

Does the Internet Make the World Worse? Depression, Aggression and Polarization in the Social Media Age

Bulletin of Science, Technology & Society
2021, Vol. 41(4) 116–135
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02704676211064567
journals.sagepub.com/home/bst


Christopher J. Ferguson 

Abstract

Since the 1990s, the influence of the internet and social media in daily communication has skyrocketed. This has brought both remarkable opportunities and perceived perils. Recent years have seen increases in suicide and mental health concerns, political polarization, and online aggression. Can such phenomenon be connected causally to communication via social media? This article reviews the evidence for perceived deleterious effects of social media on several areas of human welfare, including political polarization, depression and suicide, aggression, and cyberbullying. In addition to examining contemporary evidence from psychological studies, a historical analysis is included to examine whether we truly live in a uniquely difficult time or whether similar patterns of social behavior can be witnessed in other, pre-internet times. It is concluded that evidence may link social media to some negative social outcomes but in ways that are nuanced and complicated, often interacting with user motivations and personalities and situational variables. An increased focus on preregistered, standardized scientific methods and cautious interpretation of effect sizes can help clarify real versus phantom effects of social media.

Keywords

social media, suicide, screens, preregistration, political psychology, depression

When use of the internet and social media first became widely adopted by the general public in the early 1990s (for the internet) and 2000s (for social media), these were hailed as a technological, social, and economic revolution which some suggested to be on par with the adoption of the automobile (e.g. Stewart, 1996). Such utopian views hailed the internet and social media as having the potential to unleash global democracy, provide opportunities for new interpersonal relationships, improve the ease of purchasing, improve communication, etc. In fairness, before examining a more dystopian view of social media and the internet, it is reasonable to acknowledge that a fair amount of this utopian vision has come true. Social media and the internet have brought with them several advances, most of which we'd likely prefer to maintain even if given the opportunity to roll back time to the 1980s or before.

Nonetheless, alongside these utopian visions of the dawning social media age, warnings of a dystopian future of computer-driven surveillance and control and social isolation also promulgated. Interestingly, these utopian and dystopian visions often existed side-by-side with little cross-pollination of ideas and concerns (Howcroft & Fitzgerald, 1998).

In recent years, concerns about the impact of social media on consumers have arguably increased. These include concerns that, far from social media contributing to an international democratic conversation, social media

has contributed to political polarization and acrimony. Further, the anonymity or distance between individuals on the internet, coupled with *virtue signaling* to ingroups may have increased aggressive social behaviors in internet communications. Recent negative trends related to suicide and mental health have also been tied to the consumption of social media by some scholars and advocates. This raises several related questions. These include:

1. Does use of social media contribute causally to political polarization, aggression and the prominence of fringe political groups and ideologies?
2. Does use of social media contribute to increases in suicide rates or depression seen in some countries?
3. Has social media increased the experiencing of aggression among users?

This paper will consider each of these questions in turn, although the historical analyses will be interwoven into

Stetson University, DeLand, FL, USA

Corresponding Author:

Christopher J. Ferguson, Department of Psychology, Stetson University, 421 N. Woodland Blvd., DeLand, FL 32729, USA.

Email: CJFerguson1111@aol.com

each of the three main categories of concern. As a common thread, one concern is that in many cases researchers are put in a position of needing to make causal attributions based often on limited, correlational evidence.

Political Polarization

Political polarization occurs when individuals within a society split over ideological lines into opposing camps. These camps, in turn, adhere to ideological beliefs defined by membership in the camp, and tend to express acrimony toward or demonize and dehumanize those in the opposing camp (Manbeck et al., 2018; Pew Research Center, 2014a). Thus, polarization extends beyond merely having opposed political parties or viewpoints. Further, not all members of a society need ascribe to either camp for polarization to occur, so long as polarized acrimony becomes a defining feature of the political process. There is a general perception, backed by data (e.g. Pew Research Center, 2014a) that political polarization has increased in recent US history and may be a factor for other democratic republics as well. Such polarization can lead to a breakdown in democratic norms, unwillingness to compromise, anger-based decision making in politics and, at its worst, the potential for political violence. As such, it is a topic worth investigating and exploring why political polarization has increased in recent years.

Social Media and Political Polarization

The rise of political polarization in the US over the past two decades is undoubtedly complex. Factors involved many include the introduction of the 24-hour news cycle, a primary election system that may inadvertently empower more extreme voters (Brady & Pope, 2007), actual political and economic events such as the 9/11 attacks or 2008 economic downturn, etc. But there are concerns that the use of social media could be increasing political polarization.

Concerns about social media stem from several related ideas. The first is that, as opposed to the ubiquitous national news that dominated television and print in earlier decades, social media allows for individuals to select news and opinions that reinforce their previously held views rather than challenge them (Van Alstyne & Brynjolfsson, 2005). Second, is the concern that the internet has removed barriers for the promulgation of extreme (such as radical Islam, white supremacism, “wokism”, etc.), data-deficient (e.g., anti-Vaxxing, climate change denialism, “defund” the police) or otherwise problematic viewpoints. In effect, social media could allow for individuals to self-select only information that reinforces their personal views which could be undesirable particularly when those views are extreme.

This view of social media as “echo chamber” has both received some support and some critique. For instance, one analysis suggested that networks of social media users have created a kind of “fifth estate” of news generation, often

working symbiotically with the “fourth estate” of traditional journalism (Newman et al., 2013). This raises the possibility that news media may become, in effect, subject to the mentality of social media users. This concern became heightened after purges in editorial leadership at the New York Times and other news outlets following protests over perceived racial injustice in the US and elsewhere in 2020 and published editorials or news articles that challenged this social movement. This has raised concerns among some commentators that traditional news media is losing credibility and increasingly influenced by either far-left or far-right (depending on the outlet) moral agendas (see, for instance, Greenberg, 2020). Other analyses, however, suggest that this echo chamber effect may be limited to those who are fairly extreme partisans, with the majority of individuals consuming a diverse array of news sources (Dubois & Blank, 2018).

Thus, evidence suggests that the relationship between social media use and polarization is complex. Optimistically, some evidence suggests that social media use actually increases heterogeneity in social network affiliation although, less promisingly, this does not appear to reduce political polarization (Lee et al., 2014). Thus, although accessing social media may result in exposure to diverse views, the impact on attitudinal change appears to be minimal.

Indeed, one analysis from almost a decade prior provides some evidence for why this must be. Conover et al. (2011) examined the political content and information network of 250,000 tweets leading up to the 2010 US Congressional elections. Their analysis suggested that the flow of information was complex, though highly partisan. Information shared via retweets tended to be highly polarized with little cross-partisan contact. However, information involving mentions (e.g. hashtagging or tagging specific individuals) often drew in cross-partisan contact, though the latter did not appear to substantially reduce partisanship. Thus, it seems Twitter users remain isolated in their political camps only to come out at times en masse in conflicts over hot button issues.

Twitter has been singled out in other research as well. For instance, data suggests that politicians with the most extreme views have the highest number of followers (Hong & Kim, 2016). Twitter followers separated into opposing political camps, suggesting that “echo chamber” theories of social media are likely a better fit to the data than is the notion that social media allows for cross-pollination of diverse ideas. Other studies have likewise, generally supported these findings (e.g., Gruzd and Roy, 2014), and demonstrate this as part of a long-term trend toward increased polarization (Garimella & Weber, 2017). Further, this phenomenon extends outside of the West to Asian countries such as South Korea as well (Lee et al., 2018). This may be particularly true for Twitter given the necessity of reducing complex ideas to pithy statements that maximize likes while reducing nuance in order to fit the character limit.

The incentives for political polarization on Twitter have also been documented. For instance, politicians who express more extreme views tend to receive more campaign donations, particularly by expanding the range of donations beyond their constituencies (Hong, 2013). Twitter is also an effective platform for propaganda, given the proliferation of automated “bot” accounts, which can occupy a significant proportion of the communication traffic on politically charged issues (Stukal et al., 2017).

Questions about the causality of social media use on political polarization have also been difficult to establish. In one recent large and preregistered experimental study, Republican and Democratic respondents were randomized either to follow a Twitter bot that retweeted political messages from the other party’s perspective or to a non-political control. Randomized exposure to opposing political views actually increased polarization, particularly for the Republican participants. In essence, exposure to opposing views backfired and caused participants to become more entrenched in their political views. That these analyses were preregistered gives them particular weight. This also suggests a conundrum, since it is usually expected that interactions between opposing groups generally foster cooperation (Adachi et al., 2016) but this does not appear to work for political polarization. Why this is remains unknown. It may be that social media is ill-suited for the types of exchanges that foster cooperation or that animus between self-identified Republicans and Democrats is so great that contact between them is already too poisoned. Evidence further suggests that this partisanship is, itself, bipartisan, affecting individuals on both the right and left of the political spectrum (Clark et al., 2019; Ditto et al., 2019).

The relationship between social media use and political polarization is complex and nuanced, however, and depends in part upon the individual user. For instance, in another study (Choi & Shin, 2017) mere exposure to social media had no impact on attitudes toward political compromise. However, social media use did interact with two Big-5 personality traits. Specifically, more agreeable and more conscientious individuals were more open to political compromise the more they used social media, with inverse effects for those low in agreeableness and conscientiousness. Or, put more bluntly, less agreeable and less conscientious individuals tend to become most polarized when using social media.

Other research has indicated that individuals who feel relatively deprived compared to others in their community are particularly susceptible to certain kinds of polarization, such as populist messages. In two experiments Hameleers et al. (2018) found that individuals who felt economically and socially disadvantaged relative to the larger community were particularly susceptible to populist messages. Although selective exposure may play a role, this effect was particularly poignant in the context of attitudinal congruence. In effect, populist messages were particularly powerful to the

extent they reinforced the consumer’s preexisting beliefs and grievances.

Thus, from the extant data we can see that social media can play a role in the messaging involved in political polarization. However, it is not as simple as more time spent on social media results in political polarization. Rather, social media can be used by some individuals as a mechanism to reinforce their a priori beliefs, grievances and animosity toward others with differing views. Given that some individuals are both less agreeable about and, perhaps, less concerned with the strict accuracy of their claims (via conscientiousness), this can result in increasingly hyperbolic statements that impede rather than advance progress on important social issues by assigning blame rather than seeking compromise and fixes.

A Very Brief Historical Context for Political Polarization

Observing that social media appears to be a factor in political polarization may lead to assumptions that political polarization is uniquely acute during the social media age. Although social media may, indeed, be one factor contributing to current political polarization, as understood by moral panic theory (Bowman, 2016), it may be tempting to use social media as a scapegoat for larger, more complex, social problems and, in exclusively demonizing social media, fail to comprehensively address those larger issues. For example, within the United States, periods of high discord including threats of civil war marked the early republic’s years in the 1790s and early 1800s (Wood, 2009), then again in the 1850s, leading to the US civil war, then again in the 1880s, Depression era 1930s and the 1960s. All of these periods of marked discord were achieved without the assistance of social media. As such, an understanding of the involvement of social media in current society is aided by a fuller understanding of a natural human propensity for civilizations to at least periodically degenerated into disharmony and strife.

An exploration of history finds that political polarization within democratic or republican polities is nothing new to the current age. Granted, much of history is replete with autocracies, largely solving the issue of polarization through suppression of dissent, though even some of these were not immune to schism. It is important to note that the examples provided in the early part of this section are historical analyses, not empirical data and are selected for their illustrative nature. Alone they by no means capture a full sense of the ebb and flow of polarization in history over time. These are selected mainly to frame the issue as part of a broader historical context.

In the west, the cases of Athens and the Roman Republic are both illustrative as these each constitute variants on proto-democratic or republican political systems that became limited by acrimony and violence within their systems. Neither had a party system in the modern sense, though

various political alliances and factions would certainly rise and fall.

The case of Athens as a direct democracy of enfranchised males (which made up a minority of the population of the city-state) has often been considered a warning of the limits of direct democracy. Far from being an idyllic center of cooperation, Athenian democracy could be brutal and cruel, with contemporary observers such as Thucydides (see Thucydides, 1934) noting that decisions were often made unwisely or emotionally. At the extreme, deaths could occur such as the execution of Socrates or the execution of several Athenian naval commanders after *winning* a battle against the Spartans in 406 BCE. More common were systems like *ostracism* wherein political leaders would attempt to have their opponents voted into 10-year exile. Use of ostracism is detailed in Plutarch's *Lives* and it sometimes backfired, leading to decline in its use.

Rome's republic collapsed into empire under the weight of multiple factors, including class struggles, the increasing power of the army, and political corruption. Although a representative republic in form, it tended to be oligarchical in practice, with a gradual tendency toward highly Machiavellian political intrigues that sometimes spilled over into violence, such as the murders of the Gracchus brothers and, later, Julius Caesar (Everitt, 2013). Historical parallels are, by nature, both selective and fraught, though the decline of Rome under the weight of class polarization, corruption and military power points to some factors that could be empirically evaluated as stressors in the modern age (Ferguson, 2020a).

Indeed, my choice to highlight the Roman Republic is a particularly intentioned one for the present crisis facing the US and other republics. Other scholars have, likewise, addressed the historical parallels between the decline of the Roman Republic into Empire and the current circumstances within the United States (Hammer, 2020). Specifically, a decline in trust in public institutions, coupled with the promise of authoritarians to institute concrete change can result in the erosion of personal freedoms, erosions which can come from either side of the political spectrum. This could occur in earlier civilizations, such as Athens and Rome, which lacked the infrastructure and design of modern politics, though the tendency for partisanship and, indeed, sectarianism may arguably be aided by the emergence of modern political parties, whatever other virtues these may have.

Prototypical political parties would emerge from time to time, such as the Greens and Blues of the early Byzantine Empire. Centered around chariot racing teams, these factions also reflected the same class struggles that damaged the integrity of the Roman Republic and ultimately led to violent clashes causing the deaths of thousands (Brownworth, 2010). Modern political parties began to emerge in England in the 19th century. These had a false start with the supporters of Parliament (Roundheads) and Monarchy

(Cavaliers) during the English Civil war of 1642–51. Parliamentary forces won, though ironically shepherded in an era of military dictatorship under Cromwell. Following restoration of the monarchy after Cromwell, the Whig and Tory parties emerged, now split over more liberal/reform or conservative ideals at least as much as over class (Fraser, 2005).

It is worth asking how the concept of polarization applies in the sense of comparing historical cultures, whether Athens, Rome, Byzantium or, indeed, 19th Century England to the current turmoil of the present age, whether Trumpists versus the “woke” in the US, Leave vs Remain on Brexit in the UK, etc. However, the argument has been made that most polarization ultimately relates to class struggles (Lind, 2020), however they may otherwise ostensibly focus on other issues such as race, gender, the economy, culture wars, etc. In this sense the class struggles between the Plebeians and Patricians in Rome have considerable similarities to those of Leave vs Remain in the UK, or the “woke” coastal elites vs the “Trumpian” flyover country in the United States. Social media then becomes a new battleground for members of these classes to engage in moralistic class conflict, marked by acrimony and moral grandstanding rather than persuasion and compromise.

This distinction of more liberal versus more conservative political parties ultimately became a common feature of the American political party system. As with the English system, partisanship and violence have roughly cycled in American history. Indeed, evidence suggests that political polarization is particularly high historically among Anglophone countries (e.g. The US, UK, Canada, Australia; Smith & Mayer, 2019). The American Revolution itself is sometimes viewed as a quasi-civil war between revolutionaries and Tories, with more brutal violence committed between factions of Americans than that involving British soldiers (Allen, 2011). Most politicized violence in the United States has been between groups of individuals, such as on issues of class and race (such as considerable racial violence toward African Americans in the 20th century), but leaving political structures intact (Graham, 1970). The most obvious exception to this is the US Civil War from 1861–65, which was fought between both geographic and political factions over the issue of slavery. However, the political chaos and lower-level violence (such as the emergence of mainly left-wing domestic terror groups) of the late 1960s and early 1970s arguably also qualify as significant anti-governmental violence.

By the 1950s–1980s, polarization had decreased, and the major political parties of Republicans and Democrats had relatively broad ideological bases, allowing for compromise across party lines. Some analyses suggest this period of reduced polarization was more unusual than typical, with significant polarization more the norm in US politics (Han & Brady, 2007). The Pew Research Center (2014a) has tracked political polarization in the United States from

1994–2014. Their data confirms that political polarization has increased markedly during this time period. Those identifying as Democrats became significantly more liberal in views, whereas those identifying as Republicans became significantly more conservative. Ideological conformity and siloing (socializing only with those with similar views) also have increased. Political animosity has also increased with significant minorities in each party (27% of Democrats and 36% of Republicans) viewing the other party as a *threat* to national well-being. The data also show an interesting urban/rural divide with more liberal views among those preferring urban environments. The Pew data note that most Americans don't belong to either party extreme, but that those with more polarized views are more active in politics, often driving politics toward the extremes. It's important that this Pew data date back to 2014, before the era of President Trump. Updated graphical data from 2017 (Pew Research Center, 2017) suggest that, perhaps not unexpectedly, Democrats have moved significantly to the political left¹ since 2014. This polarization has brought with it new concerns of politicized violence, particularly white nationalist groups on the right, but also violence by left-leaning groups such as Antifa or from non-white racial groups, as well as domestic terrorism due to radical Islamist ideology.²

Data from other sources has confirmed this increasing trend in polarization of political views in the US in recent decades (Baldassarri & Gelman, 2008). Given that these trends in polarization began *before* the widespread availability of social media though concurrent with the spread of the internet more broadly, it is difficult to specifically pin down whether increased technology use plays a causal role in this process. Other analyses suggest factors such as trust in social structures and income inequality tend to be robust predictors of political polarization (Grechyna, 2016). Thus, it may be more appropriate to consider social media as having an amplifying rather than causal role in political polarization.

From this brief historical overview, we can see several themes, namely;

1. Political polarization and subsequent violence have plagued democratic and republican polities throughout history.
2. Political polarization often forms along class lines as well as extremes of conservatism and progressivism³.
3. The current polarization of politics in the US (and other nations) reflects a recent increase but has not (yet) become as critical as other points in history, such as the US Civil War.
4. Political polarization appears to cycle in waves, particularly in US history. However, some rough analogies suggest that polarization can become critical once violence is seen as an acceptable solution, norms of compromise and political process are weakened,

corruption is common, and the military is politically empowered.

We can thus place our current circumstances into a historical perspective. Current political polarization is bad relative to recent history, but not as bad as some other points in more distant history, such as the US Civil War or the collapse of the Roman Republic. Critical political polarization does not require social media to happen and, in fact, is rather common historically. But this does not mean that social media plays no role in the current political environment.

Political Politization: Conclusions

This review of the evidence regarding social media and political polarization leads to several conclusions.

1. First, social media and the internet have not produced a historically unprecedented period of political polarization. Aggressive political polarization appears to be a historical risk of politics itself.
2. The causal relationship between polarization and social media use is nuanced. Social media may have an amplification-like effect for individuals who already tend to be polarized or who feel disadvantaged in society.
3. A focus on semantics, particularly historical comparisons (particularly to the Nazis or other historical lightning rods) rather than facts of a specific case tends to increase polarization rather than stimulating meaningful and constructive debate. The *thoughtful comparison* loophole to adages such as Godwin's Law or *reductio ad Hitlerum* should be eliminated as these likely only increase self-perceived righteousness for fallacious comparisons, particular if boosted by *argument to authority* input from the adage creators. Such appeals may do little to convince, however. Critical evaluation of immediate political crises may be slowed rather than speeded by historical comparison, though more research on this would be welcome.
4. Though there is some evidence to suggest social media and political polarization are associated, stronger evidence, particularly from preregistered studies would be highly desirable.

Mental Health

Even as political and social polarization has increased, people appear to have become united in their disdain for and fear of big tech platforms such as Facebook and Instagram. In September 2021 the Wall Street Journal released a report accusing Facebook of having conducted studies linking Instagram use to mental health problems in teen girls. However, many scholars (e.g., Ritchie, 2021) soon pointed out that the studies in question were of poor quality, Facebook having asked girls what they *thought*

Instagram did to them rather than measuring clinical outcomes in any reliable way (people frequently misattribute the cause of their own issues, particularly during a moral panic, Bowman, 2016). Nonetheless, the studies were Facebook's self-inflicted wound, and fit a neat narrative of an evil company hiding damaging research much like the cigarette industry had done before. However, what does high quality research say about social media and the internet and mental health?

A Brief History of Mental Illness

Obviously, a comprehensive history of mental illness would itself take volumes to consider. Examining historical trends in mental illness is difficult given general lack of care for the mentally ill across most historical societies, lack of common nomenclature, and a focus on more extreme disorders such as psychosis, rather than stress, anxiety and depression (Ferguson, 2020a). Treatment of the mentally ill has historically been brutal and remains controversial into the modern era (e.g. Whitaker, 2002). Clear data on the prevalence of mental illness in pre-modern historical eras is entirely lacking. However, anecdotal evidence suggest that mental disorders were far from uncommon. In modern times, data suggests a generally stable level of mental illness, barring modest rises and falls, with claims of epidemics largely exaggerated and due to methodological shortcomings (McMartin et al., 2014).

The Social Media/Mental Wellness Debate

In recent years, a ferocious debate has erupted regarding the impact of social media and related screens on the mental health of youth. Arguably, this debate has begun to eclipse the similarly acrimonious debate over media violence (see Markey et al., 2015). At the very extreme, this debate has focused on whether social media is a main causal explanation for the rise in suicide rates seen in teens since the late 2000s (beginning approximately 2008–2009). On one side, psychologist Jean Twenge and colleagues have argued for such a link (Twenge & Campbell, 2019; Twenge et al., 2019). These scholars don't argue that social media is the sole or even most critical cause of teen suicide, of course, but rather than increased social media and screen use can explain increases in teen suicide over the last decade by adding a new risk factor to those already well-established for teens. By contrast, psychologists Amy Orben and Andrew Przybylski argue that the data does not support evidence for even a correlational link and these concerns are more akin to moral panic than good science (Orben & Przybylski, 2019a, 2019b). This dispute involves several features with wider implications for the conduct of psychological science more broadly.

I consider the evidence from two perspectives. The first is that of "general effects" or the belief that simple time spent

on social media can be directly quantified as a predictor for negative mental health outcomes. From such a perspective, we'd expect to see direct correlations or odds ratios linking social media use to negative mental health simply as a matter of time spent using. Second, from a nuanced effects perspective, I examine evidence that social media use is contextual, such that it may be good for some, but bad for others, or that outcomes may depend on how one uses social media with idiosyncratic effects more likely than general ones.

General Effects. From a review of the literature there are perhaps fifty or so studies that directly examine the impact of social media use on mental wellness related outcomes. It should be noted that these do not include papers that specifically examine for pathological use (or what some call "addiction" though that term may problematically sound akin to substance abuse to the general public), but rather correlations between time spent using and negative outcomes. Although not universally the case, most of the studies and much of the narrative focus specifically on suicides among teens (e.g. Twenge et al., 2019). As noted above, it is this body of literature that has been more controversial, rather than that on nuances. Thus, the question remains, whether social media is one cause of pressing public health issues such as youth suicides.

Much of the debate has focused on differing interpretations of fairly small effect sizes from a small number of large datasets. Information from outside these datasets also has been equivocal. For instance, in some cases, correlations between depression and social media use are found, although it is not clear whether social media use caused depression or depressed individuals turned to social media as a coping mechanism (e.g. Lin et al., 2016).

Perhaps most infamous was a study that suggested that changing the positive or negative valence on Facebook feeds could influence users' moods (Kramer et al., 2014). Using a massive sample of nearly 700,000 participants, the authors manipulated users' news feed with more positive or negative content, then measured the mood valence of keyed words the users typed. Participants were not informed they were in a study; indeed, the journals' editors published an expression of concern noting the ethical issues. However, the study results are a classic example of why focus on "statistical significance" and ignorance of the importance of filtering out weak effect sizes (in this case $d = .02$) damages social science and misinforms the public. The study was often presented as evidence that Facebook could influence mood, but mood was never measured, merely inferred from keyed words. Further, the effect size was so small, it would only change one word out of thousands. As one psychologist wrote, reviewing this study (Grohol, 2014) "This isn't an 'effect' so much as a statistical blip that has no real-world meaning." The authors of the original study defended interpreting such a tiny effect size by claiming aggregate effects over time could be meaningful, but this is arguably a self-serving rationalization and is certainly not based in any

data. Once again, this study highlights the danger of interpreting tiny but “statistically significant” findings that are as likely to be methodological noise as they are true effects. Such generous interpretations arguably misinform the public considerably about psychological processes.

Much of the data to suggest links between social media use and mental health problems comes from large, publicly available datasets such as Monitoring the Future or the Youth Risk Behavior Surveillance System (e.g. Twenge & Campbell, 2019; Twenge et al., 2018). These typically focus on youth or young adults as opposed to older adults. Often employing tens or even hundreds of thousands of participants, such datasets are particularly sensitive in picking up even very small effects. Others (e.g. Twenge et al., 2019) have examined suicide trends in national data, inferring that recent increases in youth suicide may be associated with cultural increases in the use of social media beginning around roughly 2009.

Often these datasets are analyzed and reanalyzed by differing scholarly groups coming to different conclusions, often based on differing interpretations of effect size and statistical significance. For instance, the effect sizes linking social media to mental health outcomes has been simultaneously compared to potatoes and eyeglasses’ impact on mental health (Orben & Przybylski, 2019a) and, opposingly, heroin and other hard drugs (Twenge et al., 2020). Thus, whatever other controversies that exist in this field, it seems uncontroversial to suggest different researchers look at the same data and come to vastly different conclusions.

It should be noted that most of this debate focuses specifically on girls, not boys. Interestingly, data from these datasets suggest that most of the positive findings (whether they are large enough to be interpreted as meaningful or not) are for teen girls only. Why this would be is somewhat mysterious at present. It could be that teen girls use social media in ways that result in, for instance, more appearance-related social comparison. Conversely, it may be that girls who are already depressed are more likely than boys to turn to social media for peer support. Or it may be that such small effect sizes are some kind of statistical artifact. Particularly based upon cross-sectional data, it can be difficult to interpret these correlations, likely leading to the significant debate we are seeing.

Another major issue is the degree to which reported effects are reliable and reproducible. Particularly for studies that are not preregistered, questionable researcher practices (QRPs) may create unreliable results (Simmons et al., 2011). With big datasets, this may involve what can be called *r*-hacking, or efforts (intentional or unintentional) to boost effect sizes with QRPs, thus making effect sizes meaningless (Orben & Przybylski, 2019a). As such, preregistered studies may be particularly valuable in reducing the risk of artificially inflated effect sizes.

However, substantial long-term effects have been difficult to establish. For instance, outside the datasets discussed

earlier, several longitudinal analyses (e.g. Heffner et al., 2019; Jensen et al., 2019) have not found that social media or screen use more generally is associated with increased mental health problems. Others have suggested a kind of “Goldilocks Effect” with best outcomes for moderate users, with worse outcomes for both heavy and non-users (Przybylski & Weinstein, 2017). However, even these quadratic relationships are so small they may be of limited clinical significance, an interpretation advanced by the authors themselves. It appears to be that, controlling for other theoretically relevant predictors such as gender, baseline mental health, or family environment, largely eliminates the predictive value of social media or screen use as a predictor of youth outcomes (Ferguson, 2017).

Other sources of data, such as time-use diary-based data likewise suggest that social media and other screens have limited direct impact on adolescent well-being (Orben & Przybylski, 2019b). Other evidence (Orben et al., 2019) suggests that some positive findings in past literature may have been spurious due to methodological choices. Quantitative reviews likewise find that relationships between social media use and wellbeing are near zero in effect size (Best et al., 2014; Huang, 2017). Others, however, strenuously debate such conclusions, suggesting that these small effect sizes are similar in magnitude to other important effects and should not be ignored (Twenge et al., 2020).

Most of the extant evidence is, of course, correlational in nature. Even longitudinal studies do not necessarily rule out third variables, though it is worth noting that longitudinal studies appear to present even less evidence for correlation than do cross-sectional studies. Thus, without a wider body of experimental work, we are also left without the ability to gain clarity on causal implications. Even if we accept that the cross-sectional effect sizes are meaningful, this does not address for us whether use of social media leads to mental health concerns or the inverse.

It appears likely that this debate, often acrimonious, is likely to continue into the near future. Given the magnitude of effects and concerns that methodological problems (to be discussed below) may be as likely to cause these as are real effects in the population, it is probably best not to consider social media use as a risk factor of youth depression in a direct sense. However, this does not rule out the potential for more nuanced effects. To this issue of more nuanced effects, we’ll now turn.

Nuanced Effects. In contrast to claims for general effects of social media on youth well-being, evidence has accumulated that social media may have more nuanced, idiosyncratic effects. Specifically, rather than raw exposure or time spent using predicting outcomes, *how* one uses social media may be more relevant. This, of course, places the user him or herself more squarely in the controlling role of how social media and mental health related.

Certain types of social media behaviors are associated with positive mental health. For instance, evidence suggests

that online authenticity, meaning the presentation of true positive aspects of one's own life online is associated with positive well-being (e.g. Reinecke & Trepte, 2014). Positive and entertaining self-disclosures can increase feelings of connectedness among colleagues on social media (Utz, 2015). Similarly, in a Chinese sample, positive self-disclosures among associates were associated with increased well-being among young adults (Wang et al., 2014). The key issue appears to be, across studies, that *positive* disclosures appear to be associated with positive well-being and mental health (Chan, 2015; Grieve & Watkinson, 2016; Qiu et al., 2012).

Naturally, if positive interactions and disclosures predict positive outcomes, negative interactions and disclosures predict more negative outcomes. For instance, negative interactions with other over social media tend to increase depressive rumination (Davila et al., 2012). In another study (Park et al., 2016) negative disclosures on Facebook were associated with greater depression, although perceived social support reduced this depression. Other data (Berryman et al., 2018) likewise find that negative statements such as *vaguebooking* (posting negative but vague thoughts so as to get attention) is associated with decreased well-being, though overall time spent on social media was not. Negative comparisons with others on social media can likewise reduce well-being (Krasnova et al., 2015). Likewise, other data has suggested nuanced relationships between social media use and clinical disorders (Rosen et al., 2013). For instance, having more Facebook friends was associated with less depression, whereas a preference for multitasking was associated with negative outcomes.

Thus, evidence suggests that interactions between social media use and mental health are nuanced with the potential for both positive and negative outcomes dependent upon how the user interacts on social media. As such, rather than preaching abstinence or reductions in use, it may wiser for clinicians to focus on healthy use and maintaining social media use in balance with other life activities.

Interpretation of Tiny Effect Sizes. As noted, one main dispute has focused on the interpretation of several large datasets of youth, each with tens or hundreds of thousands of participants. With such datasets, small correlations, typically below $r = .10$ in effect size, emerge as “statistically significant” though they may either have little practical/clinical value, or may simply be false positives brought about by methodological issues, such as single-responder bias, demand characteristics or common methods variance. The uncertainty about what to do about such small effects is by no means limited to this field but is the source of significant disagreements.

Twenge and Campbell (2019) defend effect sizes below $r = .10$ seen in such datasets for social media on depression by suggesting that using the r^2 metric for interpretation of practical significance is flawed. They suggest that some important medical effects have very low r^2 values. These comparisons

had become common for years (e.g. Rosnow & Rosenthal, 2003). However, these medical effect size comparisons have been largely discredited as based on flawed statistics (e.g. Block & Crain, 2007; Ferguson, 2009). The error came from using medical effects tables that conflated the prevalence of disease with the effectiveness of the treatment for that disease. In large samples, this produces a spuriously truncated effect size estimate that bears no relationship to the true population effect size for the treatment effectiveness. In other words, the effect sizes reported by some psychologists for medical effects are little more than statistical gibberish. For instance, the effect size for aspirin as a preventative treatment for heart disease based on the large Physician's Aspirin Study, is often miscalculated by psychologists as $r = .03$, by comparison with which many psychological effect sizes appear quite robust. However, once controlling for the prevalence of heart disease in the population, the actual effectiveness of aspirin is closer to $r = .52$ (Ferguson, 2009). Thus, comparisons to medical effects should be not used in defending weak effect sizes.

It is worth noting that arguments against r^2 in some cases may rely on data that strips out control variables or other considerations for which r^2 may have been used in the first place. Using uncontrolled raw data to repudiate r^2 runs the risk of relying too heavily on what are, in effect, bivariate correlations and, in doing so, promoting ecological fallacies (discussed below). Such graphs typically do not control for baseline mental health issues, also a serious limitation. Even with this in mind, not all datasets show such a relationship. For instance, using the data from Berryman et al. (2018), and dividing hours spent online with friends into rough quartiles, with the presence or absence of suicidal thoughts, no differences were found among high and low users of social media ($\chi^2 [n = 467] = 1.22, p = .748$). These data are represented in Table 1.

Another dataset also highlights the risks of overinterpreting small correlations. Using a representative sample of Florida youth responding to the CDC's Youth Risk Behavior Survey, Ferguson (2017) found a small correlation between screen time and depression (bivariate $r = .093$) which was statistically significant. Although smaller in magnitude, significant correlations were also found for eating potatoes ($r = .039$, and thus replicating Orben and Przybylski, 2019a, see below) and carrots ($r = .040$)⁴ on depression. Naturally, we don't warn parents about the dangers of their children eating potatoes or carrots, but these findings highlight the risks of false positive “statistically significant” results in large datasets ($n = 6089$ in this case).

Looking at the breakdown of depression prevalence (here defined as the highest quartile of responses to the depression items) we see the breakdown of increased prevalence in higher screen time quartiles as expected by advocates for causal effects. However, the differences are slight, as to be expected for such a small effect size⁵. Creating quartiles for

Table 1. Suicidal Thoughts and Social Media/Screen Time Quartile Breakdown.

	1 st (Lowest) Quartile	2 nd Quartile	3 rd Quartile	4 th Quartile
% express suicidal thoughts ^a	15.8%	11.2%	13.5%	10.3%
% highest 1/4 depression ^b	6.0%	5.7%	7.2%	9.8%
	0 Potato	3 Potato	4 Potato More	
% highest 1/4 depression ^b	7.1%	6.3%	8.3%	

Note: a = Berryman et al., 2018; b = Ferguson (2017).

potato consumption was not possible as the vast majority of kids had either eaten potatoes either zero times in the past week, or from 1–3 times in the past week. Thus, three categories were created for 0, 1–3, and 3 + potatoes. As can be seen, depression prevalence was slightly higher in the frequent potato group (indeed, only slightly lower than in the high screen time group). These differences are all very small, as would be warranted given the tiny effect sizes. However, we do not warn parents of the dangers of potatoes. By contrast, as a protective factor, eating breakfast is associated with reduced depression ($r = -.136$) with an effect size larger than either screen time or potatoes. It is not unreasonable to suggest that breakfast eating is part of a healthy lifestyle, yet once again we seldom highlight breakfast as a major preventative factor for depression⁶. Ultimately, research from large datasets related to aggression demonstrate that social science research simply lacks the precision at effect sizes below $r = .10$ to distinguish between noise and true effects. As such, such small effects should not be interpreted as hypothesis supportive (Ferguson & Heene, 2021).

These findings may help to illustrate two elements that are often lost in the acrimonious debate over screen time. The outcome for potatoes warns us that sometimes “statistically significant” effects in large datasets are simply nonsense noise, and we should be resistant to interpreting tiny effects as discussed below. Second, if we ignore the issue of tiny effects, it may be more reasonable to view overuse of social media as a tiny element of overall lifestyle issues, with the choices an individual makes on a grander scale more critical by far than mere exposure to social media. In this sense, the tiny relationships between social media and well-being seen in these datasets may be akin to a grain of sand on a beach of minuscule indicators of larger lifestyle issues. Further, the direction of effects remains unclear. Indeed, individuals who are struggling may simply turn to social media (and books, Ferguson, 2014) to cope with their struggles.

Ecological Fallacy Issues. The argument for an impact of social media on mental health is based on the assumption that social media use increased around the years 2008–2009, timed with an increase in teen suicide and depression. However, even if we allow that these two phenomena co-occurred contemporaneously, it is unclear this is anything other than a chance occurrence. For instance, many odd but strong correlations such as that between Nicholas Cage

movies and US swimming pool deaths each year exist, despite the implausibility of their being any causal connection (Vigen, 2015).

Most studies rely upon survey-based responses showing (or not showing) correlations between social media use and either concurrent or longitudinal mental health symptoms. However, the broader narrative regarding social media and mental health typically relies upon noting an increase in suicide, particularly in girls, in recent years and linking this to the perceived introduction of social media into society (e.g. Twenge et al., 2018). Such arguments rely upon the perception that two things have occurred simultaneously in society without necessarily requiring individual-level evidence.

Time series analyses, which examine data points occurring at the society level, rather than individual level, can be of value. However, they are most valuable when they control for other important variables that could explain any perceived cooccurrence, as well as employ proper statistical controls for autocorrelations (essentially spurious correlations between two data series). Even then, time series analyses can be fraught, as evidence by recent data on the television show *13 Reasons Why* which portrays a teen girl’s suicide. Once thought to be linked to real-life suicides, a careful time series analysis ultimately demonstrated this was not the case (Romer, 2020).

At present, though some narratives discuss the rise in teen girl suicide beginning in the 2010s as being linked to social media, little data has considered this from the perspective of well-controlled time series analyses controlling for other trends in society, not least of which, a simultaneous rise in suicides among middle-aged adults. The closest study thus far examined trends in mood disorder symptoms and suicide related outcomes (e.g. Twenge et al., 2019). Although the authors suggest a link with social media use, this study, in fact, provides no data on social media use, so this implication is entirely speculative.

Although teen suicides have increased in the years 2008–2017, this is not a phenomenon isolated to teens. Indeed, the highest suicide rates and greatest increases in suicide rates, according to Centers for Disease Control (CDC) data are amongst middle aged adults (i.e. 45–54). Figure 1 presents this data using CDC WISQARs data. This observation suggests that suicide is a greater issue among relatively lower tech-adopting middle-aged adults than it is among teens.

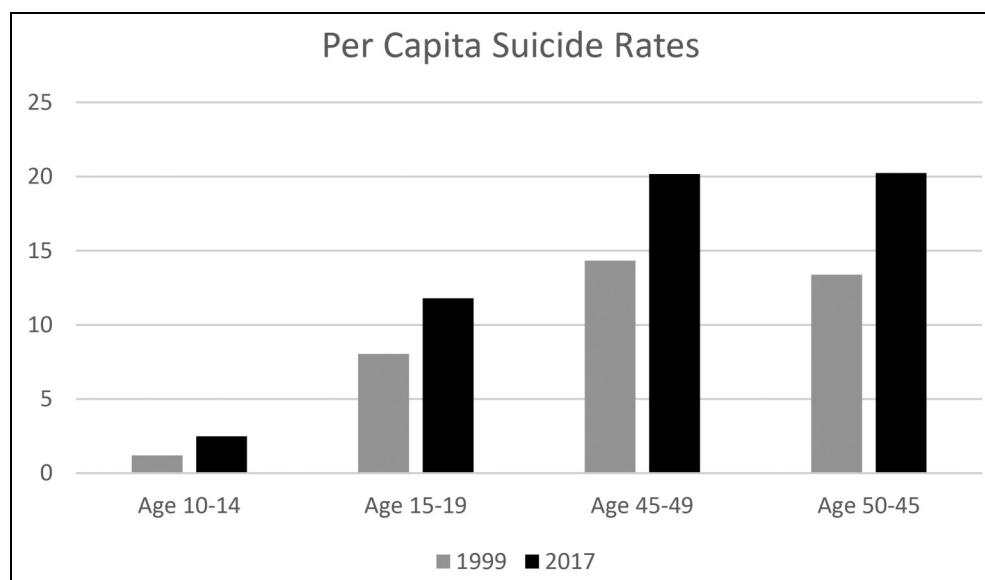


Figure 1. Per capita suicide rates in 1999 and 2017 across age categories.

This is one piece of data to suggest social media use is unlikely a major factor in increased suicide prevalence and suicide among teens may be better conceptualized as part of a larger family and social strain effecting most age categories.

Concerns about youth suicide and social media also tend to focus on exclusively US data, which tend to show an increase in teen suicide (as well as adult suicide) since around 2008. However, other high-tech adopting countries do not show the same pattern. Figures 2 and 3 present time-series suicide data per capita (1000 individuals) for the US, UK, Japan, Republic of Korea, Australia and the Czech Republic for males and females. This is a fairly random assortment of countries, though meant to represent a diverse group of industrialized nations. That data was taken from the World Health Organization's CoDQL database on mortality with the exception of the US data which were taken from the CDC's WISQARs. For boys, the US shows the expected pattern of increased suicide since 2008, although numbers were also higher at the turn of the millennium, indicating a sine-wave pattern. Asian countries show a declining suicide trend in boys. The UK remained stable aside from a 1-year increase from 2014–15. Both Australia and the Czech Republic demonstrate volatile patterns that are hard to pin down to a trend, although current per capita suicide numbers are lower than in the late 1990s. Patterns are fairly similar for girls, with the exception that the Czech Republic is also showing an increase in suicide rates. The UK has remained largely stable, with Australia showing a volatile pattern, and observable declines in the Republic of Korea and Japan. These data suggest that focusing on trends on a single country (the USA) may have inadvertently misidentified a trend that is spurious and not

consistent cross-nationally. Of course, one might reasonably suggest that different cultural issues will affect different countries' suicide trends heterogeneously, but of course this argument cuts both ways and still warrants caution in the interpretation of US suicide trends. Even for US data, the argument linking suicide to social media requires relying on 2008 as a sort of magic date. It is not clear there is any empirical reason to set this date as a critical one for examining these trends other than it is convenient for arguing for links between suicide and social media. Although penetration of social media certainly increases over time (Kolmes, 2012) this appears to be a linear rather than binary process with no critical date.

Mental Health: Conclusions

1. The debate over social media and youth suicide and depression is currently very fierce. At present, evidence for a causal relationship remains weak. Most arguments rest on an observed correlation between increased US teen suicides in the years 2008–2017 with increased use of social media. However, data from other countries does not match this trend and data from survey studies is weak and potentially better explained as methodological noise rather than true effects.
2. Most evidence to date relies on cross-sectional self-report designs, albeit in some cases with very large sample sizes. However, such studies preclude conclusions of causality or even temporal order.
3. Time spent on social media is not a good predictor of mental health outcomes.

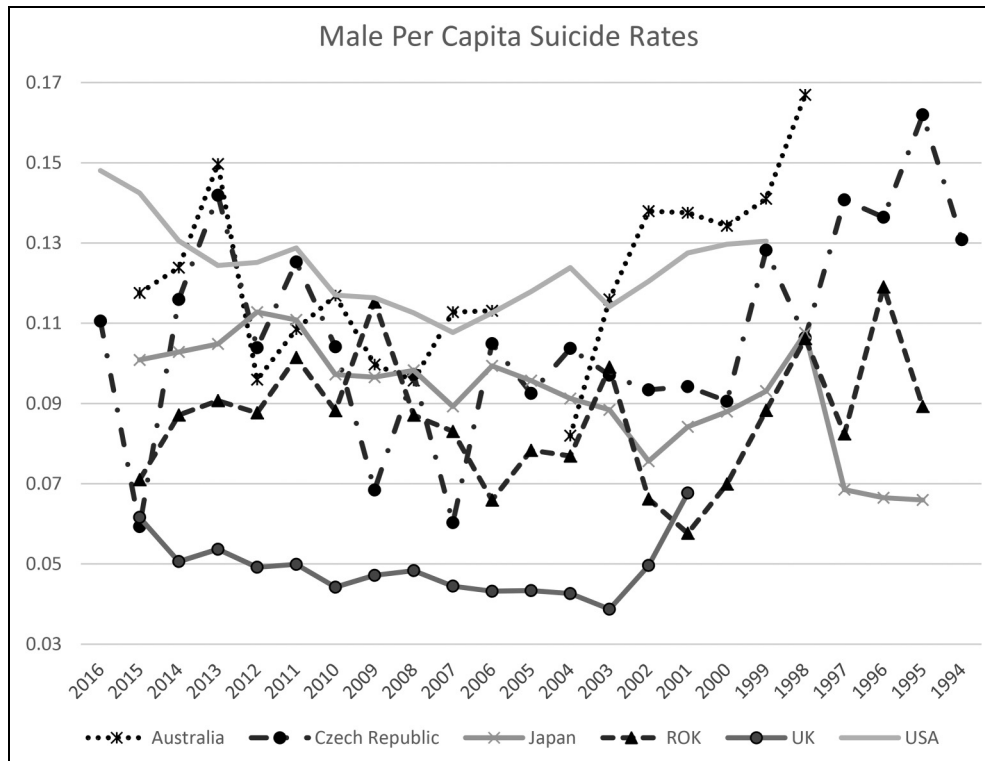


Figure 2. Cross-national per capita teen (15–19) suicide data for males.

4. How one uses social media does appear to predict mental health outcomes. Usage focused on positive disclosures with associates predicts positive outcomes whereas negative disclosures and interactions predict negative outcomes.
5. Measurement of social media use is difficult, often reliant on self-report and difficult to track over time as platforms come and go. For instance, the years after 2004 saw increased use of major platforms such as Facebook and Instagram, although focusing on these discounts the use of chatrooms and discussion boards and other platforms prior to this time.
6. This research field is in dire need of open, preregistered research designs. This may reduce the potential for r-hacking and spuriously high effect sizes. Further, data on usage from technology companies would also be welcome.

Aggression

As such, we see that the links between social media use and mental health are less clear and more nuanced than many have assumed. However, people often have concerns about harassment and other acts of aggression online. A common criticism of social media is that interactions on social media are often aggressive, sometimes to extremes such as doxxing (revealing personal information so that an individual might be harassed) and swatting (calling in false police emergencies to someone's address as a dangerous prank.) There's

little question that experiencing aggression online is common, but is social media actually causing aggression or simply removing filters for our preexisting aggression?

Aggression is typically defined as intentional behavior that causes harm to another that this other person wishes to avoid (Baron & Richardson, 1994). This of course encompasses a wide range of behavior from criminal violence to behaviors society sanctions such as sports aggression, defending oneself, or even rancorous debate. Regarding social media, it appears reasonable to expect that the most pronounced effects would regard social aggression such as bullying, harassment, ad hominem (including racist and misogynistic statements) and public shaming. Some commentators (e.g. Ronson, 2015) have suggested that social media has ushered in a new era of public shaming reminiscent of the use of such punishments in the 16th-18th century. It is possible, of course, that heightened aggression in the social sphere could spill over into physical aggression in real life as social and physical aggression tend to be highly correlated (Slawinski et al., 2019). I'll begin with a brief examination of this historical evidence on aggression as this may help us to understand how aggression works in the modern social network.

A Brief History on Aggression

Available archaeological evidence suggests that serious physical aggression was endemic in prehistorical times (McCall & Shields, 2008). Evidence likewise suggests that

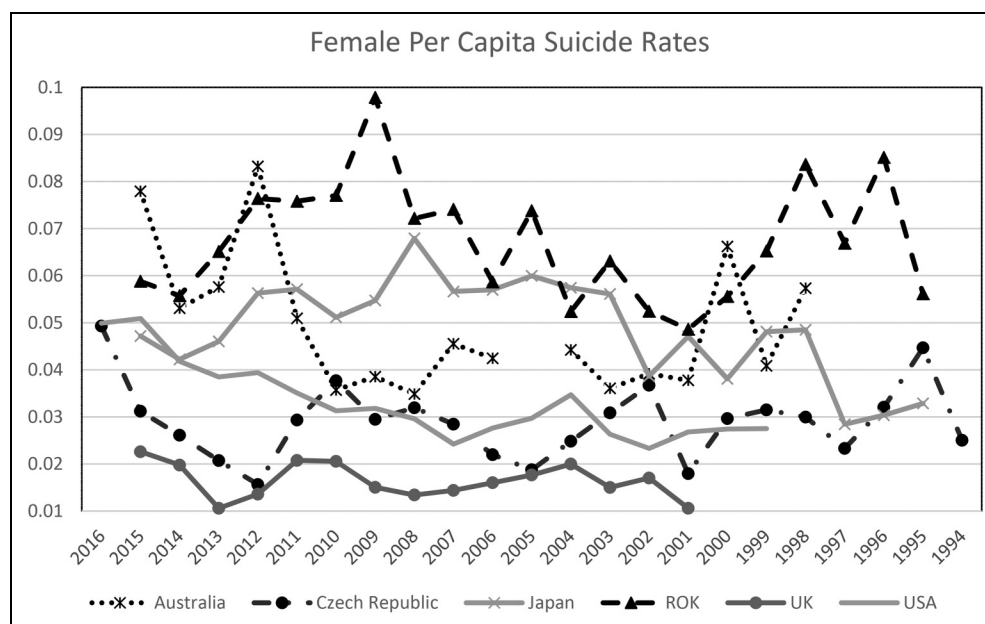


Figure 3. Cross-national per capita teen (15–19) suicide data for females.

serious physical aggression is largely an evolutionary trait which may be exacerbated in situations of childhood neglect and abuse (Boutwell et al., 2011). However, the last century of so, despite two vicious world wars, has seen an overall decrease in serious physical violence, with gradually fewer war deaths and violent crimes, including homicide, assault and rape. Indeed, as of the writing of this manuscript, we are likely living in the least violent epoch of human history (Pinker, 2011).

Why this has occurred is a subject of intense debate, and probably has multiple explanations. One of these has focused on deterrence, or the belief that engaging in aggression will be met with negative consequences, thus reducing incentives to engage in aggressive behavior. The impact of deterrence on aggression has been understood for decades (e.g. Michener & Cohen, 1973). Discourse on the internet, however, may eliminate deterrence through anonymity or distance (and, thus, reduced threat of consequences for aggressiveness). Other forms of aggression, such as bullying, appear to have decreased in recent years (Waasdorp et al., 2017). Thus, the role of social media in aggression must be viewed with an eye on the extremely high prevalence of aggression through most of human history, coupled with significant declines in recent years despite increased use of social media.

Experiencing Aggression Online

Despite overall decreases in bullying and aggression, the experiencing of aggression online is fairly common. For instance, youth experiencing of online harassment increased from 6% in 2000 to 11% in 2010 (Jones et al., 2013). Data from the Pew Research Center (2014b) finds that online harassment is

common with 27% of respondents reporting experiencing at least mild harassment. Harassment was a common experience for both men and women, though women experienced more harassment of a sexual nature than men.

It is important to note that while online harassment and cyberbullying are highly prevalent, the impacts of specific incidents of online aggression may be less than for aggression experienced in real life. In one recent large study of bullying, for instance (Przybylski & Bowes, 2017), youth commonly reported both real-life bullying and cyberbullying. Further, both were associated with greater mental health symptoms. However, real-life bullying accounted for far more variance in mental health symptoms (approximately 5%) than did cyberbullying (approximately 1%). Nonetheless, meta-analysis suggests that cyberbullying is fairly common (most prevalent estimates hover between 10–20% victimization rates), tend to overlap with real-life experiences with bullying and are associated with negative mental health outcomes (Kowalski et al., 2014).

One interesting facet of online harassment is that it is often less motivated by ideological differences than sometimes suspected. For instance, regarding the experiences of women with sexual harassment, it is often assumed that such harassment reflects hostile sexism among the harassers. However, some research indicates that harassment, aside from when occurring in specific disputes, is often random. Sexual harassment of women is associated with general harassment of men and women and appears more often motivated by sadistic amusement than any adherence to sexist gender ideology (Paananen & Reichl, 2019). This seems a crucial point in understanding online aggression. Often such aggression is used to claim larger points about sexism

or racism in society more broadly, but it may be more appropriate to understand such behaviors more as random trolling with the aggression serving more of an amusement than ideologically enforcing motive. However, some other research has found hostile sexism to be a predictor of aggression in online gaming contexts (Tang & Fox, 2016). The role of such personality factors may depend on the lens through which we look. Hostile sexism may increase the risk of harassment perpetration, but not all (or most) online sexual harassment may be perpetrated by ideological sexists. It is fair to note, as well, that a considerable amount of aggressive online content comes not from humans at all, but social bots designed to inflame emotions (Stella et al., 2018).

Experience of harassment is also linked to users' own behavior, although it is important to state that this does not mean they deserve harassment. For instance, among teens, posting public pictures or personal information, flirting with strangers online, and, curiously, having a part-time job (and perhaps depending on social media for socialization) were predictors of harassment, although having a social media account by and in itself was not (Sengupta & Chaudhuri, 2011). Other research confirms that interacting with strangers met online increases the risk of harassment (Mitchell et al., 2008).

Is Social Media Causing Aggression?

In this section, I discuss evidence pertaining to whether social media is related to increased aggressiveness. I do this by examining three related issues. First, is there a direct relationship between social media and aggression. Second, I consider the potential for idiosyncratic effects, such as that social media may interact with personality traits to cause aggression in some but not others. And third, I consider anonymity effects on aggression.

Direct Effects. Although it is clear that the experiencing of aggression on social media is fairly common, it is a more challenging question to address whether social media *causes* aggression, or merely prevents a new and broader social milieu by which individuals can come across and interact with aggressive others. Put simply: do people become more aggressive on social media?

Decades of research have examined the impact of violent media on aggression, largely without having produced any consensus on effects (Savage & Yancey, 2008). More recent evidence from preregistered studies, mostly on video game violence, suggests an absence of effects (Ferguson, 2020b). However, comparatively little evidence has examined for direct relationships between social media use and aggression.

This relatively paucity of evidence may be because social media doesn't inherently involve direct modeling opportunities between viewed violence and physical aggression, which is often the typical focus of media violence studies. This is not to say aggression can't be witnessed online, but rather it isn't social media's central quality as it would be for violent movies or video games.

Some evidence suggests that exposure to extremist content online is associated with more extreme political views and this relationship itself is associated with greater amounts of time online (Pauwels & Hardyns, 2018). However, it is not clear from this evidence that such a relationship is causal. It remains possible that those with more extreme views spend more time seeking out related content. Some evidence does suggest overlap between aggression occurring online with that occurring in real life, however much of that evidence is descriptive or poorly standardized and more work needs to be done (Patton et al., 2014).

Other studies have failed to link social media use to increased aggression (e.g. Memmedova & Selahattin, 2018). Thus, at present the evidence base from correlational studies remains relatively thin, not always of the highest quality and is, at times, inconsistent. It remains unclear whether some aggressive individuals may use social media contexts to express their aggression, or whether social media use, in and of itself, can accelerate aggressiveness.

Causal evidence, however, remains relatively slim. A search for "social media" (subject search) "aggress*" (subject search) and "experiment" (open search), returned only 1 relevant experiment (Chen, 2015). This small ($n = 75$) experiment suggested that losing face, namely criticism of rejection on social media resulted in retaliatory aggression. Although there is some survey-based research on social media use and aggression, clearly there is a need for well-designed, preregistered experiments with adequate power and standardized assessment instruments.

Personality Factors. Some evidence does suggest that personality factors do matter regarding links between social media use and aggression. For instance, individuals who are more agreeable and emotionally stable and less extraverted are less aggressive on social media (McCreery & Kathleen Krach, 2018). Other studies suggest that individuals who are already high in trait aggression are more likely to engage in later cyber aggression (Wright & Li, 2013). This appears to be particularly true for those prone to verbal aggressiveness (Hmielowski et al., 2014). This suggests that there is something of a selection effect at play but doesn't necessarily rule out socialization.

Other evidence suggests that motivational factors may also play a role. For instance, in one study, general social media use as well as use for romantic or social comparison motivations was associated with increased online social aggression (Young et al., 2017). By contrast, using social media for informational or entertainment purposes was associated with lower online aggression. It is worth noting that whether relationships also extend to non-online aggression in the "real world" was not established.

Though small, this area of research is, in some ways, more promising than that for direct effects. It may be likely that some individuals, already aggressive by nature, may use social media for their aggression. In fact, it could be argued that the typical aspects of deterrence for aggression in real-world spaces may not be in place online. This lack of

deterrence may increase the frequency of aggressive behaviors among those already aggressive rather than cause aggression to increase as a trait among individuals. With that in mind, this review now turns to potential anonymity effects.

Anonymity Effects. Some evidence does suggest that the perceived anonymity of some social media platforms may play a role in aggression. For instance, in a longitudinal study of young adults, beliefs in the anonymity of online posts as well as their non-permanency was associated with increased aggressiveness in posting (Wright, 2013). Similar results are found for teens (Wright, 2014) with the added observation that the perceived normativeness of aggression online also predicts aggressive behaviors for the individual.

However, this observed relationship with anonymity may reverse with non-anonymous individuals exhibiting more aggression in some circumstances (Rosz et al., 2016). This is specifically true in circumstances in which aggression toward a perceived bad actor (e.g. a big business, or politician who has misbehaved in some way) allows aggression as a form of moral virtue signaling in which the aggression will be approved of by the non-anonymous individual's followers or social group. In this sense, the motivation for using anonymity to reduce perceived consequences is eliminated and, in the face of low deterrence, aggressiveness becomes more common. Put simply, anonymity decreases deterrence for general aggression, but anonymity is not necessary for aggression that signals moral virtue at another's expense.

Aggression: Conclusions

1. Correlational evidence suggests that some individuals, particularly those lower in agreeableness or emotional stability and high in extraversion may be more inclined to engage in aggression on social media.
2. Further, aggression is more likely under certain circumstances. Anonymity appears to increase aggression rates although, paradoxically, when aggression offers an opportunity to appear morally virtuous, anonymity may be inversely related to aggression. Some aggression is also reactive in terms of individuals responding to criticism or losing face online.
3. Further, evidence is mixed whether identity-based aggression such as sexual harassment is based in actual gender ideology, or rather is due to non-ideological trolling for amusement.
4. What is also evident is that most data thus far are correlational in nature. There is a significant dearth of experimental evidence examining causal elements of social media and aggression. As with other fields, pre-registered designs would be highly desirable.

General Conclusions

In recent years, initial optimism about the role of social media and the internet more generally in public life has faded into

multiple concerns, even as usage has soared. Specific to the current article, expressed concerns include the potential impact of social media on political polarization, mental health and aggression. This review finds a mixed body of evidence. In general, apocalyptic fears of social media causing unique perils for modern society do not appear warranted. However, this does not mean all concerns should be dismissed. Evidence does point to the potential for certain smaller, nuanced impacts on social media.

One of the conclusions to emerge from all three avenues of consideration is that social media impacts are not general or predictable but tend to interact with the motivations and experiences of the user. This observation is consistent with *Uses and Gratifications Theory* (Sherry, 2013) which suggests, briefly, that the interaction between media user and content is more critical than content itself. Different users may have very different responses to the same media based upon their differing motivations. Outcomes may be idiosyncratic rather than predictable with the user a driving force in the exchange rather than a passive recipient of media messages.

This approach is also consistent with Co-construction Theory (e.g. Subrahmanyam et al., 2008). Briefly, this theory posits that off-line and on-line social networks overlap and online users bring their real-life contexts and experience into social media networks, helping to shape them. Thus, social media users are instrumental in developing and shaping online social spaces even as these spaces may have impact on them. This differs from approaches of traditional social science which often posit users as passive recipients of media effects, including social media.

These findings suggest that psychologists might exercise some care when discussing social media effects, neither exaggerating impacts, nor dismissing them entirely. One of the worrisome developments in recent years has been potentially exaggerated claims linking social media use directly to significant harms such as youth suicide. While the good intentions of these scholars are not in doubt, significant data at present would appear to warrant more cautious dialogue. Evidence is weak and correlational and often inconsistent. Exaggerated warnings could result in warning fatigue (wherein the public becomes inured to continued and varied warnings of harm) and distract parents and policy makers from more serious concerns such as privacy issues involving social media. In this sense, the potential for moral panic (a tendency to spuriously blame social problems on technology) appears to be significant for this field as with many related to media effects (Bowman, 2016).

That said, a discussion of more nuanced effects can still provide roadmaps for technology companies and policy makers to reduce negative user experiences and reduce negative impacts. For instance, some behaviors such as vaguebooking (Berryman et al., 2018) could serve as warning signs for significant mental illness, particularly when repetitive in nature. Social media platforms could include options

by which users could be referred to available services. Similarly, platforms could move away from anonymous posting in acknowledgement that these circumstances tend to be related to increased aggression. In other areas, such as political polarization, platforms could look for ways to encourage civil debate and reduce acrimony. Naturally, free speech protections limit the degree to which all adverse social media experiences can be restrained. Indeed, a bias toward regulatory paternalism should never be used to reduce free speech values. But there may be ways to promote positive exchanges without limiting open expression. Naturally, threats and harassment that rise to the level of criminal transgressions should be met with legal deterrence as appropriate. But, more generally, an approach that favors promoting civil exchange and responsible use of social media over heavy-handed restriction and regulation would be desirable.

Recommendations for Researchers

One thing to emerge from this review is that two basic issues continue to haunt research in this field. First, too much research remains non-transparent, non-preregistered and, as such, likely non-replicable. Second, researchers still struggle with what to do with effect sizes that are tiny in magnitude. With these issues in mind, I offer several straightforward suggestions.

1. Preregistration should be routinely used for most research designs. Preregistration should include all hypotheses, measures, and data analytic plans.
2. Standardized assessment measures should be used where applicable, particularly for outcomes which will be generalized to clinical phenomenon such as mental health or aggression.
3. Psychologists should reevaluate whether defensive arguments in favor of interpreting weak effect sizes are serving the field well. These include arguments that weak effects accumulate over time, that weak effects are significant when applied to a large population, or comparisons with medical effects. The first and last arguments have been contradicted either by longitudinal analyses, or reexamination of the data used to calculate medical effect sizes. The second argument is largely speculative and misinterprets the degree to which magnitude of effect among individuals can be used to estimate prevalence of an outcome in the population. It is also worth noting that arguments about interpretation of effect sizes are themselves undermined by the observation that effect sizes themselves are often not reliably estimated. This is particularly true for studies employing single-responder, self-report survey methodology, wherein effect sizes may be significantly over or underestimated.

4. Similar to Orben and Przybylski (2019a), an effect size cut-off of $r = .10$ should be used, below which effect sizes should not be interpreted as evidence in support of a hypothesis even if “statistically significant.” Although more data would be welcome and is likely forthcoming, anecdotal evidence in this field suggests a high rate of false positives for effect sizes below $r = .10$. The range from $r = .10$ to $.20$ should be evaluated with caution with $r = .20$ more likely to indicate clinical significance (Ferguson, 2009).
5. Psychologists may do well to examine their own biases. For instance, when concerning political polarization, it may help for psychologists to remember that their field tends to be weighted toward the left side of the political spectrum. Further, psychologists may be served well to remember their biases may lead toward overinterpreting trivial effects, as well as a tendency toward warning bias, or claiming alarming outcomes, even if evidence is weak or mixed.
6. Given how often secondary data analysis is used in this field, it is recommended that researchers understand and employ techniques for increasing the transparency of secondary data analysis. This includes the preregistration of secondary data analysis (i.e. preregistering analyses plans prior to looking at the data), provision of codebooks, analytic scripts, etc., to make replicating analyses easier (see Weston et al., 2019).

Final Comments

Concerns about social media are unlikely to abate in the near future. Psychologists can be a positive part of this discussion and help promote healthy use of social media. This contribution could be substantial, particularly if psychologists consider some modest and reasonable adjustments to how science is done in this field in order to promote rigor and transparency. To date, evidence suggests that social media's influence on users is nuanced, defying either gloom-and-doom predictions or complete dismissals of effects. Historical analyses suggests that the social media age has not created a period of unique social problems for human societies. Further, high-quality evidence will continue to elucidate the positive and negative involvements of social media with behavioral outcomes for users. A data-based approach, rather than one rooted in traditional fears of new technology and social science biases toward exaggeration of negative impacts will be crucial to guiding policy on social media.

Ultimately, in asking whether the internet has made the world worse, it must be recognized that it is difficult to disentangle the influence of the internet from other changes and historical events that have occurred during the same decades. For instance, if we are seeing a rise in suicide, is that due to social media, the 2008 financial crisis from

which many in the lower classes never recovered, increased globalization, etc. Beginning in 2020, left versus right class struggles became even more acute than they'd been, but it is unclear how much of this can be attributed to the internet as opposed to the covid19, strong disagreements about whether "systemic racism" is real or the best frame for addressing race relations, arguably poor leadership in both main political parties in the US, etc. It can be very difficult to disentangle one factor such as the internet, from these many other systems and issues, particularly when effect sizes in the social science research on the internet and social media tend to be so small. Thus, considering the internet or social media may simply be the wrong question and we'll be best served by understanding the internet as one piece of a larger, more complex system and one which, unfortunately, appears to be enduring a chaotic period.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD

Christopher J. Ferguson  <https://orcid.org/0000-0003-0986-7519>

Notes

1. What constitutes the political left and political right shifts across history and cultures and are generally difficult to define satisfactorily. The terms, used in this paper, are meant to reflect only current US politics and may not apply perfectly to even other contemporary industrialized democracies. They are also meant to be relative in terms of liberal/progressive versus conservative views rather than absolutely precise.
2. Arguing which groups are worst as promoters of domestic terrorism is itself a partisan debate made more complex by unclear definitions and data and the (thankfully) relatively low frequency of terrorism deaths relative to crime-related homicides. Conclusions also differ depending upon whether one assesses number of attempts versus fatalities and can vary from year to year. Most domestic terrorists are US nationals, whatever their ideology. It is probably safest to say that white nationalism and Islamic radicalism remain the two primary sources of domestic terrorism at this time, with smaller contributions from left-leaning groups or those dedicated to racial but non-white radicalism (Government Accountability Office, 2017).
3. Although liberal is often used to contrast with conservative, I use the term progressive here to avoid liberalist philosophy, which is focused on equality, liberty and consent and is not mutually exclusive with conservatism.
4. Not being fond of carrots, I am entirely willing to believe these cause depression.

5. This also highlights that graphs which appear to document large group differences despite small effect sizes should be eyed with suspicion.
6. Indeed, the current author remains largely contented despite regularly skipping breakfast.

References

- Adachi, P. J. C., Hodson, G., Willoughby, T., Blank, C., & Ha, A. (2016). From outgroups to allied forces: Effect of intergroup cooperation in violent and nonviolent video games on boosting favorable outgroup attitudes. *Journal of Experimental Psychology: General*, 145(3), 259–265. <https://doi-org.stetson.idm.oclc.org/10.1037/xge0000145>
- Allen, T. (2011). *Tories: Fighting for the king in America's First Civil War*. Harper Paperbacks.
- Baldassarri, D., & Gelman, A. (2008). Partisans without constraint: Political polarization and trends in American public opinion. *American Journal of Sociology*, 114, 408–446. <https://doi.org/10.1086/590649>
- Baron, R., & Richardson, D. (1994). *Human aggression*. Plenum Press.
- Berryman, C., Ferguson, C. J., & Negy, C. (2018). Social media use and mental health among young adults. *Psychiatric Quarterly*, 89(2), 307–314. <https://doi.org/10.1007/s11126-017-9535-6>
- Best, P., Manktelow, R., & Taylor, B. (2014). Online communication, social media and adolescent wellbeing: A systematic narrative review. *Children and Youth Services Review*, 41, 27–36. <https://doi-org.stetson.idm.oclc.org/10.1016/j.childyouth.2014.03.001>
- Block, J., & Crain, B. (2007). Omissions and errors in 'Media violence and the American public.'. *American Psychologist*, 62, 252–253. <https://doi.org/10.1037/0003-066X.62.3.252>
- Boutwell, B. B., Franklin, C. A., Barnes, J. C., & Beaver, K. M. (2011). Physical punishment and childhood aggression: The role of gender and gene-environment interplay. *Aggressive Behavior*, 37(6), 559–568. <https://doi-org.stetson.idm.oclc.org/10.1002/ab.20409>
- Bowman, N. D. (2016). The rise (and refinement) of moral panic. In Kowert, R., & Quandt, T. (Eds.), *The video game debate: Unravelling the physical, social, and psychological effects of digital games* (pp. 22–38). Routledge/Taylor & Francis Group.
- Brady, D. H., & Pope, J. (2007). Primary elections and candidate ideology: Out of step with the primary electorate? *Legislative Studies Quarterly*, 32, 79–105. <https://doi.org/10.3162/036298007X201994>
- Brownworth, L. (2010). *Lost to the west: The forgotten byzantine empire that rescued western civilization*. Broadway Books.
- Chan, M. (2015). Mobile phones and the good life: Examining the relationships among mobile use, social capital and subjective well-being. *New Media & Society*, 17(1), 96–113. <https://doi-org.stetson.idm.oclc.org/10.1177/1461444813516836>
- Chen, G. M. (2015). Losing face on social media: Threats to positive face lead to an indirect effect on retaliatory aggression through negative affect. *Communication Research*, 42(6), 819–838. <https://doi-org.stetson.idm.oclc.org/10.1177/0093650213510937>
- Choi, D.-H., & Shin, D.-H. (2017). Exploring political compromise in the new media environment: The interaction effects of social media use and the Big Five personality traits. *Personality and*

- Individual Differences*, 106, 163–171. <https://doi-org.stetson.idm.oclc.org/10.1016/j.paid.2016.11.022>
- Clark, C., Liu, B., Winegard, B., & Ditto, P. (2019). Tribalism is human nature. *Current Directions in Psychological Science*, 28(6), 587–592. <https://doi.org/10.1177/0963721419862289>
- Conover, M., Ratkiewicz, M., Francisco, B., Goncalves, A., & Flammini, M. (2011). Political polarization on Twitter. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2847/3275>.
- Davila, J., Hershenberg, R., Feinstein, B. A., Gorman, K., Bhatia, V., & Starr, L. R. (2012). Frequency and quality of social networking among young adults: Associations with depressive symptoms, rumination, and corumination. *Psychology of Popular Media Culture*, 1(2), 72–86. <https://doi-org.stetson.idm.oclc.org/10.1037/a0027512>
- Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., Celniker, J. B., & Zinger, J. F. (2019). At least bias is bipartisan: A meta-analytic comparison of partisan bias in liberals and conservatives. *Perspectives on Psychological Science*, 14(2), 273–291. <https://doi-org.stetson.idm.oclc.org/10.1177/1745691617746796>
- Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society*, 21, 729–745. <https://doi.org/10.1080/1369118X.2018.1428656>
- Everitt, A. (2013). *The rise of Rome*. Random House.
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532–538. <https://doi.org/10.1037/a0015808>
- Ferguson, C. J. (2014). Is reading “banned” books associated with behavior problems in young readers? The influence of controversial young adult books on the psychological well-being of adolescents. *Psychology of Aesthetics, Creativity, and the Arts*, 8, 354–362. <https://doi.org/10.1037/a0035601>
- Ferguson, C. J. (2017). Everything in moderation: Moderate use of screens unassociated with child behavior problems. *Psychiatric Quarterly*, 88(4), 797–805. <https://doi.org/10.1007/s1126-016-9486-3>
- Ferguson, C. J. (2020a). *How madness shaped history*. Rowman & Littlefield.
- Ferguson, C. J. (2020b). Aggressive video games research emerges from its replication crisis (sort of). *Current Opinion in Psychology*, 36, 1–6. <https://doi.org/10.1016/j.copsyc.2020.01.002>
- Ferguson, C. J., & Heene, M. (2021). Providing a lower-bound estimate for psychology’s “crud factor”: The case of aggression. *Professional Psychology: Research and Practice*. Advance online publication. <https://doi.org/10.1037/pro0000386>
- Fraser, R. (2005). *The story of Britain: From the Romans to the present: A narrative history*. W. W. Norton & Company.
- Garimella, V., & Weber, I. (2017). A long-term analysis of polarization on Twitter. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15592/14846>.
- Government Accountability Office (2017). *Countering violent extremism*. <https://www.gao.gov/assets/690/683984.pdf>.
- Graham, H. (1970). The paradox of American violence: A historical appraisal. *The Annals of the American Academy of Political and Social Science*, 391, 74–82. <https://doi.org/10.1177/000271627039100107>
- Grechyna, D. (2016). On the determinants of political polarization. *Economic Letters*, 144, 10–14. <https://doi.org/10.1016/j.econlet.2016.04.018>
- Greenberg, D. (2020). The New York Times used to be a model of diverse opinion. What happened? *Politico*. <https://news.yahoo.com/york-times-survived-1960s-105730411.html>.
- Grieve, R., & Watkinson, J. (2016). The psychological benefits of being authentic on Facebook. *Cyberpsychology, Behavior, and Social Networking*, 19(7), 420–425. <https://doi-org.stetson.idm.oclc.org/10.1089/cyber.2016.0010>
- Grohol, J. (2014). Emotional contagion on Facebook? More like bad research methods. *PsychCentral*. <https://psychcentral.com/blog/emotional-contagion-on-facebook-more-like-bad-research-methods/>.
- Gruzd, A., & Roy, J. (2014). Investigating political polarization on twitter: A Canadian perspective. *Policy and Internet*, 6, 28–45. <https://doi.org/10.1002/1944-2866.POI354>
- Hameleers, M., Bos, L., & de Vreese, C. H. (2018). Selective exposure to populist communication: How attitudinal congruence drives the effects of populist attributions of blame. *Journal of Communication*, 68(1), 51–74. <https://doi-org.stetson.idm.oclc.org/10.1093/joc/jqx001>
- Hammer, D. (2020). The roman republic and the crisis of American democracy: Echoes of the past. *Polis: The Journal for Ancient Greek and Roman Political Thought*, 37, 95–122. <https://doi.org/10.1163/20512996-12340259>
- Han, H., & Brady, D. (2007). A delayed return to historical norms: Congressional party polarization after the second world War. *British Journal of Political Science*, 37, 505–531. <https://doi.org/10.1017/S0007123407000269>
- Heffner, T., Good, M., Daly, O., MacDonell, E., & Willoughby, T. (2019). The longitudinal association between social-media use and depressive symptoms among adolescents and young adults: An empirical reply to Twenge et al. (2018). *Clinical Psychological Science*, 7(3), 462–470. <https://doi.org/10.1177/2167702618812727>
- Hmielowski, J. D., Hutchens, M. J., & Cicchirillo, V. J. (2014). Living in an age of online incivility: Examining the conditional indirect effects of online discussion on political flaming. *Information, Communication & Society*, 17(10), 1196–1211. <https://doi-org.stetson.idm.oclc.org/10.1080/1369118X.2014.899609>
- Hong, S. (2013). Who benefits from Twitter? Social media and political competition in the U.S. House of representatives. *Government Information Quarterly*, 30, 464–472. <https://doi.org/10.1016/j.giq.2013.05.009>
- Hong, S., & Kim, S. (2016). Political polarization on Twitter: Implications for the use of social media in digital governments. *Government Information Quarterly*, 33, 777–782. <https://doi.org/10.1016/j.giq.2016.04.007>
- Howcroft, D., & Fitzgerald, B. (1998). From Utopia to Dystopia: the twin faces of the internet. https://www.researchgate.net/publication/228817859_From_Utopia_to_Dystopia_the_twin_faces_of_the_Internet.
- Huang, C. (2017). Time spent on social network sites and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 20(6), 346–354. <https://doi-org.stetson.idm.oclc.org/10.1089/cyber.2016.0758>
- Jensen, M., George, M., Russel, M., & Odgers, C. (2019). Young adolescents’ digital technology use and adolescents’ mental health symptoms: Little evidence of longitudinal or daily

- linkages. *Clinical Psychological Science*, 7(6), 1416–1433. <https://doi.org/10.1177/2167702619859336>
- Jones, L. M., Mitchell, K. J., & Finkelhor, D. (2013). Online harassment in context: Trends from three youth internet safety surveys (2000, 2005, 2010). *Psychology of Violence*, 3(1), 53–69. <https://doi-org.stetson.idm.oclc.org/10.1037/a0030309>
- Kolmes, K. (2012). Social media in the future of professional psychology. *Professional Psychology: Research and Practice*, 43(6), 606–612. <https://doi-org.stetson.idm.oclc.org/10.1037/a0028678>
- Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, 140(4), 1073–1137. <https://doi-org.stetson.idm.oclc.org/10.1037/a0035618>
- Kramer, A., Guillory, J., & Hancock, J. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 24, 8788–8790. www.pnas.org/cgi/doi/10.1073/pnas.1320040111. <https://doi.org/10.1073/pnas.1320040111>
- Krasnova, H., Widjaja, T., Buxmann, P., Wenninger, H., & Benbasat, I. (2015). Why following friends can hurt you: An exploratory investigation of the effects of envy on social networking sites among college-age users. *Information Systems Research*, 26(3), 585–605. <https://doi-org.stetson.idm.oclc.org/10.1287/isre.2015.0588>
- Lee, C., Shin, J., & Hong, A. (2018). Does social media use really make people politically polarized? Direct and indirect effects of social media use on political polarization in South Korea. *Telematics and Informatics*, 35, 245–254. <https://doi.org/10.1016/j.tele.2017.11.005>
- Lee, J. K., Choi, J., Kim, C., & Kim, Y. (2014). Social media, network heterogeneity, and opinion polarization. *Journal of Communication*, 64(4), 702–722. <https://doi-org.stetson.idm.oclc.org/10.1111/jcom.12077>
- Lin, L. y., Sidani, J. E., Shensa, A., Radovic, A., Miller, E., Colditz, J. B., Hoffman, B. L., Giles, L. M., & Primack, B. A. (2016). Association between social media use and depression among US young adults. *Depression and Anxiety*, 33(4), 323–331. <https://doi.org/10.1002/da.22466>
- Lind, M. (2020). *The New Class War*. Portfolio.
- Manbeck, K. E., Kanter, J. W., Kuczynski, A. M., Fine, L., Corey, M. D., & Maitland, D. W. M. (2018). Improving relations among conservatives and liberals on a college campus: A preliminary trial of a contextual-behavioral intervention. *Journal of Contextual Behavioral Science*, 10, 120–125. <https://doi-org.stetson.idm.oclc.org/10.1016/j.jcbs.2018.10.006>
- Markey, P. M., Males, M. A., French, J. E., & Markey, C. N. (2015). Lessons from Markey et al. (2015) and Bushman et al. (2015): Sensationalism and integrity in media research. *Human Communication Research*, 41(2), 184–203. <https://doi.org/10.1111/hcre.12057>
- McCall, G., & Shields, N. (2008). Examining the evidence from small-scale societies and early prehistory and implications for modern theories of aggression and violence. *Aggression and Violent Behavior*, 13, 1–9. <https://doi.org/10.1016/j.avb.2007.04.001>
- McCreery, M. P., & Kathleen Krach, S. (2018). How the human is the catalyst: Personality, aggressive fantasy, and proactive-reactive aggression among users of social media. *Personality and Individual Differences*, 133, 91–95. <https://doi-org.stetson.idm.oclc.org/10.1016/j.paid.2017.06.037>
- McMartin, S. E., Kingsbury, M., Dykxhoorn, J., & Colman, I. (2014). Time trends in symptoms of mental illness in children and adolescents in Canada. *Canadian Medical Association Journal*, 186(18), E672–E678. <https://doi.org/10.1503/cmaj.140064>
- Memmedova, K., & Selahattin, E. L. (2018). Effects of the technology use on anxiety and aggression levels of youth conducting their higher education studies abroad. *Quality & Quantity: International Journal of Methodology*, 52(Suppl 1), S501–S507. <https://doi-org.stetson.idm.oclc.org/10.1007/s11135-017-0630-4>
- Michener, H. A., & Cohen, E. D. (1973). Effects of punishment magnitude in the bilateral threat situation: Evidence for the deterrence hypothesis. *Journal of Personality and Social Psychology*, 26(3), 427–438. <https://doi-org.stetson.idm.oclc.org/10.1037/h0034467>
- Mitchell, K. J., Wolak, J., & Finkelhor, D. (2008). Are blogs putting youth at risk for online sexual solicitation or harassment? *Child Abuse & Neglect*, 32(2), 277–294. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chiabu.2007.04.015>
- Newman, N., Dutton, W., & Blank, G. (2013). Social media in the changing ecology of news: The fourth and fifth estates in Britain. *International Journal of Internet Science*, 7(1), 6–22.
- Orben, A., Dienlin, T., & Przybylski, A. K. (2019). Social media's enduring effect on adolescent life satisfaction. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 116(21), 10226–10228. <https://doi.org/10.1073/pnas.1902058116>
- Orben, A., & Przybylski, A. (2019a). The association between adolescent well-being and digital technology use. *Nature Human Behavior*, 3, 173–182. <https://doi.org/10.1038/s41562-018-0506-1>
- Orben, A., & Przybylski, A. K. (2019b). Screens, teens, and psychological well-being: Evidence from three time-use-diary studies. *Psychological Science*, 30(5), 682–696. <https://doi.org/10.1177/0956797619830329>
- Paananen, A., & Reichl, A. J. (2019). Gendertrolls just want to have fun, too. *Personality and Individual Differences*, 141, 152–156. <https://doi-org.stetson.idm.oclc.org/10.1016/j.paid.2019.01.011>
- Park, J., Lee, D. S., Shablack, H., Verduyn, P., Deldin, P., Ybarra, O., Jonides, J., & Kross, E. (2016). When perceptions defy reality: The relationships between depression and actual and perceived Facebook social support. *Journal of Affective Disorders*, 200, 37–44. <https://doi-org.stetson.idm.oclc.org/10.1016/j.jad.2016.01.048>
- Patton, D. U., Hong, J. S., Ranney, M., Patel, S., Kelley, C., Eschmann, R., & Washington, T. (2014). Social media as a vector for youth violence: A review of the literature. *Computers in Human Behavior*, 35, 548–553. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chb.2014.02.043>
- Pauwels, L. J. R., & Hardyns, W. (2018). Endorsement for extremism, exposure to extremism via social media and self-reported political/religious aggression. *International Journal of Developmental Science*, 12(1–2), 51–69. <https://doi-org.stetson.idm.oclc.org/10.3233/DEV-170229>
- Pew Research Center (2014a). *Political polarization in the American public*. <https://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/>
- Pew Research Center (2014b). *Online Harassment*. <https://www.pewinternet.org/2014/10/22/online-harassment/>

- Pew Research Center (2017). *Political polarization, 1994-2017*. <https://www.people-press.org/interactives/political-polarization-1994-2017/>.
- Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. Viking, New York.
- Przybylski, A., & Bowes, L. (2017). Cyberbullying and adolescent well-being in England: A population-based cross-sectional study. *Lancet: Child and Adolescent Health*, 1, 19–26. [https://doi.org/10.1016/S2352-4642\(17\)30011-1](https://doi.org/10.1016/S2352-4642(17)30011-1)
- Przybylski, A. K., & Weinstein, N. (2017). A large-scale test of the Goldilocks hypothesis: Quantifying the relations between digital-screen use and the mental well-being of adolescents. *Psychological Science*, 28(2), 204–215. <https://doi.org/10.1177/0956797616678438>
- Qiu, L., Lin, H., Leung, A. K., & Tov, W. (2012). Putting their best foot forward: Emotional disclosure on Facebook. *Cyberpsychology, Behavior, and Social Networking*, 15(10), 569–572. <https://doi-org.stetson.idm.oclc.org/10.1089/cyber.2012.0200>
- Reinecke, L., & Trepte, S. (2014). Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in SNS communication. *Computers in Human Behavior*, 30, 95–102. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chb.2013.07.030>
- Ritchie, S. (2021). Is Instagram really bad for teenagers? *Unherd*. <https://unherd.com/2021/09/facebooks-bad-science/>.
- Romer, D. (2020). Reanalysis of the Bridge, et al. study of suicide following release of 13 Reasons Why. *PLoS One*. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227545>.
- Ronson, J. (2015). *So you've been publicly shamed*. Riverhead Books.
- Rosen, L. D., Whaling, K., Rab, S., Carrier, L. M., & Cheever, N. A. (2013). Is Facebook creating “iDisorders”? The link between clinical symptoms of psychiatric disorders and technology use, attitudes and anxiety. *Computers in Human Behavior*, 29(3), 1243–1254. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chb.2012.11.012>
- Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 57(3), 221–237. <https://doi.org/10.1037/h0087427>
- Rosz, K., Stahel, L., & Frey, B. S. (2016). Digital social norm enforcement: Online firestorms in social media. *PLoS ONE*, 11(6).
- Savage, J., & Yancey, C. (2008). The effects of media violence exposure on criminal aggression: A meta-analysis. *Criminal Justice and Behavior*, 35, 1123–1136. <https://doi.org/10.1177/0093854808316487>
- Sengupta, A., & Chaudhuri, A. (2011). Are social networking sites a source of online harassment for teens? Evidence from survey data. *Children and Youth Services Review*, 33(2), 284–290. <https://doi-org.stetson.idm.oclc.org/10.1016/j.childyouth.2010.09.011>
- Sherry, J. L. (2013). The challenge of audience reception: A developmental model for educational game engagement. In Blumberg, F. C., & Fisch, S. M. (Eds.), *Digital games: A context for cognitive development*. (Vol. 2013, pp. 11–20). Jossey-Bass.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Slawinski, B. L., Klump, K. L., & Burt, S. A. (2019). No sex differences in the origins of covariation between social and physical aggression. *Psychological Medicine*, 49(15), 2515–2523. <https://doi-org.stetson.idm.oclc.org/10.1017/S0033291718003392>
- Smith, E., & Mayer, A. (2019). Anomalous anglophones? Contours of free market ideology, political polarization, and climate change attitudes in English-speaking countries, Western European and post-communist states. *Climatic Change*, 152, 17–34. <https://doi.org/10.1007/s10584-018-2332-x>
- Stella, M., Ferrara, E., & De Domenico, M. (2018). Bots increase exposure to negative and inflammatory content in online social systems. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 115(49), 12435–12440. <https://doi.org/10.1073/pnas.1803470115>
- Stewart, T. A. (1996). Boom time on the new frontier. In Kling, R. (Ed.), *Computerization and controversy* (pp. 67–74). Academic Press.
- Stukal, D., Sanovich, S., Bonneau, R., & Tucker, J. (2017). Detecting bots on Russian political Twitter. *Big Data*, 5, <https://doi.org/10.1089/big.2017.0038>
- Subrahmanyam, K., Reich, S. M., Waechter, N., & Espinoza, G. (2008). Online and offline social networks: Use of social networking sites by emerging adults. *Journal of Applied Developmental Psychology*, 29(6), 420–433. <https://doi-org.stetson.idm.oclc.org/10.1016/j.appdev.2008.07.003>
- Tang, W. Y., & Fox, J. (2016). Men's harassment behavior in online video games: Personality traits and game factors. *Aggressive Behavior*, 42(6), 513–521. <https://doi-org.stetson.idm.oclc.org/10.1002/ab.21646>
- Thucydides (1934). *The complete writings of Thucydides*. Translated by R. Crawley. Random House.
- Twenge, J., Haidt, J., Joiner, T., & Campbell, K. (2020). Underestimating digital media harm. *Nature Human Behavior*, 4, 346–348. <https://doi.org/10.1038/s41562-020-0839-4>
- Twenge, J. M., & Campbell, W. K. (2019). Media use is linked to lower psychological well-being: Evidence from three datasets. *Psychiatric Quarterly*, 90(2), 311–331. <https://doi.org/10.1007/s11126-019-09630-7>
- Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of Abnormal Psychology*, 128(3), 185–199. <https://doi.org/10.1037/abn0000410.supp>
- Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, 6(1), 3–17. <https://doi.org/10.1177/2167702617723376>
- Utz, S. (2015). The function of self-disclosure on social network sites: Not only intimate, but also positive and entertaining self-disclosures increase the feeling of connection. *Computers in Human Behavior*, 45, 1–10. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chb.2014.11.076>
- Van Alstyne, M., & Brynjolfsson, E. (2005). Global village or cyber-Balkans? Modeling and measuring the integration of electronic communities. *Management Science*, 51(6), 851–868. <https://doi.org/10.2307/20110380>
- Vigen, T. (2015). *Spurious correlations*. Hachette Book Group.
- Waasdorp, T. E., Pas, E. T., Zablotnik, B., & Bradshaw, C. P. (2017). Ten-year trends in bullying and related attitudes among 4th- to 12th-graders. *Pediatrics*, 139(6), 1–8. <http://>

- search.ebscohost.com.stetson.idm.oclc.org/login.aspx?direct=true&db=psyh&AN=2017-34833-001&site=ehost-live. <https://doi.org/10.1542/peds.2016-2615>
- Wang, J.-L., Jackson, L. A., Gaskin, J., & Wang, H.-Z. (2014). The effects of Social Networking Site (SNS) use on college students' friendship and well-being. *Computers in Human Behavior*, 37, 229–236. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chb.2014.04.051>
- Weston, S., Richie, S., Rohrer, J., & Przybylski, A. (2019). Recommendations for increasing the transparency of analysis of preexisting data sets. *Advances in Methods and Practices in Psychological Science*, 2(3), 214–227. <https://doi.org/10.1177/2515245919848684>
- Whitaker, R. (2002). *Mad in America: Bad science, bad medicine, and the enduring mistreatment of the mentally ill*. Perseus Publishing.
- Wood, G. (2009). *Empire of liberty*. Oxford University Press.
- Wright, M. F. (2013). The relationship between young adults' beliefs about anonymity and subsequent cyber aggression. *Cyberpsychology, Behavior, and Social Networking*, 16(12), 858–862. <https://doi-org.stetson.idm.oclc.org/10.1089/cyber.2013.0009>
- Wright, M. F. (2014). Predictors of anonymous cyber aggression: The role of adolescents' beliefs about anonymity, aggression, and the permanency of digital content. *Cyberpsychology, Behavior, and Social Networking*, 17(7), 431–438. <https://doi-org/10.1089/cyber.2013.0457>
- Wright, M. F., & Li, Y. (2013). Normative beliefs about aggression and cyber aggression among young adults: A longitudinal investigation. *Aggressive Behavior*, 39(3), 161–170. <https://doi-org.stetson.idm.oclc.org/10.1002/ab.21470>
- Young, R., Len-Ríos, M., & Young, H. (2017). Romantic motivations for social media use, social comparison, and online aggression among adolescents. *Computers in Human Behavior*, 75, 385–395. <https://doi-org.stetson.idm.oclc.org/10.1016/j.chb.2017.04.021>

Author Biography

Christopher J. Ferguson is a professor of psychology at Stetson University. He has worked on issues related to media effects on youth for approximately 20 years. He lives in Orlando with his wife and son.