Computers in Human Behavior Reports 10 (2023) 100281

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Leaving traces behind: Using social media digital trace data to study adolescent wellbeing

Mubashir Sultan [a,b,c,*], Christin Scholz [d], Wouter van den Bos [a,e]

[a] Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany
[b] Department of Psychology, Humboldt University of Berlin, Berlin, Germany
[c] Institute for Interdisciplinary Studies, University of Amsterdam, Amsterdam, the Netherlands
[d] Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, the Netherlands
[e] Department of Developmental Psychology, University of Amsterdam, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Adolescence
Wellbeing
Digital trace data
Social media
Engagement
Screen time

ABSTRACT

Adolescents spend a significant amount of time on social media and there is a great public worry, from parents to policy makers, about the effect of social media on healthy development. Public interest has fuelled ample research on the impact of social media use and wellbeing during adolescence, yet, numerous reviews and meta-analyses report mixed findings that are nested in myriad limitations. One key limitation is an overreliance on high-level measures, such as screen time, as a proxy for the multi-dimensional set of experiences that constitute social media use. In line with a trend moving away from simple but crude measures, we argue that a more nuanced approach that captures the breadth of each individual's behaviours and experience of social media (i.e., their digital phenotype) could benefit the field. In this review, we synthesise what we have learned about the relationship between social media use and adolescent wellbeing and identify outstanding challenges. We then highlight the richness of social media digital trace data and discuss concrete solutions for making optimal use of this data within a structuring framework for future research. Finally, with the particular vulnerability of adolescents in mind, we discuss practical and ethical challenges and limitations of this new approach.

Approximately 3.8 billion people use social media (Kemp, 2020), a relatively recent invention that has rapidly transformed the ways humans interact. Given the exceptional popularity of social media, many researchers have tasked themselves with understanding the effects social media is having on its users. The outcomes of this research are especially pertinent for 'digitally native' adolescents. Present day adolescents are born and raised with advanced digital tools, including social media. They make up the majority of social media users and increasingly conduct their social interactions on social media (Pew Research Center, 2018, 2021). Mirroring historical public responses to the introduction of other media, like radio and television, the advent of social media, and especially the fast and heavy adoption among young users, has been met with intense scepticism (Conley, 2011; Orben, 2020a). The past decade has seen multiple generations of studies investigating relationships between social media use and adolescent wellbeing, largely focusing on potential negative effects. Initial findings highlight that there are indeed relationships between social media use and wellbeing (e.g., Course-Choi & Hammond, 2021; Odgers & Jensen, 2020). However, existing evidence is mixed regarding the directionality of these effects, and untangling and explaining these contradictory findings is difficult because of two key limitations of existing data: an overreliance on high-level measures of social media use, most often screen time, and on cross-sectional correlational methods (e.g., Ellis, Davidson, Shaw, & Geyer, 2019; Parry et al., 2021). There is now a growing awareness in the literature that a full understanding of the relationship between social media use and wellbeing requires us to not only ask how much time users spend online, but to define and study digital phenotypes that describe in much more detail who does what and when while utilising the complex environments offered by social media, and to what effect (Bayer, Ellison, Schoenebeck, Brady, & Falk, 2018; Course-Choi & Hammond, 2021; Valkenburg, van Driel, & Beyens, 2021).

In this review, we argue that social media digital trace data is an immensely rich and precise means to study adolescent (online) behaviour and can help solve the puzzle that has captivated this academic field and the public. While emerging work is beginning to show the potential of this class of methodologies (e.g., Burke & Kraut, 2016), the field is

* Corresponding author. Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195, Berlin, Germany.
E-mail address: mubashirsultan@outlook.com (M. Sultan).

https://doi.org/10.1016/j.chbr.2023.100281
Received 5 September 2022; Received in revised form 25 March 2023; Accepted 27 March 2023
Available online 31 March 2023
2451-9588/

lacking a systematic framework under which to efficiently plan, conduct, and integrate this type of work with other methodologies. Here, we synthesise what we have learned about the relationship between social media use and adolescent wellbeing to date, highlighting outstanding challenges. We then discuss (social media) digital trace data and present an overview of practical insights regarding the use of such trace data for research purposes, culminating in a framework that outlines multiple levels of analyses using trace data. One of the great benefits of digital trace data, the enormous richness, is also its greatest challenge. That is, it constitutes a significant challenge, ethically, conceptually, and methodologically, to make efficient use of this data, while protecting the privacy and rights of research participants. We present digital phenotypes as a practical example for structuring and utilising research data in a theoretically meaningful and statistically valid manner and outline key ethical considerations in this type of work.

1. The puzzle of social media use and wellbeing

As with its popularity among users, research examining the psychological ramifications of social media use has seen a steep rise. Although social media are increasingly popular among most age groups, adolescents—defined here as the transition period between childhood and adulthood (roughly 10–22 years of age)—have adopted this new technology most enthusiastically. For example, more than 70% of 13–24 year olds in the United States and the United Kingdom, on average, spend multiple hours on social media every day (Pew Research Centre, 2018; Pew Research Centre, 2021; Ofcom, 2021; Ilakkuvan, Johnson, Villanti, Evans, & Turner, 2019). This heavy use in combination with rapid physical, social, behavioural, and cognitive development throughout adolescence has moved adolescents into the focus of researchers interested in the nature of the relationships between this new medium of hyper-connectivity and wellbeing (for a review of key areas within social media research, see Fox & McEwan, 2020 and Kross et al., 2020). Overall, as we outline below, reviews on social media use and wellbeing across multiple study designs tend to report mixed (i.e., positive, negative, and null) findings.

Cross-sectional and correlational studies of social media use and wellbeing, be it emotional or social wellbeing, make up the majority of the field. They tend to include general self-report measures of social media use where participants retrospectively quantify the amount or the frequency of social media use, and to a lesser extent, specific engagement related questions (e.g., posting, liking). Orben (2020b), in their review of reviews, highlight that the associations between social media use and the more 'emotional' sides of wellbeing (e.g., depression, anxiety, self-esteem, and loneliness), while mixed (positive and negative correlations), tends to lean on the negative side, ranging from $r = -0.15$ to $r = -0.10$. Appel, Marker, and Gnambs (2020), in their meta-analysis of meta-analyses, when looking at social media use and depression, self-esteem, loneliness, and life-satisfaction as indicators of wellbeing also report a similar negative association ($r = -0.09$ to $r = -0.12$). Finally, in another example, Odgers and Jensen (2020), looking at both meta-analytic and review-based evidence on social media use and depression, anxiety, self-esteem, and loneliness are inconclusive in their search for a direction and refer to inconsistent findings (small positive, negative and null associations). At first glance, the above may, if anything, suggest stronger evidence for a small negative association between social media use and wellbeing. However, when Orben (2020b) and Appel et al. (2020) consider the more 'social' aspects of wellbeing (e.g., social capital, social support, and social connectedness), positive effects of social media use are extant and with larger correlations; although findings in this domain are also mixed.

Next to these cross-sectional data, some scholars have implemented longitudinal designs, often relying on panel or experience sampling designs. These studies also make use of retrospective self-reported frequency of social media use, but measure each participant multiple times, ranging from a couple of weeks to multiple years. Like cross-sectional work, the body of work employing longitudinal methods has produced mixed results (e.g., Course-Choi & Hammond, 2021; Odgers & Jensen, 2020). Finally, experimental evidence remains scarce. Some of the current experimental studies on social media use and wellbeing include 'detox studies', where participants are asked to refrain from a specific social media related activity (e.g., Hall, Xing, Ross, & Johnson, 2021); other studies tend to be a limited simulation of the experience of social media where, for example, participants view 10 photos on an Instagram feed (e.g., Kleemans, Daalmans, Carbaat, & Anschütz, 2018) or are told to expect comments from others (Tobin, Vanman, Verreynne, & Saeri, 2015). As experimental studies are limited, so are reviews that synthesise their findings. Nonetheless, Orben (2020b) highlights that the findings of experimental studies are inconclusive.

Across many studies and methodologies, research on social media use and adolescent wellbeing is characterised by mixed findings and uncertainty about the directionality of effects. These mixed findings are likely due to multiple factors, making it difficult to draw overall definitive conclusions. Firstly, social media is not a homogeneous entity and refers to different platforms (e.g., Instagram, Tik-Tok, Twitter, Facebook, Snapchat) that critically differ (over time) in their user demographics, structure, engagement, and usage norms. Such an appreciation for differences in social media platforms has largely been missing from past studies (e.g., reviews generalise across multiple social media platforms) and researchers have only recently begun to move away from the idea of an aggregate social medium (Kross et al., 2020). Secondly, researchers use different definitions (e.g., emotional versus social) and (often unvalidated) operationalisations of wellbeing (Griffioen, Rooij, Lichtwarck-Aschoff, & Granic, 2020). Finally, the level of detail of conclusions that can be drawn has been limited by the crude nature in which social media use has been operationalised. Most research in this field makes use of some form of self-reported frequency-based measure of social media use and this measure is very often screen time. On the one hand, the validity of measuring self-reported screen time has increasingly been called into question (e.g., Ellis, 2019; Parry et al., 2021; more detail on this below), and on the other hand, screen time does not adequately capture the multi-dimensional set of experiences that constitutes social media use. As an analogy to this point: It is unlikely that time spent at a dinner table impacts health; rather, what is being eaten is likely to be impactful.

When engagement is the subject of assessment in studies, it is often the case that only a specific engagement activity is looked at. For example, a small group of studies have investigated different effects of active versus passive social media use (i.e., differences in social media use patterns). Active social media use entails engaging in social media, such that it facilitates direct engagement with others (e.g., status updates, commenting), whereas passive social media use, also referred to as 'lurking', involves monitoring the activities of others without direct engagement (e.g., scrolling through a feed; browsing profiles; Verduyn, Ybarra, Résibois, Jonides, & Kross, 2017). It is thought that active engagement with social media may lead to increased wellbeing, whereas passive social media use may be associated with a decrease in wellbeing (for a review, see Verduyn et al., 2017). These early findings seem to provide a first step in the direction of a solution to the puzzle of mixed findings regarding the relationship of screen time and wellbeing. Yet, similar to the screen time evidence, different reviews of the active versus passive use literature highlight inconsistent evidence (Course-Choi & Hammond, 2021; Liu, Baumeister, Yang, & Hu, 2019; Valkenburg, Driel, & Beyens, 2021).

Thus, although this work is an important step in the right direction, we argue that the active versus passive social media use dichotomy is not nuanced enough. Indeed, if both social media *activity and content* were to be taken into account, it would become clear that the hypothesis that active social media use leads to improvements in wellbeing is too rigid (Kross et al., 2020; as cited in Valkenburg et al., 2021). For example, not all active social media use may be beneficial for wellbeing. Cyberbullying (and responding to bullies) is active engagement but also an anti-social means of engaging with social media. As such, studies on cyberbullying have shown that active social media use has been associated with a decrease in wellbeing (Kowalski, Giumetti, Schroeder, & Lattanner, 2014).

To appropriately measure the impact of social media use on wellbeing (1) a multi-dimensional approach is required that can capture multiple types of experiences, as well as (2) an appropriate methodology that can capture subtle characteristics of social media engagement (e.g., the activity and the content) with high objective ecological validity. We believe that both of these requirements can be addressed by a combination of social media digital trace data and traditional social science methods. Such a multi-dimensional approach will support the development of more nuanced models of the relationship between social media use and wellbeing and has the potential to resolve inconsistencies in existing evidence. We next delineate what we mean by digital trace data and present some of the many ways social media trace data can be measured.

2. Social media digital trace data analysis framework

Computational social science (CSS) is a new largely interdisciplinary effort that draws heavily from methods in computer science, social network analyses, and communication science. Methods in CSS are able to capture granular, high precision, reliable, and objective data reflecting human behaviour, all-the-while trying to minimise measurement bias (Lazer et al., 2009). The type of data often collected in CSS is referred to as digital trace data, defined as activity records as processed through an online environment. One can, for example, collect digital trace data through Application Programming Interfaces (APIs)[1] offered by platforms (e.g., TikTok and YouTube) or other extraction methods, such as web scraping. These data come in various forms, ranging from time of posting, the content, number of likes, and number of shares, and are very scalable given the billions of people who use social media. As may be apparent, data at such a breadth and scale presents a major challenge when it comes to appropriate data management and theory-driven and valid dimension reduction. A framework for how to do this is presented next (for insight on how to collect social media trace data and its challenges, like the unexpected shutdown of Twitter's API, see Collecting Social Media Data).

We aim to provide an organising framework for the study of effects of social media use on psychological wellbeing using digital trace data. To this end, our framework is focused on accounts (also referred to as a node), the digital representation of an individual.[2] Each account encompasses meta-data (e.g., profile 'about' info), activity data (e.g., number of posts), content data (e.g., post content), and relational data (e.g., how meta-, activity, and content data link to other nodes). Pieces of data from accounts can be used to describe higher-level entities (e.g., the account/user itself, a campaign, or event). Further, multiple nodes can form groups, and finally, multiple groups can combine to form (sub) populations. One dataset drawn from social media can, thus, support analyses of constructs at three broad levels: account (e.g., user), group, and population (Fig. 1). These constructs are also situated in time (t_1) and allow one to compare data between levels of analyses and at different times (t_n; e.g., data pre- and post-pandemic). Each level of analysis is discussed in further detail below.

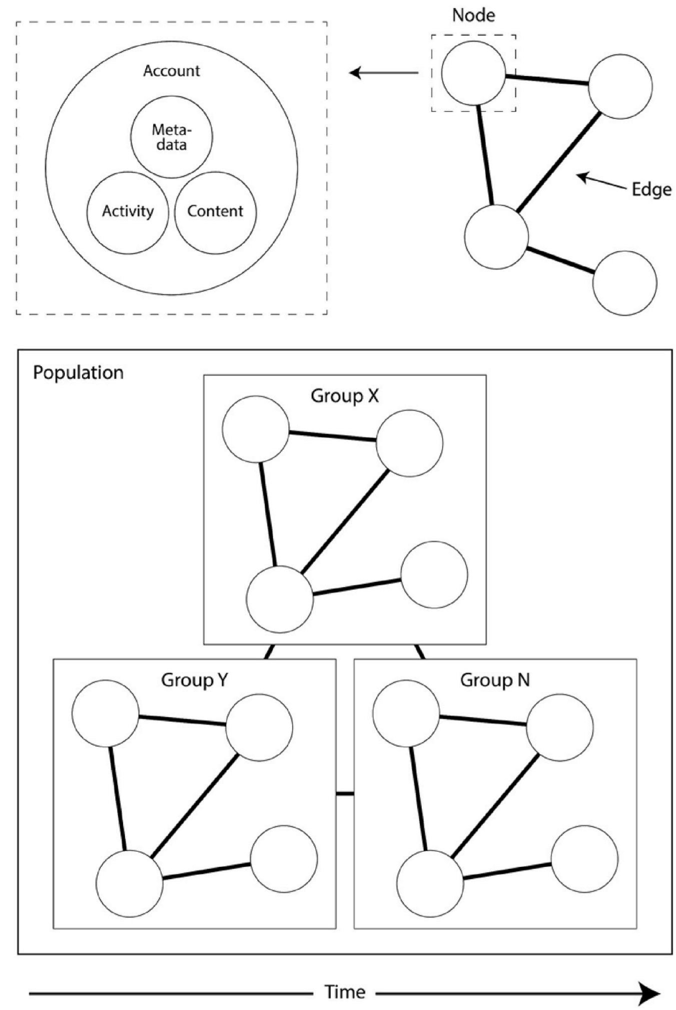

Fig. 1. Levels of analysis showing how an account can provide meta-data, activity data, and content data. Each account (e.g., user, group) is also part of a network and connected to other accounts via edges. Meta-, activity, and content data can offer insight into a specific account, data from multiple accounts can form a group that can be compared to other groups, and multiple groups can form a (sub)population that can then be compared to other (sup) populations. While we take the user account as the fundamental building block, such an approach can also account for inferences on every post using a certain hashtag, given that posts originate from an account.

2.1. Account-level analysis

Via meta-data, activity data, and content data, each account in a social media context offers a multitude of measurement possibilities. Looking at the content of the post (e.g., text, image) could provide insights into the topic of the post, emotion or sentiment conveyed, and/or stance or opinion expressed. Posts also include other metrics (e.g., number of likes, comments, shares, and time of posting) that provide information about activity-related behaviour. By combining information from accounts on a specific social media platform, such as the users' usage/posting behaviour or content of their posts, it becomes possible to generate measures of subtler individual characteristics at scale. Please note that the meta-, activity, and content data mentioned here are not only limited to the publicly visible aspects of social media but also include private interactions (e.g., instant messages; effectively treatable as posts) that are afforded by social media. Private interactions have also been known to vary in their frequency and nature (Valkenburg et al., 2021; Waterloo, Baumgartner, Peter, & Valkenburg, 2018) and need to be considered when operationalising social media use.

[1] In the context of social media, APIs are software that allow other applications and websites to pull social media data and integrate functionality with their site or application.

[2] Other fields relying on CSS may choose a different focus to better serve their research purpose. For instance, work on the spread of misinformation is likely to centre analyses around pieces of content and networks of content rather than individual users.

2.2. Group-level analysis

Groups of interest to the study of social media use and wellbeing can be constructed or identified in multiple ways. For instance, groups may be naturally occurring (where users explicitly describe themselves as members) or researcher-defined/inferred. Naturally occurring groups can include public or private groups (e.g., Facebook groups) that are based around shared characteristics (e.g., age, gender, school/university) and interests (e.g., hobbies, social and political topics, memes, cute animals). In other instances, researchers may define groups that are not explicitly defined outside the research context, for instance, based on network connectivity (e.g., by network analysis; edges in Fig. 1). Some examples of inferred groups are based on follower/friend networks, content networks (e.g., shares), and hashtag networks. To illustrate, popular hashtags can be used to identify communities in which discussions are dominated by a certain topic such as #mentalhealthawareness, #fitgirl, #bodypositive, and #anxiety.

Once a group of interest is established, digital trace data can be used in a multitude of ways to describe the group in great detail. For instance, researchers may describe groups based on activity data, such as the number of members in the group, group demographics, level of group participation, timing of participation (e.g., school week versus on weekends; holidays versus not). In a similar vein, groups can be distinguished based on content data, including topic prevalence (e.g., body types displayed on photos with #fitgirl vs. #bodypositive), sentiment (e.g., positive versus negative), and opinions expressed. Further approaches include describing groups by focusing on information flow between group members and between networks of members (e.g., between groups). Specifically, these methods can answer questions like: Does information travel mostly via direct communication or to a general audience (known as narrowcasting and broadcasting, respectively; e.g., Barasch & Berger, 2014); is information flow driven by influential members (i.e., via opinion leaders; Lazarsfeld, Berelson, & Gaudet, 1948; e.g., Turcotte, York, Irving, Scholl, & Pingree, 2015) or is it more disparate; and what is the speed of information flow (i.e., information cascades; e.g., Kim, 2021)? As both activity and content are not static, one could also study the aforementioned in a dynamic manner (e.g., change or prevalence of topic over time, across events). Furthermore, once groups are identified, they can also be approached to participate in further experimental research focused on specific questions regarding a specific group or group comparisons. For example, the image of an ideal body being thin is more easily accessible than before through the use of hashtags. Past research has shown that exposure and social comparison to thin-ideal imagery on social media is associated with increased body dissatisfaction and worsened mood among both adolescent girls (Kleemans et al., 2018) and adult women (Brown & Tiggemann, 2016; McComb & Mills, 2021). Interestingly, comparing thin-ideal with fit-ideal hashtag imagery, it was found that, in some cases, fit-ideal imagery worsens body image even more so than thin-ideal imagery (Betz & Ramsey, 2017; Robinson et al., 2017). Ultimately, by combining meta-data, activity data, and content data at a group level, one could derive critical aspects relating to adolescent behaviour and their internal dynamics.

2.3. Population-level analysis

One could extend the scale of group-level analyses while controlling for geographical location to infer national and cross-cultural insights (i.e., population-level analyses). Therefore, social media trace data is well-suited for studying and tracking population-level measurements, and indeed can (and has) been used to study phenomena relating to political movements (e.g., #MeToo, #BlackLivesMatter), opinions, cultural attitudes, and even wellbeing (e.g., Manikonda, Beigi, Kambhampati, & Liu, 2018; Mitchell, Frank, Harris, Dodds, & Danforth, 2013).

3. Digital phenotype approach: A proposal

Digital trace data may help solve a key issue in existing work on the relationship between social media and adolescent wellbeing, namely the lack of specificity in describing diverse online experiences and activities. The framework outlined above provides a structured overview of the overwhelming vastness of information included in social media digital trace data. Yet, the novelty of this type of data comes with a lack of scientific consensus regarding efficient and objective ways of utilising it. There is simply too much data for a full, researcher-driven review and analysis. We propose that one efficient way of utilising these data, one that lends itself well to both theory-driven and data-driven approaches, and one that provides a strong backbone for research within an account-centred framework, like the one described above, is through the concept of digital phenotypes. Put simply, digital phenotypes describe types of social media users (represented by accounts) that are clustered based on features of their digital trace data. As such, a digital phenotype is a user characteristic that can represent different social media use experiences and activities which tend to co-occur (Fig. 2). The selecting and clustering of trace data in the construction of digital phenotypes will result in a significant reduction of the data and creation of useful parameters suitable for further statistical analyses. Here we discuss two ways of constructing digital phenotypes, namely, from a theory-driven versus a data-driven perspective.

3.1. Theory-driven approach

Theory-driven creation of digital phenotypes would use existing theories of human behaviour to carefully select features from digital trace data. One such example can be found in the work of Burke and Kraut (2016). They looked at theories of belongingness, relationship maintenance, signals of relational investment, social support, and social comparison for predictions on wellbeing as operationalised via social media interactions. Overall, they differentiated between: (1) directed written communication involving a specific person (e.g., wall-post, comment), (2) low-effort yet direct communication (e.g., likes), and (3) written communication for a broader audience (e.g., status update). This selection of features was based on the idea that those who engage in a lot of directed written communication as compared to those who engage in mostly written communication for a broader audience are likely to be different social media users. That is, while their absolute screen time measurements may be identical, their experience of social media may not (for more details and results, see Burke & Kraut, 2016). Theory driven research may also focus on group processes or group membership. Here, one could use specific hashtags in order to group people by, for instance, political affiliation, or interest. This simple high-level grouping may of course be extended to create subgroups, such as using geolocation to further divide people in urban versus rural areas, thus testing more specific hypotheses.

Another theory-driven approach would be to use computational modelling to extract latent constructs from raw trace data. Computational models can be seen as formalised theories that are able to use raw trace data to make inferences about latent states or variables that drive observed patterns of social media use (for a general primer on the usefulness of computational models, see Smaldino, 2017). One such example can be taken from Lindström et al. (2021). They hypothesised that a specific form of reward learning, or reinforcement learning, would explain posting behaviour on Instagram. Their theoretical model presupposes that there is an intimate link between posting latency and the number of likes received on a post. Using social media activity data representing a million posts, across 4000 participants, they firstly used trace data to show that individuals' posting behaviour could be explained by an attempt to maximise the average rate of reward. Subsequently, they used the parameters of their model to identify different user types in the data. Although the dataset of Lindström et al. (2021) did not include measures on wellbeing, it would have been very

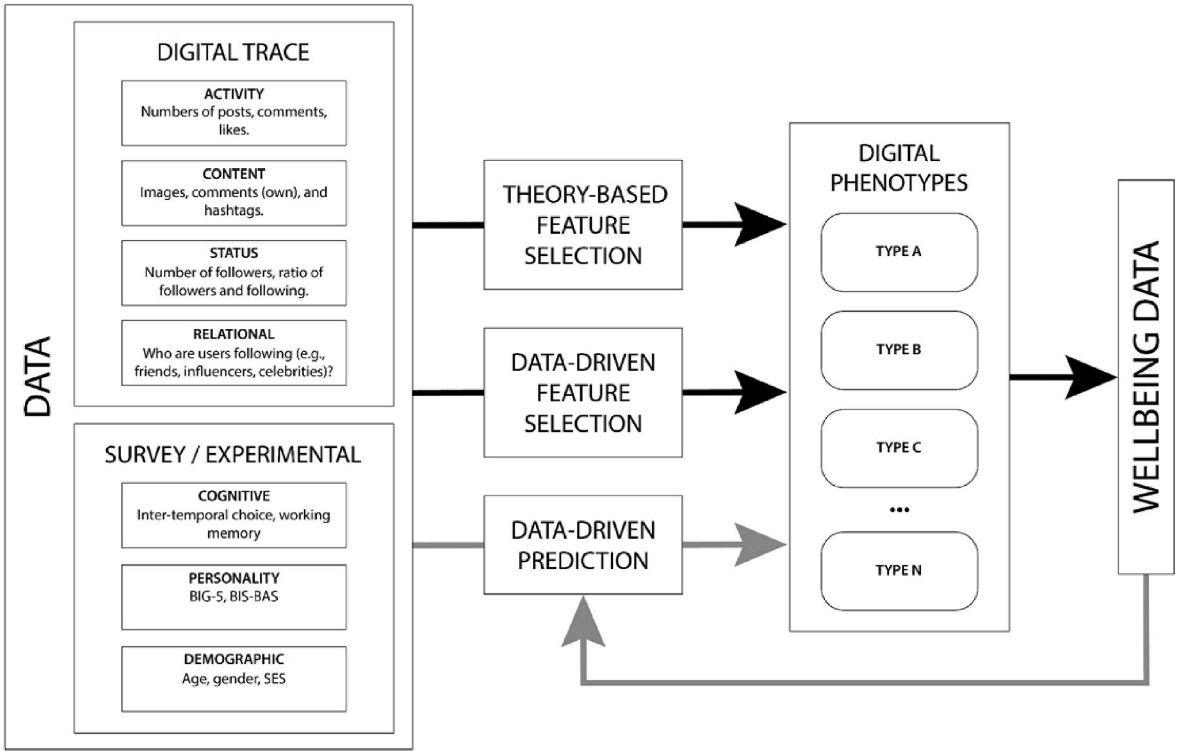

Fig. 2. Theory-based and data-driven approaches can be used to establish digital phenotypes, which can then be used to inform our understanding of wellbeing. The figure also demonstrates how the more traditional data collection methods, via surveys for example, can nonetheless be used to augment trace data (for further explanation, see Enriching Trace Data).

revealing to relate the different user types, partly identified on their sensitivity to likes, to wellbeing. Furthermore, such a computational approach would fit well with the currently emerging field of computational psychiatry, which aims to leverage latent variables in the understanding and treatment of mental ill health (Montague, Dolan, Friston, & Dayan, 2012; Wang & Krystal, 2014).

3.2. Data-driven driven approach

A theory-driven approach can be very useful, as it has been in psychological science for over a century, but it does not make use of the full richness of trace data and is necessarily biased by prior hypotheses. Fully theory-driven approaches may miss unexpected links between social media use and wellbeing. Some reasons include: There is not always a clear mapping between existing theoretical constructs and continuously evolving online behaviours and theoretical constructs may not yet exist for the novel behaviours that the online environment affords (e.g., having multiple social media accounts that simultaneously present different identities). It is also increasingly being argued that psychological science could benefit from a more data-driven approach, one with a focus away from inference and toward prediction (Dwyer, Falkai, & Koutsouleris, 2018; Yarkoni & Westfall, 2017). We believe that such complementary approaches that rely on recent advances in machine learning and artificial intelligence will prove useful for identifying digital phenotypes.

Unsupervised learning techniques can be used to find hidden patterns in the data. For example, one could focus on the following trace data categories: (1) activity data (e.g., number of posts, comments, likes, and log-ins), (2) content data (e.g., images, comments, and hashtags), (3) status data (e.g., number of followers, ratio of followers and following), and (4) relational data (e.g., who are users following). An initial filtering of such is recommended to avoid the phenomenon known as the curse of dimensionality. Generally, the amount of data required for inferences increases exponentially with increased numbers of inputs in the feature space. With these already filtered data, clustering and data reduction techniques—such as k-means clustering (Talasbek et al., 2020) and support vector machines (Pratama & Sarno, 2015)—can be utilised to extract multiple user phenotypes.

As different social media platforms can serve different needs and afford different interactions, data from multiple platforms can also be used to come to more nuanced and comprehensive digital phenotypes. For example, a recent study by Skowron, Tkalčič, Ferwerda, and Schedl (2016) has shown that combining data from multiple platforms can improve the nature of the emerging phenotypes. In their case, it improved the accuracy of predictions about personality traits based on the phenotypes. Such clustering techniques may reveal unexpected user types based on content, sentiment, and activity across different social media platforms. In a subsequent step, these potentially novel phenotypes can be used to explore the relationship with wellbeing and provide novel insights into the relation between social media use and wellbeing.

Alternatively, wellbeing measures, derived from digital trace data, could also be used to directly identify digital phenotypes. One popular machine learning method is using regularised regression in combination with cross-validation. Here, one could enter many, if not all variables,[3] extracted from trace data as initial predictors in the model, and depending on the tuning variable, the model fit would be punished based on the number of variables (or beta value). Based on several iterations of cross-validation, the best set of predictors will be identified. This technique, and similar ones like random forests, have already been applied to both activity and content trace data to predict the presence (and severity) of mental health disorders or symptoms in social media users (e.g., depression, anxiety, subjective mood, wellbeing; for a review see, Chancellor & De Choudhury, 2020). The feature selection

[3] Even for these techniques the number of data points can be overwhelming given the richness of the trace data, thereby, this technique could also be combined with previously mentioned clustering techniques for reduction.

methodology used in these studies is promising and can help explore and identify those elements in the trace data that are the best predictors of wellbeing (see also Kristensen et al., 2017).

3.3. Enriching Trace Data

As we have argued here, social media digital trace data provides rich objective insights about behaviours and experiences of social media users. However, social media digital trace data alone cannot provide a full picture of social media use. For instance, they insufficiently represent subjective experiences of users which are not explicitly described in their content. We next present some ways in which social media digital trace data can be enriched for better inference. We also then touch upon how digital trace data should not substitute other methods of studying social media use, but rather complement them.

3.3.1. Passive social media use

The social media digital trace data described above does not necessarily provide access to the chronology of what participants are exposed to on their social media feed. In a similar fashion, discerning participants who are passively exploring Instagram (e.g., lurking) versus actively using social media (e.g., liking, commenting) is complicated. While it may be possible to construct social media use measures with some social media platforms (e.g., Facebook data logs; see Burke & Kraut, 2016), this data is not easily extracted from raw social media digital trace data provided by many other social media platforms (e.g., Twitter, Instagram). There, nonetheless, have been external developments in the literature, drawing from computer science, that show how such measurements are possible (e.g., built-in screen time metrics [via a phone or a desktop web browser plugin]; Ellis et al., 2019). Furthermore, participants can install software on their computers or portable devices that make it possible to record chronological data, as done in the Screenomics Project (Ram et al., 2020). The premise of Screenomics Project is to capture recordings of individual user screens via a sequence of screenshots, including that of social media use. A software is then used to extract text and images, making the screenshots a searchable database (see also Epstein & Lin, 2022).

3.3.2. Demographics

To make claims about an adolescent population one must know the age of their sample. Yet this information—along with other demographic information—is not directly available on most social media platforms. There are nonetheless ways to infer this information. One could simply ask participants about demographic information via surveys along with their social media data during recruitment (for examples of work linking digital trace data with social media, see work by Stier, Breuer, Siegers, & Thorson, 2020; Al Baghal, Sloan, Jessop, Williams, & Burnap, 2020). Alternatively, suitable for working at a larger scale, researchers have also successfully inferred demographic information from trace data itself, including age (Al Zamal et al., 2012; Chen, Wang, Agichtein, & Wang, 2015; Han, Lee, Jang, Jung, & Lee, 2016), gender (Chen et al., 2015; Al Zamal et al., 2012), race (Culotta, Ravi, & Cutler, 2016; Preoţiuc-Pietro & Ungar, 2018), and education level (Culotta et al., 2016). These approaches include first establishing a ground-truth, for example, via crowdsourcing platforms such as MTurk or Prolific (Appen, 2020; Amazon Mechanical Turk, 2020), which is then used to train larger datasets using machine learning algorithms. These techniques are not as accurate as simply asking participants for their age—which may be impractical to do at scale—they, nonetheless, offer a promising starting point. To get an idea of how such an algorithm may work, see M3 (https://github.com/euagendas/m3inference), a deep learning system that infers demographic attributes directly from social media profiles.

3.3.3. Surveys, experience sampling, and experiments

Just as one can ask for demographic information of participants via surveys, one could also marry other self-report measures of interest with digital trace data, thereby adding further robustness to their inferences ('data-thickening'; Latzko-Toth, Bonneau, & Millette, 2017). Examples of self-report areas of interest can include cognitive factors (e.g., working memory, attention, intertemporal choice) and personality traits (e.g., the BIG-5). Furthermore, when measuring wellbeing, for example via self-report survey, a more temporally sound method can be used. Rather than asking survey questions about wellbeing at predefined times, event-triggered surveys can be deployed. Here, participants would be asked survey questions after social media interactions, such as after liking a tweet, retweeting, or posting content (e.g., Bayer et al., 2018; Griffioen et al., 2020). Self-reported perceptions and experiences can then be linked to digital trace data about those content and activities (e.g., via sentiment analysis) with a temporal resolution that is appropriate given the natural occurrence of relevant behaviours. Another related approach that can be utilised is Trace Interviews (Dubois & Ford, 2015). Trace Interviews is an actor-centred method that takes participants' trace data related behaviour as its starting point (e.g., search histories, social media trace data). Participants are then asked, via semi-structured interviews, to reflect on their online behaviour. Based on qualitative analyses, the Trace Interview method can be used to draw richer insights into the experiences of social media use and wellbeing that may be overlooked by the more quantitative-dominant methods used in the field.

Further, many of the research avenues mentioned thus far are correlational in nature. Causal inferences are, nonetheless, possible. We can once again look to the study by Lindström et al. (2021) for reference. After applying and verifying their computational model on digital trace data, the authors conducted a further experiment where they simulated Instagram to confirm that received 'likes' indeed causally influenced subsequent posting behaviour. The authors were able to make causal claims because of their methodological triangulation: social media trace data was used to test a computational model of reinforcement learning and resulting insights were then used to test mechanisms via a social media simulation with in-real-life participants. The takeaway here is that, in a similar fashion, observations that are gathered from social media digital trace data regarding social media use (e.g., via digital phenotypes) can then be used to test participants in labs using experimental designs, working all the way up to neuronal data (see also Hendriks, de Nooy, Gebhardt, & van den Putte, 2021).

Above, we have described digital phenotypes as a way to organise, reduce, and productively use social media digital trace data to study adolescent wellbeing within the framework we have laid out in this article. We argue that digital phenotypes allow researchers to efficiently and validly make use of the vastness of social media digital trace data to describe social media users, groups, and populations in theory-driven or data-driven ways. Using these tools may help the field to keep moving beyond coarse measures of social media use, indulging the complexity of real-world experiences. Phenotypes that are found to be linked to wellbeing indicators would ultimately provide rich insights into the types of social media use (defined by activity, content, status, and relational data) that may be beneficial, harmful, or unrelated to wellbeing.

In the examples above, we also hope to have gotten across the versatility of digital trace data, showing how it can also be combined to be used with traditional methods in social science, namely with self-report, longitudinal (e.g., experience sampling), and experimental designs. The use of social media trace data does not imply a rejection of traditional social science methods, rather, they can complement each other. At this point in time, the field of research relying on social media digital trace data is brimming with potential, but still in its infancy.

4. Challenges and future directions

The richness of social media digital trace data is also one of its challenges. We have attempted to address this challenge by proposing a framework to organise this richness in a valid manner. There are additional challenges that researchers will need to deal with when working with digital trace data, especially as vulnerable adolescent populations are involved. These include, but are not limited to, collecting and making sense of social media trace data in an ever-changing environment, the unknown dynamics of propriety platforms and issues of generalisability, and importance of keeping ethical consideration at the forefront.

4.1. Collecting Social Media Data

One can collect digital trace data through Application Programming Interfaces (APIs) offered by platforms (e.g., TikTok and YouTube), use third-party APIs, or one can make use of web scrapers/algorithms to extract data. Researchers can also incentivise participants to install plugins and software onto their devices to collect data, be it via a desktop, mobile, or another internet-based device, or to download their own user profiles (e.g., Facebook timeline). Getting one's hand on social media data is not without its limitations, however. One must decide the mode of data collection: Should researchers simply ask participants for the data? For larger studies, should researchers make use of APIs, use web scraping tools, cooperate directly with social media companies, purchase data from data resellers who actively collect social media data, or make use of already collected datasets? Each of these methods come with (dis)advantages, are subject to constant changes in the dynamic social media environment, and use-cases will likely depend on the research question at hand (Breuer, Bishop, & Kinder-Kurlanda, 2020). If we take the example of APIs, they return a JavaScript Object Notation (JSON) data file, which needs technical know-how to convert for data analysis. Luckily multiple open-source tools exist that can help with streamlining data collection.[4]

Social media platforms are also not fully committed to making data available to researchers. APIs are also subject to unexpected changes and researchers and tools have to constantly adapt to these changes. For example, in light of the Cambridge Analytica scandal, Facebook has largely shut-down its API (although, see upcoming 'Researcher API'). As another example, at the beginning of the review process of this paper, Twitter was a prime example of a more-open research-friendly API, which was reflected in high volumes of active research and publications, but throughout the review process, including Elon Musk's takeover of Twitter, the landscape has changed drastically. Their API has been suspended, halting all API related research. There is, however, talk of a paid version being introduced soon; the finality of this is nonetheless uncertain. Thinking in a similar vein, some scholars have even been prompted to write of the age of 'post-API' (Freelon, 2018) or an 'API-calypse' (Bruns, 2019). Overall, API research can be highly beneficial because it is supported by the companies, in line with their policies, and usually provides nicely formatted data, but historically, web scraping is the most independent/flexible and sustainable method of data collection for public data. While digital trace data is rich, obtaining it can be a tricky endeavour, especially in light of a changing playing field. Researchers will be required to constantly adapt their data collection toolkit until a more transparent relationship between the platforms and researchers is made.[5]

4.2. The secrecy of social media platforms and issues of representation

Social media platforms are not necessarily transparent in their decision making, for instance, regarding the functioning of their content ranking algorithms, content moderation policies, the data that is made available to researchers (e.g., via their APIs; Pfeffer, Mayer, & Morstatter, 2018), and regarding internal experiments conducted by their platforms. For example, social media platforms routinely perform A/B testing (or split testing) to assess their products. Generally, A/B testing involves allocating users into random groups that are shown a different variation of the social media platform. The extent and frequency of A/B tests are largely unknown and results are only rarely publicly reported. The outcome of one such A/B test, where one group of users had their 'like' counts on posts masked, is that Instagram users can now hide their 'like' counts. Importantly, the depths of their findings were reported as: 'seeing like counts was beneficial for some, and annoying to others' (Instagram, 2021). Instagram did not choose to disclose the exact experimental setup or release their data for study by researchers, data that could help shed light on important questions such as 'for whom are the like counters beneficial' and 'for whom were the like counters annoying', or topics such as narcissism and social-comparison. While it is empowering that users can now control this aspect of Instagram, this example highlights the conflict between motivations inherent in running a for-profit business and social science research. Sadly, this conflict is largely symptomatic of the secretive nature in which social media platforms operate.

Relatedly, social media platforms exercise content moderation and algorithm-driven personalisation. This can influence what users are less exposed to (e.g., suppression of different body types) and which content is amplified (e.g., emotional content). Content moderation can serve important purposes like the suppression of hate speech, but may also serve as a form of intentional or unintentional censorship, which can, for example, lead to the suppression of minority voices (The Feminist Data Manifest-No; Cifor et al., 2019). There is also the issue of bots and purchased followers and 'likes', which can make certain accounts, posts, and trends appear more influential than they are in reality. These examples highlight that social media and social media use are complex phenomena in themselves rather than purely objective measurement tools. As with any data collection method, digital trace data is more suited to answering certain types of research questions than others. For instance, while bot interference, content moderation and hidden algorithms make it difficult to predict what a given individual will see and why, and limitations in terms of representation limit the generalisability of conclusions, digital trace data is a powerful tool for those who want to quantify objectively what a user was exposed to and what a user did (e.g., likes, page views). Especially in combination with other data sources via 'data-thickening' (Latzko-Toth et al., 2017, pp. 199–214; e.g., via self-reported experiences after content exposure), digital trace data can be a powerful tool to social science researchers.

4.3. Ethics

The novelty, dynamic nature, and diversity of digital trace data have led to extensive and quickly developing discussions about ethical concerns for research using these types of data, especially in vulnerable populations, such as adolescents. Like any research project, an in-depth consideration of the ethical implications of any project utilising digital trace data is paramount. A full discussion of potential ethical concerns in this space is beyond the scope of the current review and extensively covered elsewhere (e.g., The Association of Internet Researchers [AoIR]; 2020; The Feminist Data Manifest-No; Cifor et al., 2019; Sloan, Jessop, Al Baghal, & Williams, 2020). Here, we briefly discuss four categories of ethical concerns: representation, informed consent, privacy and data security.

In the 'Demographics' section above, we alluded to how social media digital trace data can be used to create a probabilistic approximation of

[4] For an overview of available tools, see https://socialmedialab.ca/apps/social-media-research-toolkit-2/.

[5] One approach that some researchers employ is that of working directly with social media platforms (e.g., Social Science One; www.socialscience.one). While this can be an optimal solution to many data gathering limitations, it nonetheless perpetuates elitism: Some researchers will have unfair and unlimited access to data that is sought after by many.

user demographics using classifiers applied to large samples. It is important to realise, actively acknowledge, and analyse limitations of such analyses in terms of representation. First, not all groups in society are equally well represented among the users of any given social media platform. Second, algorithms and classifiers used to infer user characteristics are generally trained on corpora with their own inherent issues in representativeness, leading to biases in the interpretation of new datasets (Buolamwini & Gebru, 2018; Fosch-Villaronga, Poulsen, Søraa, & Custers, 2021). For instance, most gender classifiers represent a binary definition of gender and ignore those who identify as anything other than heterosexual males or females (e.g., LGBTQAI+). To minimise unwarranted harm or negative consequences (for more on this, see also The Feminist Data Manifest-No; Cifor et al., 2019), researchers who utilise digital trace data will need to consciously and explicitly reflect on the representativeness of their data and the biases inherent in their analysis methods. Additional validity and accuracy checks, for instance in combination with other data sources, such as self-reports, can help to quantify and ultimately minimise bias.

Another key issue with a unique manifestation in the context of digital trace data is informed consent, especially of minors. Informed consent is a hallmark of ethical science and, at a minimum, study participants should understand the purpose of the study they are involved in, and that they can withdraw from it. When studying minors, informed consent should (also) be obtained. These requirements are usually easily met in small-scale studies with few. Studies at a larger scale, however, can include millions of users, posts, and data points, making it impractical to individually reach out to millions of users for direct consent, including the consent of their legal guardians. In many cases, large-scale studies make use of publicly available profiles and posts that are easily accessible to researchers and may be seen as part of the public space, which many Institutional Review Boards, and fields of research, deem irrelevant for ethical review, thereby not always requiring consent from individuals. Yet, not all users may be aware of, or expect this potential use of their data, and there may be systematic disparities in media literacy that further complicate this issue. Researchers working in this emerging space have a special responsibility to uphold the high scientific standard of informed consent to the extent possible. To date, some guidance exists on how one should approach the ethics of informed consent at large (e.g., AoIR, 2019; Sloan et al., 2020).

That said, and as mentioned, obtaining informed consent from each individual whose data is being processed in research is not always feasible or even possible. For instance, the sample size, and thus, the generalisability and impact of studies of large amounts of public content would be diminished with the requirement of explicit, individual consent for all. Similarly, in-depth study of an individual's social media profile or feed is complicated by the involvement of content from non-consenting third parties (e.g., their friends) that is visible within the participant's data. Instead of dogmatic statements and policies about informed consent in social media research, it will therefore be essential to evaluate the possibilities and ethical trade-offs of individual research projects and to continuously evaluate policies against the ever-changing social media environment. Key principles that should guide these discussions include: harm to individuals whose data is used in research should be avoided by all means available to the researcher (e.g., through anonymisation efforts and data safety procedures), and any potential risk taken (e.g., in studies without individual consent) should ideally be justified with a potential positive (e.g., pro-social) impact of the research (e.g., improvements in adolescent wellbeing).

When studying vulnerable populations (at scale) using social media data, one also has to think about privacy and data security. Privacy achieved through anonymity is a key consideration when using public data from unconsented individuals. For instance, when studying millions of Tweets on a particular political debate, one may assume that individual users who published those Tweets remain anonymous to the researchers and to the public. Similarly, anonymity is almost always promised to consenting participants in social science research. Participant's name or other unique identifiers are usually replaced to eliminate possible participant identification. However, the richness of social media trace data may allow the identification of an individual through the combination of multiple data points, even if each datapoint on its own is not per se identifiable (e.g., by combining the time the tweet was created, tweet content, number of likes, number of retweets). For example, Ayers, Caputi, Nebeker, and Dredze (2018) found that a tweet could be re-traced back to its user 84% of the time in studies quoting a tweet. As a means to avoid participant disclosure, researchers at present reduce the granularity of trace data and share summary/aggregate level data, including paraphrasing and methods known as ethical fabrication (e.g., Ayers et al., 2018; Markham, 2012). This is, of course, not optimal and comes with disadvantages relating to transparency and reproducibility, hallmarks of the Open Science movement (Open Science Collaboration, 2015). Removing too much data can result in important aspects being masked, leaving in too much data can result in the possibility of re-identification. The field is actively grappling with these issues and some promising approaches have been developed. For instance, one approach would be to create a platform like Facebook's Researcher API (in development), which would be a controlled environment that grants researchers access to raw data (e.g., via a VPN), giving them access to conduct their analyses, but only allowing the export of results. Such actions, along with novel approaches, will be needed to come to a better and more transparent solution to share data taken from an adolescent population without increasing risk of disclosure and misrepresentation (for more on this, see also Sloan et al., 2020).

This section has briefly covered key topics when working with social media digital trace data, including being mindful of underrepresented groups, informed consent, privacy and data security, and transparency. Generally, technological progress tends to run ahead of ethics and guidelines, which are often created in reaction to said progress (e.g., General Data Protection Regulation legislation). In many cases the ethics and guidelines may not even exist at (research) institutes and researchers will have to take responsibility for their choices. To minimise societal harm, we strongly encourage researchers to be mindful of these important ethical and social concerns and to drive innovation in dealing with them. For more guidelines, see AoIR (2019) and The Feminist Data Manifest-No (Cifor et al., 2019).

5. Conclusion

With the exceptional uptake of social media has come an exceptional increase in research on social media use. These rapid developments have brought about ample uncertainty and unsolved puzzles. The puzzle that faces the field of social media use and wellbeing in adolescent populations is that of mixed findings nested on myriad limitations, limitations that will nevertheless require attention even in light of digital trace data. While the last two decades have brought us many insights, in this review, we argue that social media digital trace data has been largely overlooked, offers analyses at multiple levels, and can help bring this field some definitive and well-sought-after answers. One promising way forward for future research can be realised via the digital phenotypes proposal. Yet, while using digital trace data boasts ample opportunities for research, the field of CSS is in its infancy and many crucial challenges and limitations warrant careful consideration (e.g., ethics and informed consent). The field of CSS is also rapidly evolving, however, and keeping up with the discussion in this area will prove important for translating the use of digital trace data for developmental research. We hope that researchers wanting to make use of social media digital trace data to study social media use and adolescent wellbeing will make themselves aware of these limitations and developments to drive novel innovation, helping this emerging field mature.

Funding

Contribution by C.S. is part of an NWO Veni project (with project number VI.veni.191G.034), which is (partly) financed by the Dutch Research Council (NWO). W.v.d.B. is supported by the European Research Council grant number (ERC2018-StG-803338) and the Netherlands Organization for Scientific Research grant number (NWO-VIDI 016.Vidi.185.068).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Al Baghal, T., Sloan, L., Jessop, C., Williams, M. L., & Burnap, P. (2020). Linking twitter and survey data: The impact of survey mode and demographics on consent rates across three UK studies. *Social Science Computer Review, 38*(5), 517–532. https://doi.org/10.1177/0894439319828011

Appel, M., Marker, C., & Gnambs, T. (2020). Are social media ruining our lives? A review of meta-analytic evidence. *Review of General Psychology, 24*(1), 60–74. https://doi.org/10.1177/1089268019880891

Ayers, J. W., Caputi, T. L., Nebeker, C., & Dredze, M. (2018). Don't quote me: Reverse identification of research participants in social media studies. *Npj Digital Medicine, 1*(1). https://doi.org/10.1038/s41746-018-0036-2

Barasch, A., & Berger, J. (2014). Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research, 51*(3), 286–299. https://doi.org/10.1509/jmr.13.0238

Bayer, J., Ellison, N., Schoenebeck, S., Brady, E., & Falk, E. B. (2018). Facebook in context(s): Measuring emotional responses across time and space. *New Media & Society, 20*(3), 1047–1067. https://doi.org/10.1177/1461444816681522

Betz, D. E., & Ramsey, L. R. (2017). Should women be "All about that Bass?": Diverse body-ideal messages and women's body image. *Body Image, 22*, 18–31. https://doi.org/10.1016/j.bodyim.2017.04.004

Breuer, J., Bishop, L., & Kinder-Kurlanda, K. (2020). The practical and ethical challenges in acquiring and sharing digital trace data: Negotiating public-private partnerships. *New Media & Society, 22*(11), 2058–2080. https://doi.org/10.1177/1461444820924622

Brown, Z., & Tiggemann, M. (2016). Attractive celebrity and peer images on Instagram: Effect on women's mood and body image. *Body Image, 19*, 37–43. https://doi.org/10.1016/j.bodyim.2016.08.007

Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society, 22*(11), 1544–1566. https://doi.org/10.1080/1369118X.2019.1637447

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of the 1st conference on fairness, accountability and transparency. *Proceedings of Machine Learning Research, 81*, 77–91. Available from: https://proceedings.mlr.press/v81/buolamwini18a.html.

Burke, M., & Kraut, R. E. (2016). The relationship between Facebook use and well-being depends on communication type and tie strength. *Journal of Computer-Mediated Communication, 21*(4), 265–281. https://doi.org/10.1111/jcc4.12162

Chancellor, S., & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: A critical review. *Npj Digital Medicine, 3*(1), 1–11. https://doi.org/10.1038/s41746-020-0233-7

Chen, X., Wang, Y., Agichtein, E., & Wang, F. (2015). A comparative study of demographic attribute inference in twitter. *Proceedings of the International AAAI Conference on Web and Social Media, 9*(1), 590–593. https://ojs.aaai.org/index.php/ICWSM/article/view/14656.

Cifor, M., Garcia, P., Cowan, T. L., Rault, J., Sutherland, T., Chan, A., et al. (2019). Feminist data manifest-No. Retrieved from: https://www.manifestno.com/.

Conley, D. (2011). Wired for distraction: Kids in social media. *About Time Magazine*. Retrieved from http://content.time.com/time/magazine/article/0,9171,2048363,00.html.

Course-Choi, J., & Hammond, L. (2021). Social media use and adolescent well-being: A narrative review of longitudinal studies. *Cyberpsychology, Behavior, and Social Networking, 24*(4), 223–236. https://doi.org/10.1089/cyber.2020.0020

Culotta, A., Ravi, N. K., & Cutler, J. (2016). Predicting twitter user demographics using distant supervision from website traffic data. *Journal of Artificial Intelligence Research, 55*, 389–408. https://doi.org/10.1613/jair.4935

Dubois, E., & Ford, H. (2015). Qualitative political communication| trace interviews: An actor-centered approach. *International Journal of Communication, 9*, 25.

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology, 14*(1), 91–118. https://doi.org/10.1146/annurev-clinpsy-032816-045037

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies, 130*, 86–92. https://doi.org/10.1016/j.ijhcs.2019.05.004

Epstein, Z., & Lin, H. (2022). Yourfeed: Towards open science and interoperable systems for social media. *arXiv preprint*. https://doi.org/10.48550/arXiv.2207.07478

Fosch-Villaronga, E., Poulsen, A., Søraa, R., & Custers, B. (2021). A little bird told me your gender: Gender inferences in social media. *Information Processing & Management, 58*(3), Article 102541. https://doi.org/10.1016/j.ipm.2021.102541

Fox, J., & McEwan, B. (2020). Social media. In M. B. Oliver, A. Raney, & Bryant (Eds.), *Media effects: Advances in theory and research* (4th ed., pp. 373–388). Routledge.

Freelon, D. (2018). Computational research in the post-API age. *Political Communication, 35*(4), 665–668. https://doi.org/10.1080/10584609.2018.1477506

Griffioen, N., Rooij, M. van, Lichtwarck-Aschoff, A., & Granic, I. (2020). Toward improved methods in social media research. *Technology, Mind, and Behavior, 1*(1). https://doi.org/10.1037/tmb0000005

Hall, J. A., Xing, C., Ross, E. M., & Johnson, R. M. (2021). Experimentally manipulating social media abstinence: Results of a four-week diary study. *Media Psychology, 24*(2), 259–275. https://doi.org/10.1080/15213269.2019.1688171

Han, K., Lee, S., Jang, J. Y., Jung, Y., & Lee, D. (2016). Teens are from mars, adults are from venus: Analyzing and predicting age groups with behavioral characteristics in instagram. *Proceedings of the 8th ACM Conference on Web Science*, 35–44. https://doi.org/10.1145/2908131.2908160

Hendriks, H., de Nooy, W., Gebhardt, W. A., & van den Putte, B. (2021). Causal effects of alcohol-related Facebook posts on drinking behaviour: Longitudinal experimental study. *Journal of Medical Internet Research, 23*(11), Article e28237. https://doi.org/10.2196/28237

Ilakkuvan, V., Johnson, A., Villanti, A. C., Evans, W. D., & Turner, M. (2019). Patterns of social media use and their relationship to health risks among young adults. *Journal of Adolescent Health, 64*(2), 158–164. https://doi.org/10.1016/j.jadohealth.2018.06.025

Instagram. (2021). Giving people more control on instagram and Facebook. Retrieved https://about.instagram.com/blog/announcements/giving-people-more-control. (Accessed 30 July 2021).

Kemp, S. (2020). Digital 2020: 3.8 billion people use social media. *We Are Social*. https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media.

Kim, H. S. (2021). How message features and social endorsements affect the longevity of news sharing. *Digital Journalism, 9*(8), 1162–1183. https://doi.org/10.1080/21670811.2020.1811742

Kleemans, M., Daalmans, S., Carbaat, I., & Anschütz, D. (2018). Picture perfect: The direct effect of manipulated instagram photos on body image in adolescent girls. *Media Psychology, 21*(1), 93–110. https://doi.org/10.1080/15213269.2016.1257392

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/10.1037/a0035618

Kristensen, J. B., Albrechtsen, T., Dahl-Nielsen, E., Jensen, M., Skovrind, M., & Bornakke, T. (2017). Parsimonious data: How a single Facebook like predicts voting behavior in multiparty systems. *PLoS One, 12*(9). https://doi.org/10.1371/journal.pone.0184562

Kross, E., Verduyn, P., Sheppes, G., Costello, C. K., Jonides, J., & Ybarra, O. (2021). Social media and well-being: Pitfalls, progress, and next steps. *Trends in Cognitive Sciences, 25*(1), 55–66. https://doi.org/10.1016/j.tics.2020.10.005

Latzko-Toth, G., Bonneau, C., & Millette, M. (2017). *Small data, thick data: Thickening strategies for trace-based social media research*. The SAGE Handbook of Social Media Research Methods.

Lazarsfeld, L., Berelson, B., & Gaudet, H. (1948). *The people's choice: How the voter makes up his mind in a presidential campaign*. New York: Columbia University Press.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., et al. (2009). Computational social science. *Science, 323*(5915), 721–723. https://doi.org/10.1126/science.1167742

Lindström, B., Bellander, M., Schultner, D. T., Chang, A., Tobler, P. N., & Amodio, D. M. (2021). A computational reward learning account of social media engagement. *Nature Communications, 12*(1), 1311. https://doi.org/10.1038/s41467-020-19607-x

Liu, D., Baumeister, R. F., Yang, C., & Hu, B. (2019). Digital communication media use and psychological well-being: A meta-analysis. *Journal of Computer-Mediated Communication, 24*(5), 259–273. https://doi.org/10.1093/jcmc/zmz013

Manikonda, L., Beigi, G., Kambhampati, S., & Liu, H. (2018). #metoo through the lens of social media. In *International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation* (pp. 104–110). Cham: Springer. https://doi.org/10.1007/978-3-319-93372-6_13.

Markham, A. (2012). Fabrication as ethical practice: Qualitative inquiry in ambiguous internet contexts. *Information, Communication & Society, 15*(3), 334–353. https://doi.org/10.1080/1369118X.2011.641993

McComb, S. E., & Mills, J. S. (2021). Young women's body image following upwards comparison to Instagram models: The role of physical appearance perfectionism and cognitive emotion regulation. *Body Image, 38*, 49–62. https://doi.org/10.1016/j.bodyim.2021.03.012

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS One, 8*(5), Article e64417. https://doi.org/10.1371/journal.pone.0064417

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences, 16*(1), 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Odgers, C. L., & Jensen, M. R. (2020). Annual research review: Adolescent mental health in the digital age: Facts, fears, and future directions. *Journal of Child Psychology and Psychiatry, 61*(3), 336–348. https://doi.org/10.1111/jcpp.13190

Ofcom. (2021). Children and parents: Media use and attitudes report 2020/21. Ofcom.org.uk. https://www.ofcom.org.uk/_data/assets/pdf_file/0025/217825/children-and-parents-media-use-and-attitudes-report-2020-21.pdf

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). https://doi.org/10.1126/science.aac4716

Orben, A. (2020a). The Sisyphean cycle of technology panics. *Perspectives on Psychological Science, 15*(5), 1143–1157. https://doi.org/10.1177/1745691620919372

Orben, A. (2020b). Teenagers, screens and social media: A narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology, 55*(4), 407–414. https://doi.org/10.1007/s00127-019-01825-4

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, 1–13. https://doi.org/10.1038/s41562-021-01117-5

Pew Research Center. (2018). Teens, social media & technology 2018. *Pew research center: Internet. Science & Techniques*. https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/.

Pew Research Center. (2021). Social media use in 2021. Pew research center*: Internet. Science & Techniques*. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/.

Pfeffer, J., Mayer, K., & Morstatter, F. (2018). Tampering with Twitter's sample API. *EPJ Data Science, 7*(1), 1–21. https://doi.org/10.1140/epjds/s13688-018-0178-0

Pratama, B. Y., & Sarno, R. (2015). Personality classification based on Twitter text using Naive Bayes. In *KNN and SVM. 2015 international conference on data and software engineering (ICoDSE)*. https://doi.org/10.1109/icodse.2015.7436992

Preoţiuc-Pietro, D., & Ungar, L. (2018). User-level race and ethnicity predictors from twitter text. *Proceedings of the 27th International Conference on Computational Linguistics*, 1534–1545. https://aclanthology.org/C18-1130.

Ram, N., Yang, X., Cho, M.-J., Brinberg, M., Muirhead, F., Reeves, B., et al. (2020). Screenomics: A new approach for observing and studying individuals' digital lives. *Journal of Adolescent Research, 35*(1), 16–50. https://doi.org/10.1177/0743558419883362

Robinson, L., Prichard, I., Nikolaidis, A., Drummond, C., Drummond, M., & Tiggemann, M. (2017). Idealised media images: The effect of fitspiration imagery on body satisfaction and exercise behaviour. *Body Image, 22*, 65–71. https://doi.org/10.1016/j.bodyim.2017.06.001

Skowron, M., Tkalčič, M., Ferwerda, B., & Schedl, M. (2016). Fusing social media cues: Personality prediction from twitter and instagram. In *Proceedings of the 25th international conference companion on world wide web*. https://doi.org/10.1145/2872518.2889368

Sloan, L., Jessop, C., Al Baghal, T., & Williams, M. (2020). Linking survey and twitter data: Informed consent, disclosure, security, and archiving. *Journal of Empirical Research on Human Research Ethics, 15*(1–2), 63–76. https://doi.org/10.1177/1556264619853447

Smaldino, P. E. (2017). Models are stupid, and we need more of them. *Computational Social Psychology*, 311–331. https://doi.org/10.4324/9781315173726-14

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review, 38*(5), 503–516. https://doi.org/10.1177/0894439319843669

Talasbek, A., Serek, A., Zhaparov, M., Yoo, S.-M., Kim, Y.-K., & Jeong, G.-H. (2020). Personality classification experiment by applying k-means clustering. *International Journal of Emerging Technologies in Learning (iJET), 15*(16), 162. https://doi.org/10.3991/ijet.v15i16.15049

The Association of Internet Researchers (AoIR. (2020). Internet research: Ethical guidelines 3.0. https://aoir.org/reports/ethics3.pdf.

Tobin, S. J., Vanman, E. J., Verreynne, M., & Saeri, A. K. (2015). Threats to belonging on Facebook: Lurking and ostracism. *Social Influence, 10*(1), 31–42. https://doi.org/10.1080/15534510.2014.893924

Turcotte, J., York, C., Irving, J., Scholl, R. M., & Pingree, R. J. (2015). News recommendations from social media opinion leaders: Effects on media trust and information seeking. *Journal of Computer-Mediated Communication, 20*(5), 520–535. https://doi.org/10.1111/jcc4.12127

Valkenburg, P. M., Driel, I. I. van, & Beyens, I. (2021). The associations of active and passive social media use with well-being: A critical scoping review. *PsyArXiv*. https://doi.org/10.31234/osf.io/j6xqz

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Social Issues and Policy Review, 11*(1), 274–302. https://doi.org/10.1111/sipr.12033

Wang, X. J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron, 84*(3), 638–654. https://doi.org/10.1016/j.neuron.2014.10.018

Waterloo, S. F., Baumgartner, S. E., Peter, J., & Valkenburg, P. M. (2018). Norms of online expressions of emotion: Comparing Facebook, twitter, instagram, and WhatsApp. *New Media & Society, 20*(5), 1813–1831. https://doi.org/10.1177/1461444817707349

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zamal, F. A., Liu, W., & Ruths, D. (2012). Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors. Sixth international AAAI conference on weblogs and social media. *Sixth International AAAI Conference on Weblogs and Social Media*. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4713.