

# Do we still need psychological self-report questionnaires in the age of the Internet of Things?

Christian Montag<sup>1</sup> · Paul Dagum<sup>2</sup> · Brian J. Hall<sup>3</sup> · Jon D. Elhai<sup>4,5</sup>

Received: 10 October 2021 / Accepted: 23 November 2021

Published online: 04 January 2022

© The Author(s) 2022 [OPEN](#)

## Abstract

Digital data are abundantly available for researchers in the age of the Internet of Things. In the psychological and psychiatric sciences such data can be used in myriad ways to obtain insights into mental states and traits. Most importantly, such data allow researchers to record and analyze behavior in a real-world context, a scientific approach which was expensive and difficult to conduct until only recently. Much research in recent years linked digital footprints to self-report questionnaire data, likely to demonstrate proof of concept(s)—for instance linking socializing on the smartphone to self-reported extraversion (a personality trait linked to socializing)—in the sciences investigating the human mind. The present perspective piece reflects on this approach by revisiting recent work which has been carried out mining smartphone log and social media data and questions if and when self-report data will still be of relevance in psychological/psychiatric research in the near future.

**Keywords** Digital phenotyping · Mobile sensing · Personality · Self-report · Digital footprints · Big data · Smartphone

For several years, the psychological and psychiatric sciences have increasingly relied on the study of digital footprints to obtain insights into human nature [1–4]. With digital footprints we refer to data left by humans interacting with their digital environment in the age of the Internet of Things (IoT). In particular the smartphone has been a game changer, because this device is currently owned by more than six billion people<sup>1</sup> and the devices are carried around wherever they go. With its manifold built-in sensors, the smartphone provides information on the environment a person currently is navigating, and how one interacts with the device provides insight into diverse mental states and traits (see also the smartphone-psychology-manifesto by Miller [5]). In this context, the terms “digital phenotyping” and “mobile sensing” are often used [6]. Digital phenotyping describes the prediction of mental states and traits using digital footprints from the IoT (including smartphones). Mobile sensing is narrower in its terminology, involving the detection of one’s mental states and traits via mobile devices such as the smartphone. The research path probably most chosen at the moment to obtain insight into mental states/traits from digital footprints is to study digital data or traces such as smartphone log data or from social media and link them to self-report data assessing variables such as mood or personality [7–10]. Is the reliance on participant self-report still an approach for the future?

<sup>1</sup> <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/> (accessed on 20th December 2021).

✉ Christian Montag, mail@christianmontag.de | <sup>1</sup>Department of Molecular Psychology, Institute of Psychology and Education, Ulm University, Helmholtzstr. 8/1, 89081 Ulm, Germany. <sup>2</sup>Applied Cognition, Los Altos, CA, USA. <sup>3</sup>Center for Global Health Equity, New York University (Shanghai), Shanghai, People’s Republic of China. <sup>4</sup>Department of Psychology, University of Toledo, Toledo, OH, USA. <sup>5</sup>Department of Psychiatry, University of Toledo, Toledo, OH, USA.

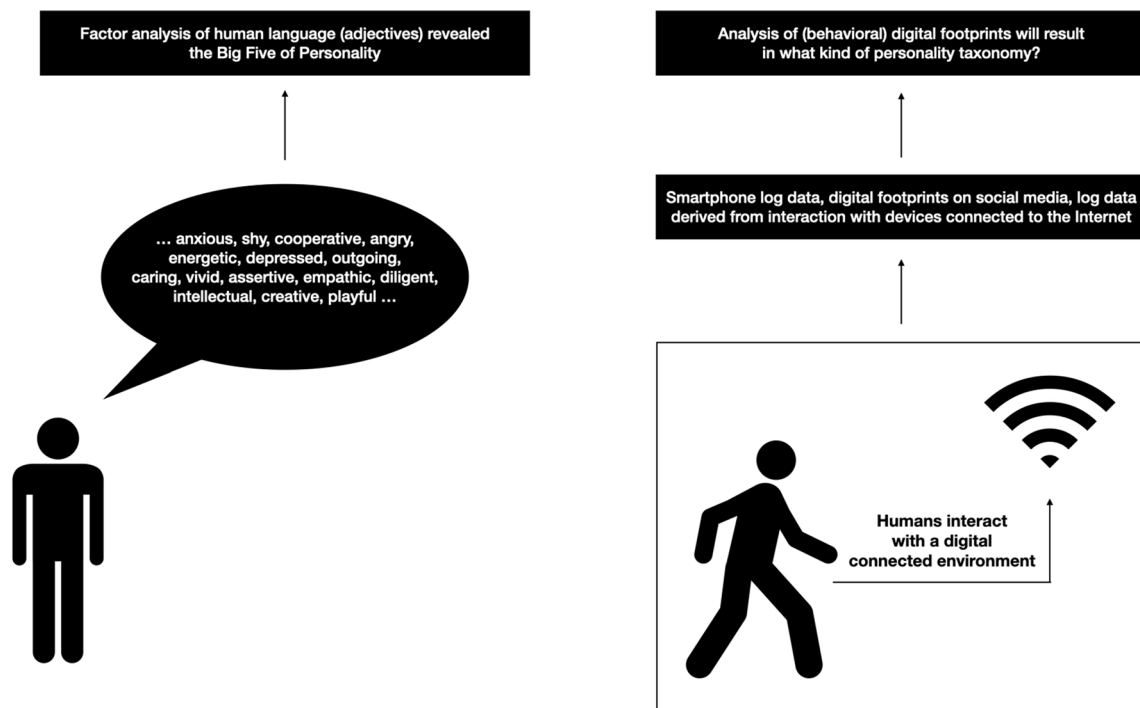


From our perspective, the answer to this question will depend on the stance a researcher takes regarding the validity, reliability and objectivity of self-report measurements. This should be illustrated in the following. A first example touches upon predicting personality from digital footprints. In recent years a large number of studies linked digital footprints on social media/mobile devices with self-reported personality (via questionnaires) [11, 12]. For instance, a meta-analysis by Marengo and Montag, see [13] showed that extraversion in particular could be associated with Facebook data. But what does such a robust link actually mean? Well, first it means that behavior on social media is linked to self-rated scores of the personality dimension extraversion. Hence a person viewing himself/herself as extraverted, and possessing the traits of being social and outgoing, behaves in certain ways on social media as reflected in their digital footprint (e.g., being more active on these platforms). It is well known though that self-report has many problems: some people have difficulty with introspection, so they are simply not good at rating themselves. Furthermore, social desirability might have an influence on how people present themselves on a questionnaire, and also some versions of personality questionnaires might be easier to understand and filled in than others (see a good overview on the self-report method in [14]). Against this background the question arises if digital data should be really validated against such inventories to predict personality or if we should consider new avenues such as simply relying exclusively on the digital data themselves (see also a recent work by Boyd et al. [15] coming to similar conclusions). As behavior—and self-report—ultimately is the output of a biological system, one should also take the importance to link digital data and biological data into account, resulting in digital biomarkers [16].

There is a high consensus that to witness core features of extraversion (as an example) one would need to observe stable social, assertive, and outgoing behavior, as the example above already illustrates. Shouldn't it be desirable—and perhaps even more ecologically valid—to build personality profiles by relying on actual behavior? For instance, one might build such “digital behavioral extraversion scores” from the size of one's social network as derived from the number of telephone numbers on a person's smartphone or one's social network site account (see a study [17] that investigated extraversion and size of different social network layers). Another example might be to record how frequently a person makes contact with their social networks [7]. This would seem to be the more objective way to obtain insight into personality. In this realm, it will be interesting to learn if factor analysis (or other data reduction methods) of digital behavior will yield different (or similar) personality factors as the factor analysis of adjectives, which resulted in the taxonomy of the Big Five of Personality; for a short history of the Big Five see [18]. Interestingly, the recent past has also seen research endeavors applying the lexical analysis approach to social media text data [19]. A different taxonomy might still be a result of the study of digital footprints, when not only text-mining is applied, but actual behavior will be the focus of analysis. See also Fig. 1 illustrating some of these ideas.

From our view, the points discussed above do not mean to suggest that self-report is or will become obsolete when applied studies are focused on the investigation of digital footprints or even psychological/psychiatric research questions as a broader category. Despite the inherent limitations, asking a person about their thoughts/well-being is still the most direct and easiest path to obtain insights into mental states/traits. This said, clinical research at the moment relies to a certain degree on asking persons about their current (diminished) well-being and at the moment such mental states cannot be fully derived by mining data from the IoT (e.g., the smartphone). Beyond this, self-report data without doubt, has proven to yield valid research: if we stick with the example of personality, it has been shown that psychometrically sound measured personality (via self-report) predicted or was associated with important real world outcomes (e.g., diet behavior, financial decision making, online social media use, etc., [20–24]). Therefore, the self-report method has led to important insights over many years. But one could still say: why should we measure personality via questionnaires to predict behavior, if the behavior itself can be readily recorded? Harari et al. [25] speak in this context of physical data, social behavior and smartphone use data to be studied. The future will show what kind of personality taxonomies or other psychological taxonomies appear from the study of different sets of digital data. And this will also shed light on the validity of different methods used—including the self-report method—to investigate the human mind.

In sum, with the abundance of digital data being available in the age of the Internet of Things, the question arises, whether behavioral (digital) data alone should be the primary source of investigation in future research endeavors. In many ways, studies linking digital footprints in a meaningful way to psychological constructs as assessed via self-report questionnaires reflected proof of concept research (see also the extraversion example above). Although interesting, should we move onwards and conduct research differently in this area from now on? The answer is yes and no. Beyond what has been mentioned so far, and taking a different perspective, we believe that there is indeed good reason to always include standardized self-report data in psychological/psychiatric research dealing with digital footprints in upcoming years. For instance, in the study of technological use disorders it is well-known that humans have problems in correctly estimating their own technology use [26]. Work by Turel et al. [27] showed that persons



**Fig. 1** An exemplarily research question for a psychologist in the Internet of Things illustrating how changes in methods applied might impact on established psychological knowledge: the analysis of language (adjectives) led to the Big Five of Personality taxonomy in the psychological sciences (left side of the figure). Will the statistical analysis of behavioral data—as recorded via the Internet of Things—lead to the same taxonomy (right side of the figure)?

**Table 1** Examples of (traditional) self-report approaches and digital phenotyping/mobile sensing approaches to obtain insights into human mental processes and related behaviors (each position in the left and right part of the table is not meant as a direct contrast)

Traditional self-report approaches (some examples of current applied methods)	Digital phenotyping/mobile sensing (some examples of current digital sources to be studied)
Administration of self-report questionnaires	Tracking of smartphone log data providing insights into human–machine–interaction
Conducting structured interviews	Tracking of social media data such as status updates, posted profile photos, Likes, comments <sup>a</sup> , etc.
Conducting unstructured interviews	Tracking of the onboard diagnostics in cars

<sup>a</sup>Please note that language posted on social media could be also seen as a form of self-report (and this underlines our argument that even self-report will happen in an Internet of Things, but it will be mostly of an unstructured nature and consists of everyday-human-communications)

with higher tendencies towards disordered Facebook use had “upward time estimate bias” (p. 84) regarding the passed time, when they filled in questionnaires related to their Facebook use. This phenomenon shows that from a clinical psychologist’s or psychiatrist’s perspective, it is always of great interest to contrast how a person sees or perceives the world and what objectively can be observed with respect to the variable of interest. Discrepancies between objective digital data and one’s subjective views might be the most interesting variables to be studied in the near future. From such a perspective, self-report data will be needed in every case. It might be also true that the role of self-report variables will then change and take on a different role in future work in the psychological and psychiatric sciences. Moreover, at least for the moment, some digital footprints may give more valid insights than others into a person’s characteristics. For instance, it might be the case that a person behaves in a certain way on social media and behaves differently in an “offline” context (a condition not existing anymore in a fully connected world to the IoT). Then, the researcher has to think about what data might be more relevant for his/her research question (for examples of traditional self-report and digital sources to be studied see Table 1) or we come back to the ideas of Walter Mischel [28] using if–then functions to solve the “personality paradox” (p. 250), hence that a person with a

certain personality trait does not always behave consistently according to his/her personality in different environments. A fictitious example to illustrate this would be that a person on social media shows impulsive behavior, but in an offline context is less impulsive. Clearly interesting and important research questions will arise in this context in the near future (also regarding how people cultivate images of themselves in the digital space).

Please note, that in the present short perspective, we cannot discuss the many barriers to be overcome in this rapidly evolving research field. It is not trivial to deal with Big Data from a methodological aspect [29] coming in the three V's (different velocity, variety and volume) requiring analysis methods going beyond inferential statistics. And Big Data is what researchers usually are facing when studying data from the IoT. Finally it is not trivial to deal with the many ethical aspects arising when applying digital phenotyping/mobile sensing, in particular privacy concerns, which need to be dealt with [30–32]. Without fostering trust towards such new methods in patients and study participants, progress in this vivid research area will be hindered.

Finally, we want to mention that the present work mainly drew on examples from personality psychology/clinical psychology. For other disciplines of psychology, such as social psychology, but also cognitive and developmental psychology, where both structured/unstructured interviews or self-report questionnaires are applied, we believe the here mentioned arguments to be equally true. For instance, social psychologists are interested in the study of group behavior. Here, social media data represent a highly relevant research opportunity, to investigate how opinions are formed and how groups interact with each other in political debates. Moreover, smartphone log data could present insights both into developing and declining cognitive functions, such as via reaction times while typing messages or the size of a person's vocabulary used in everyday life. Hence, we are of the opinion that the psychological sciences as a general discipline will see tremendous changes in methods and analysis strategies applied.

**Authors' contributions** CM drafted the first version of this perspective piece. This work was critically revised by PD, BJH and JDE. All authors read and approved the final manuscript.

#### Declarations

**Competing interests** None. Nevertheless, for reasons of transparency, the following is reported. CM mentions that he has received (to Ulm University and earlier University of Bonn) grants from agencies such as the German Research Foundation (DFG). CM has performed grant reviews for several agencies; has edited journal sections and articles; has given academic lectures in clinical or scientific venues or companies; and has generated books or book chapters for publishers of mental health texts. For some of these activities he received royalties, but never from gaming or social media companies. CM mentions that he is part of a discussion circle (Digitalität und Verantwortung: <https://about.fb.com/de/news/h/gespraechskreis-digitalitaet-und-verantwortung/>) debating ethical questions linked to social media, digitalization and society/democracy at Facebook. In this context, he receives no salary for his activities. Finally, he mentions that he currently also functions as independent scientist on the scientific advisory board of the Nymphenburg group. This activity is financially compensated. At the moment he also receives funding from a research project on digital phenotyping by Mindstrong Health, Mountain View, CA, USA. JDE notes that he receives royalties for several books published on posttraumatic stress disorder (PTSD); is a paid, full-time faculty member at University of Toledo; occasionally serves as a paid, expert witness on PTSD legal cases; and receives research funding from the US National Institutes of Health. PD is the founder of Mindstrong Health, a company developing digital phenotyping products for mental healthcare delivery. He served as its Chief Executive Officer from its founding in 2013 through October 2019 and was awarded five US patents on digital phenotyping and digital biomarkers. PD is currently co-founder and CEO of Applied Cognition developing diagnostic and therapeutic solutions for Alzheimer's disease. PD owns stock in Mindstrong Health and in Applied Cognition.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Insel TR. Digital phenotyping: technology for a new science of behavior. *JAMA*. 2017;318:1215–6.
2. Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. *Am Psychol*. 2015;70:543–56.

3. Montag C, Elhai JD, Dagum P. Show me your smartphone ... and then I will show your brain structure and brain function. *Human Behav Emerg Technol*. 2021. <https://doi.org/10.1002/hbe2.272>.
4. Montag C, Duke É, Markowetz A. Toward psychoinformatics: computer science meets psychology. *Comput Math Methods Med*. 2016;2016:2983685. <https://doi.org/10.1155/2016/2983685>.
5. Miller G. The smartphone psychology manifesto. *Perspect Psychol Sci*. 2012;7:221–37. <https://doi.org/10.1177/1745691612441215>.
6. Baumeister H, Montag C, editors. *Digital phenotyping and mobile sensing: new developments in psychoinformatics*. Cham: Springer; 2019.
7. Montag C, Baumeister H, Kannen C, Sariyska R, Meßner E-M, Brand M. Concept, possibilities and pilot-testing of a new smartphone application for the social and life sciences to study human behavior including validation data from personality psychology. *J Multidiscip Sci J*. 2019;2:102–15.
8. Peterka-Bonetta J, Sindermann C, Elhai JD, Montag C. How objectively measured Twitter and Instagram use relate to self-reported personality and tendencies towards internet use/smartphone use disorder. *Human Behav Emerg Technol*. 2021. <https://doi.org/10.1002/hbe2.243>.
9. Seabrook EM, Kern ML, Fulcher BD, Rickard NS. Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. *J Med Internet Res*. 2018;20:e168.
10. Stachl C, Hilbert S, Au J-Q, Buschek D, Luca AD, Bischl B, et al. Personality traits predict smartphone usage. *Eur J Personal*. 2017;31:701–22. <https://doi.org/10.1002/per.2113>.
11. Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proc Natl Acad Sci USA*. 2013;110:5802–5.
12. Lambiotte R, Kosinski M. Tracking the digital footprints of personality. *Proc IEEE*. 2014;102:1934–9.
13. Marengo D, Montag C. Digital phenotyping of big five personality via facebook data mining: a meta-analysis. *Digit Psychol*. 2020;1(1):52–64. <https://doi.org/10.24989/dp.v1i1.1823>.
14. Paulhus DL, Vazire S. The self-report method. *Handbook of research methods in personality psychology*. New York: The Guilford Press; 2007. p. 224–39.
15. Boyd RL, Pasca P, Lanning K. The personality panorama: conceptualizing personality through big behavioural data. *Eur J Pers*. 2020;34:599–612. <https://doi.org/10.1002/per.2254>.
16. Montag C, Elhai JD, Dagum P. On blurry boundaries when defining digital biomarkers: how much biology needs to be in a digital biomarker? *Front Psychiatry*. 2021;12:1690. <https://doi.org/10.3389/fpsy.2021.740292>.
17. Pollet TV, Roberts SGB, Dunbar RIM. Extraverts have larger social network layers. *J Individ Differ*. 2011;32:161–9. <https://doi.org/10.1027/1614-0001/a000048>.
18. Montag C, Elhai JD. A new agenda for personality psychology in the digital age? *Personal Individ Differ*. 2019;147:128–34.
19. Kern ML, Eichstaedt JC, Schwartz HA, Dziurzynski L, Ungar LH, Stillwell DJ, et al. The online social self: an open vocabulary approach to personality. *Assessment*. 2014;21:158–69. <https://doi.org/10.1177/1073191113514104>.
20. Bogg T, Roberts BW. Conscientiousness and health-related behaviors: a meta-analysis of the leading behavioral contributors to mortality. *Psychol Bull*. 2004;130:887–919. <https://doi.org/10.1037/0033-2909.130.6.887>.
21. Montag C, Błaskiewicz K, Sariyska R, Lachmann B, Andone I, Trendafilov B, et al. Smartphone usage in the 21st century: who is active on WhatsApp? *BMC Res Notes*. 2015;8:331. <https://doi.org/10.1186/s13104-015-1280-z>.
22. Friedman H, Kern M. Personality, well-being, and health. *Annu Rev Psychol*. 2014;65:719–42.
23. Plieger T, Grünhage T, Duke É, Reuter M. Predicting stock market performance: the influence of gender and personality on financial decision making. *J Individ Differ*. 2021;42:64–73.
24. Nga JKH, Ken YL. The influence of personality trait and demographics on financial decision making among Generation Y. *Young Consum*. 2013;14:230–43. <https://doi.org/10.1108/YC-11-2012-00325>.
25. Harari GM, Vaid SS, Müller SR, Stachl C, Marrero Z, Schoedel R, et al. Personality sensing for theory development and assessment in the digital age. *Eur J Pers*. 2020;34:649–69. <https://doi.org/10.1002/per.2273>.
26. Parry DA, Davidson BI, Sewall CJR, Fisher JT, Mieczkowski H, Quintana DS. A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nat Hum Behav*. 2021. <https://doi.org/10.1038/s41562-021-01117-5>.
27. Turel O, Brevers D, Bechara A. Time distortion when users at-risk for social media addiction engage in non-social media tasks. *J Psychiatr Res*. 2018;97:84–8.
28. Mischel W, Shoda Y. A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychol Rev*. 1995;102:246–68.
29. Elhai JD, Montag C. The compatibility of theoretical frameworks with machine learning analyses in psychological research. *Curr Opin Psychol*. 2020;36:83–8.
30. Dagum P, Montag C. Ethical considerations of digital phenotyping from the perspective of a healthcare practitioner. In: Baumeister H, Montag C, editors. *Digital phenotyping and mobile sensing: new developments in psychoinformatics*. Cham: Springer; 2019. p. 13–28. [https://doi.org/10.1007/978-3-030-31620-4\\_2](https://doi.org/10.1007/978-3-030-31620-4_2).
31. Matz SC, Appel RE, Kosinski M. Privacy in the age of psychological targeting. *Curr Opin Psychol*. 2020;31:116–21.
32. Montag C, Sindermann C, Baumeister H. Digital phenotyping in psychological and medical sciences: a reflection about necessary prerequisites to reduce harm and increase benefits. *Curr Opin Psychol*. 2020;36:19–24.