs to explain changes over time.

Third, social media evolution is very fast-paced. New platforms emerge and disappear every year. Five of the six platforms considered here continue to increase their user base (Twitter is the exception) and additional users inevitably change the social profile.

Fourth, other research has shown cross-national differences in social media users (e.g. Blank, 2016). Additional cross-national studies would further our understanding of the national contexts in which social media users operate.

Overall, this study contributes to the research of social media by being one of the first investigations of Internet social media divides in Great Britain, and – to our knowledge – the first to systematically assess with a high quality, broad sample how the user base of different platforms differ, with implications for representativeness and generalization.

References

Archambault, A., & Grudin, J. (2012). A longitudinal study of facebook, linkedin, & twitter use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2741-2750). ACM.

Blank, G. & Dutton, W.H. (2012). Age and trust in the Internet: The centrality of experience and attitudes toward technology in Britain. *Social Science Computer Review*. 30, 135-151. DOI:10.1177/0894439310396186.

Blank, G. (2013). Who creates content? Stratification and content creation on the Internet. *Information, Communication & Society, 16*(4), 590-612.

Blank, G. (2016). The Digital Divide Among Twitter Users and Its Implications for Social Research. *Social Science Computer Review*. Advance online publication. doi: 10.1177/0894439316671698

Blank, G., & Groselj, D. (2015). Examining Internet Use Through a Weberian Lens. *International Journal of Communication, 9*, 2763-2783.

Blank, G., & Groselj, D. (2016b). A dynamic phenomenon: The types and uses of social network sites, 2011-2013. Unpublished research paper.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of Human Behavior* (Vol. 4, pp. 71-81). New York, NY: Academic Press.

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday, 11*(9). Retrieved from http://firstmonday.org/article/view/1394/1312

Beldad, A., De Jong, M., & Steehouder, M. (2010). How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. *Computers in Human Behavior, 26*(5), 857-869.

Bonfadelli, H. (2002). The Internet and knowledge gaps: A theoretical and empirical investigation. *European Journal of communication, 17*(1), 65-84.

Correa, T. (2010). The Participation Divide Among 'Online Experts': Experience, Skills and Psychological Factors as Predictors of College Students' Web Content Creation. *Journal of Computer-Mediated Communication, 16*(1), 71–92.

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). Social Media Update 2014'. *Pew Research Center*, January 2015. Retrieved from http://www.pewinternet.org/2015/01/09/social-media-update-2014/

Haight, M., Quan-Haase, A., & Corbett, B. A. (2014). Revisiting the digital divide in Canada: the impact of demographic factors on access to the internet, level of online activity, and social networking site usage. *Information, Communication & Society, 17*(4), 503-519.

Hampton, K., Goulet, L. S., Rainie, L., & Purcell, K. (2011). Social networking sites and our lives. *Pew Research Center*, June 2011. Retrieved from http://www.pewinternet.org/2011/06/16/social-networking-sites-and-our-lives/

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, 7(4), Retrieved from http://firstmonday.org/article/view/942/864

Hargittai, E. (2007). Whose space? Differences among users and non-users of social network sites. *Journal of Computer-Mediated Communication, 13*(1), 276-297.

Hargittai, E. (2010). Digital na(t)ives? Variation in Internet skills and uses among members of the "net generation" *Sociological Inquiry*, 8(1), 92-113.

Hargittai, E. (2015). Is bigger always better? Potential biases of big data derived from social network sites. *The ANNALS of the American Academy of Political and Social Science, 659*(1), 63-76.

Hargittai, E., & Litt, E. (2011). The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults. *New Media & Society, 13*(5), 824-842.

Hargittai, E., & Walejko, G. (2008). The participation divide: Content creation and sharing in the digital age. *Information, Communication & Society, 11*(2), 239–256.

Hoffmann, C. P., Lutz, C., & Meckel, M. (2015). Content creation on the Internet: A social cognitive perspective on the participation divide. *Information, Communication & Society, 18*(6), 696-716.

Huberly, M. (2015). Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting, 31*, 992–1007

Korupp, S. E., & Szydlik, M. (2005). Causes and trends of the digital divide. *European Sociological Review, 21*(4), 409-422.

Lutz, C. (2016). A social milieu approach to the online participation divides in Germany. *Social Media + Society, 2*(1), 1-14.

ONS – Office for National Statistics (2015). Divorces in England and Wales: 2013. Statistical Bulletin. Retrieved from http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/divorce/bulletins/divorcesinenglandandwales/2013

Pearce, K. E. (2015). Counting to Nowhere: Social Media Adoption and Use as an Opportunity for Public Scholarship and Engagement. *Social Media + Society, 1*(1), 1-3.

Pearce, K. E., & Rice, R. E. (2013), Digital Divides From Access to Activities: Comparing Mobile and Personal Computer Internet Users. *Journal of Communication, 63*(4), 721-744.

Robinson, L. (2009). A taste for the necessary: A Bourdieuian approach to digital inequality. *Information, Communication & Society, 12*(4), 488-507.

Rogers, E. M. (2010). *Diffusion of innovations*. New York, NY: Simon and Schuster.

Sims, C. (2014). From differentiated use to differentiating practices: Negotiating legitimate participation and the production of privileged identities. *Information, Communication & Society, 17* (6), 670-682.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. *MIS Quarterly, 35*(4), 989-1016.

Statista (2017). Number of monthly active Instagram users from January 2013 to December 2016 (in millions). Retrieved from https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/

Tufekci, Z. (2008). Grooming, gossip, Facebook and MySpace: What can we learn about these sites from those who won't assimilate? *Information, Communication & Society, 11*(4), 544–564.

Van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly, 28*(4), 695-704.

Van Deursen, A., & Van Dijk, J. (2014). The digital divide shifts to differences in usage. *New Media & Society, 16*(3), 507-526.

Van Deursen, A. J., & Van Dijk, J. A. (2015). Internet skill levels increase, but gaps widen: a longitudinal cross-sectional analysis (2010–2013) among the Dutch population. *Information, Communication & Society, 18*(7), 782-797.

Van Dijck, J. (2013). 'You have one identity': performing the self on Facebook and LinkedIn. *Media, Culture & Society, 35*(2), 199-215.

Wells, T., & Link, M. (2014). Facebook user research using a probability-based sample and behavioral data. *Journal of Computer-Mediated Communication, 19*(4), 1042–1052.

Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A review of Facebook research in the social sciences. *Perspectives on Psychological Science, 7*(3), 203-220.

Author Biographies

Grant Blank (PhD, University of Chicago) is the survey research fellow at the Oxford Internet Institute, University of Oxford, United Kingdom. He is a sociologist specializing in the political and social impact of computers and the Internet, the digital divide, statistical and qualitative methods, and cultural sociology. He is currently working on a project asking how cultural hierarchies are understood in online reviews of cultural attractions. His other project

studies the effects of search engines on politics. He can be reached at grant.blank@oii.ox.ac.uk; visit also http://www.oii.ox.ac.uk/people/blank/

Christoph Lutz is assistant professor at the Nordic Centre for Internet & Society, BI Norwegian Business School (Oslo). His research interests include online participation, privacy, the sharing economy, social robots, and digital serendipity [email: christoph.lutz@bi.no].

Appendix

Item	**Wording (Scale)**
	How confident do you feel that you are able to…
a	*…judge the reliability of online content?*
b	*…remove a virus that infected your computer?*
c	*…participate in a discussion online?*
d	*…make new friends online?*
e	*…upload photos to a website?*
f	*…download and save music (MP3s)?*

Table A1: Wording of the self-efficacy items