Received: 16 June 2022 | Revised: 23 January 2023 | Accepted: 29 January 2023

DOI: 10.1111/jopy.12817

ORIGINAL ARTICLE

WILEY

Predicting Big Five personality traits from smartphone data: A meta-analysis on the potential of digital phenotyping

Davide Marengo[1] | **Jon D. Elhai**[2,3] | **Christian Montag**[4]

[1]Department of Psychology, University of Turin, Turin, Italy

[2]Department of Psychology, The University of Toledo, Toledo, Ohio, USA

[3]Department of Psychiatry, The University of Toledo, Toledo, Ohio, USA

[4]Department of Molecular Psychology, Institute of Psychology and Education, Ulm University, Ulm, Germany

Correspondence
Davide Marengo, Department of Psychology, University of Turin, Via Verdi 10, Turin 1014, Italy.
Email: davide.marengo@unito.it

Abstract

Objective: Since the first study linking recorded smartphone variables to self-reported personality in 2011, many additional studies have been published investigating this association. In the present meta-analyses, we aimed to understand how strongly personality can be predicted via smartphone data.

Method: Meta-analytical calculations were used to assess the association between smartphone data and Big Five traits. Because of the lack of independence of many included studies, analyses were performed using a multilevel approach.

Results: Based on data collected from 21 distinct studies, extraversion showed the largest association with the digital footprints derived from smartphone data ($r = .35$), while remaining traits showed smaller associations (ranging from 0.23 to 0.25). For all traits except neuroticism, moderator analyses showed that prediction performance was improved when multiple features were combined together in a single predictive model. Additionally, the strength of the prediction of extraversion was improved when call and text log data were used to perform the prediction, as opposed to other types of smartphone data

Conclusions: Our synthesis reveals small-to-moderate associations between smartphone activity data and Big Five traits. The opportunities, but also dangers of the digital phenotyping of personality traits based on traces of users' activity on a smartphone data are discussed.

KEYWORDS

Big Five, data, digital phenotyping, extraversion, personality, smartphone

1 | INTRODUCTION

Currently, more than five billion people use a smartphone to communicate via phone calls and social media, to search for information and navigate unknown territory. Given the fact that smartphones are a constant companion in everyday life for many, it is of interest that a bit more than ten years ago the Smartphone Psychology Manifesto was published (Miller, 2012). In this work, Miller (2012) foresaw that the smartphone might have a dramatic impact on psychological research because it can provide researchers with private insights into the lives of smartphone users in a longitudinal fashion via studying digital traces left by human-smartphone interactions. These traces may include logs of call and texting behaviors, information about the usage of smartphone applications, as well as footprints left by simply carrying the smartphone around, such as GPS location and mobility data, and accelerometer and gyroscope logs. In this context, it has been discussed several years ago that psychodiagnostics might strongly profit from including digital footprints (Markowetz et al., 2014) and this led also to many studies linking digital footprints from the smartphone to personality traits. Personality has been shown to be of relevance to better understand relevant variables such as health/physical inactivity (Strickhouser et al., 2017; Sutin et al., 2016) or job performance (Zell & Lesick, 2022), making it a relevant variable for both the health sector and business areas. In spite of the growing amount of research on predicting personality from digital footprints collected via smartphone, a meta-analysis addressing this research area is lacking.

Predicting psychological states and traits from digital footprints on the smartphone has been coined in the literature as digital phenotyping or mobile sensing (Baumeister & Montag, 2019). This part of psychodiagnostics could be seen as a research endeavor falling in the realm of a new research discipline called Psychoinformatics (Markowetz et al., 2014; Montag et al., 2016; Yarkoni, 2012), where psychologists and computer scientists cooperate to better understand the human mind. In this area, most of the studies so far exploited digital traces left on social media to predict personality traits (Azucar et al., 2018; Kosinski et al., 2013; Marengo & Montag, 2020), but also examining depressive tendencies by studying Facebook posts via text mining have been a focus of recent research (Eichstaedt et al., 2018). The aforementioned research from Psychoinformatics led to insights that personality measures correlate around $r = .34$ with individual differences in digital footprints left on Facebook (for a meta-analysis, see Marengo & Montag, 2020). When taking a broader look at different social media platforms, similar results can be observed with an upper limit of around $r = .40$ (Azucar et al., 2018). Interestingly, studies predicting personality directly from smartphone variables tend to be more scarce than those using digital footprints from social media (for a narrative review see, Sariyska & Montag, 2019), possibly because the recording of diverse digital footprints from one's smartphone requires unique developed apps, whereas monitoring of social media activities is technically in parts more easy due to the open character of much of what is posted, but in parts also relies on open APIs (application programming interfaces), which are often restricted (e.g., Meta's Facebook API; see also a call for more access to the platforms for independent academics by Montag et al., 2021). Despite the existing obstacles in conducting smartphone tracking studies, for the present meta-analysis, we were able to detect a total of 25 papers from the literature investigating whether and how strongly individual differences in personality can be assessed using diverse smartphone variables.[1]

To our knowledge, the first study linking personality to recorded smartphone variables was published in 2011 and observed many bivariate correlations of low effect sizes between unique smartphone features and the Big Five personality traits (Chittaranjan et al., 2011). Among emerging associations were total duration of incoming calls with extraversion (0.20) or uses of the calendar with agreeableness (−0.18). This early study also used a machine-learning approach to show how the combination of available smartphone information leads to higher prediction rates. In this work, dimensional personality scores were reduced via median split in either low or high personality scorers, and the highest prediction rates could be observed when predicting extraversion. This said, many research endeavors dealing with smartphone-log-data-personality-associations used different versions of Big Five measures (e.g., TIPI, NEO-FFI, TSDI, etc.) and included different sets of recorded variables (e.g., call activity, social media use, gaming, etc.) from the smartphone to shed light on human personality (e.g., Montag et al., 2015, 2019; Stachl et al., 2017). Here, a recent work by Stachl et al. (2020) is in particular noteworthy, because it included a myriad of smartphone variables such as communication/music/day-night activity and location in a large sample of 624 volunteers, and personality was assessed at both domain and facet level. Interestingly, in this work, the same upper limit of personality prediction observed from social media applications emerged. They observed that a machine-learning approach led to a median of $r = .37$ between smartphone-data and self-reported personality scores. Facets of personality could be predicted even a bit higher with a median of $r = .40$ from available smartphone data.

We mentioned above that personality is linked to a myriad of important life variables such as health behavior, longevity, and job performance (to name a few; e.g., Soto, 2019; Ozer & Benet-Martinez, 2006; see also Montag & Elhai, 2019). Therefore, we believe it to be very important to understand personality prediction rates from smartphones because these might provide insights into a person's personality without having to rely on self-report personality assessment (although this is a very optimistic assumption; for more reflections on self-report in the Internet of Things see Montag, Dagum, et al., 2022). This said, assessing personality from digital traces could be misused for micro-targeting campaigns (Matz et al., 2020; Montag et al., 2020). In this context, Shoshana Zuboff raised awareness for privacy concerns and the problems arising from surveillance capitalism, hence a tech-industry basing their business cases on the surveillance of humans while using online platforms (Zuboff, 2019). As such, it is of high relevance to understand with what kind of precision a relevant variable such as personality can be derived from digital footprints of the smartphone.

In light of these considerations, the present work aims to understand the overall strength of association between Big Five personality traits and smartphone data when the available literature is summarized by means of meta-analysis, thus providing insights into how far we are in the field in prediction of personality from digital footprints derived via the smartphone, and whether some traits may be better predicted than others. Note that personality traits from the Big Five framework have been referred to using many names and definitions over time; here, we use those adopted by Costa and McCrae (1992) for the five broad domains of personality, namely *agreeableness*, referring to individual differences in cooperativeness, trustfulness and altruism; *conscientiousness*, referring to individual differences in dutifulness, self-discipline, and deliberation; *extraversion*, referring to individual differences in warmth, assertiveness, and overall energy; *openness*, referring to individual differences in imagination, creativity, and aesthetic sensibility; and *neuroticism*, referring to individual differences in self-consciousness, emotional vulnerability, and proneness to anxiety, depression, and anger.

Beyond establishing the central tendency of the personality-smartphone data associations, we also investigate potential moderator effects related to the time of publication, methods used to investigate associations between personality and features extracted from smartphone data by selected studies, i.e., use of predictive models versus bivariate correlations, and the type of smartphone data analyzed by the studies. Because of significant changes in mobile technology over time, including increasingly faster Internet connection services, new features and applications, one would expect these changes to have affected the way smartphone-derived data relates to personality traits, influencing the strength of associations. Regarding the importance of different methodologies employed in the selected studies, and in keeping with previous meta-analyses investigating the use of digital footprints to predict personality (Azucar et al., 2018; Settanni et al., 2018), we expect that studies combining multiple features for the prediction of personality should improve over studies reporting simple bivariate correlations in terms of overall predictive performance. Finally, following the established notion that personality traits show trait-specific associations with the use of different smartphone functionality (e.g., Burtăverde et al., 2021), we hypothesize that specific types of smartphone data might show differential associations with personality traits, improving prediction. The rationale for this hypothesis relates to the fact that different underlying motives for smartphone use tend to result in different smartphone usage behaviors—for example, social versus process (non-social) smartphone use (Elhai et al., 2017), and these differences are expected to be in part related to individual differences in Big Five personality traits. For example, extraverted individuals are expected to show a higher frequency of calling behaviors than introverted individuals due to increased social motives for smartphone use, while neuroticism tends to be positively related to the amount of time spent consuming media due to heightened escapism tendencies and mood modification purposes (Stachl et al., 2017). Please note that this meta-analysis was not pre-registered.

2 | METHODS

2.1 | Literature search

In order to identify papers investigating the association between Big Five personality traits and features extracted from smartphone data, we followed the PRISMA guidelines (Moher et al., 2015) and performed multiple searches in several databases, using multiple groups of keywords. Identified papers were then screened based on specific inclusion and exclusion criteria. To perform the database searches, the authors collaborated in creating a list of keywords that could be used to identify papers referring to smartphones, namely: *smartphone**, *mobile phon**, *mobile-phone**, *mobilephone**, *smart phone**, *smart-phone**, *phone**. Another group of keywords was used to identify papers investigating features extracted from smartphone data, namely: *sensing*, *sensor**, *application**, *app**, *touch*, *log*, *passive*, *data*, *us**, *pattern**, *record**. The asterisk symbol was used as a wildcard during the database searches to allow for different forms of each term (e.g., plural forms) to be detected. These resulting groups of keywords were combined with the following keywords referring to Big Five personality traits: *personality*, *traits*, *Big 5*, *Big five*, *five-factor model*, *extraversion*, *introversion*, *neuroticism*, *emotional stability*, *openness*, *conscientiousness*, *agreeableness*, *extrovert*, *introvert*, *neurotic*, *open*, *agreeable*, *conscientious*, *emotionally stable*.

Using the resulting set of keywords, we conducted a broad literature search using the following databases: Scopus, and ISI Web of Science. By focusing on these databases, we focused on published studies, therefore excluding unpublished studies (e.g., pre-prints) from the searches. The database literature search was first finalized in November 2020, and later updated in March 2022. A flowchart illustrating the selection process is shown in Figure 1.

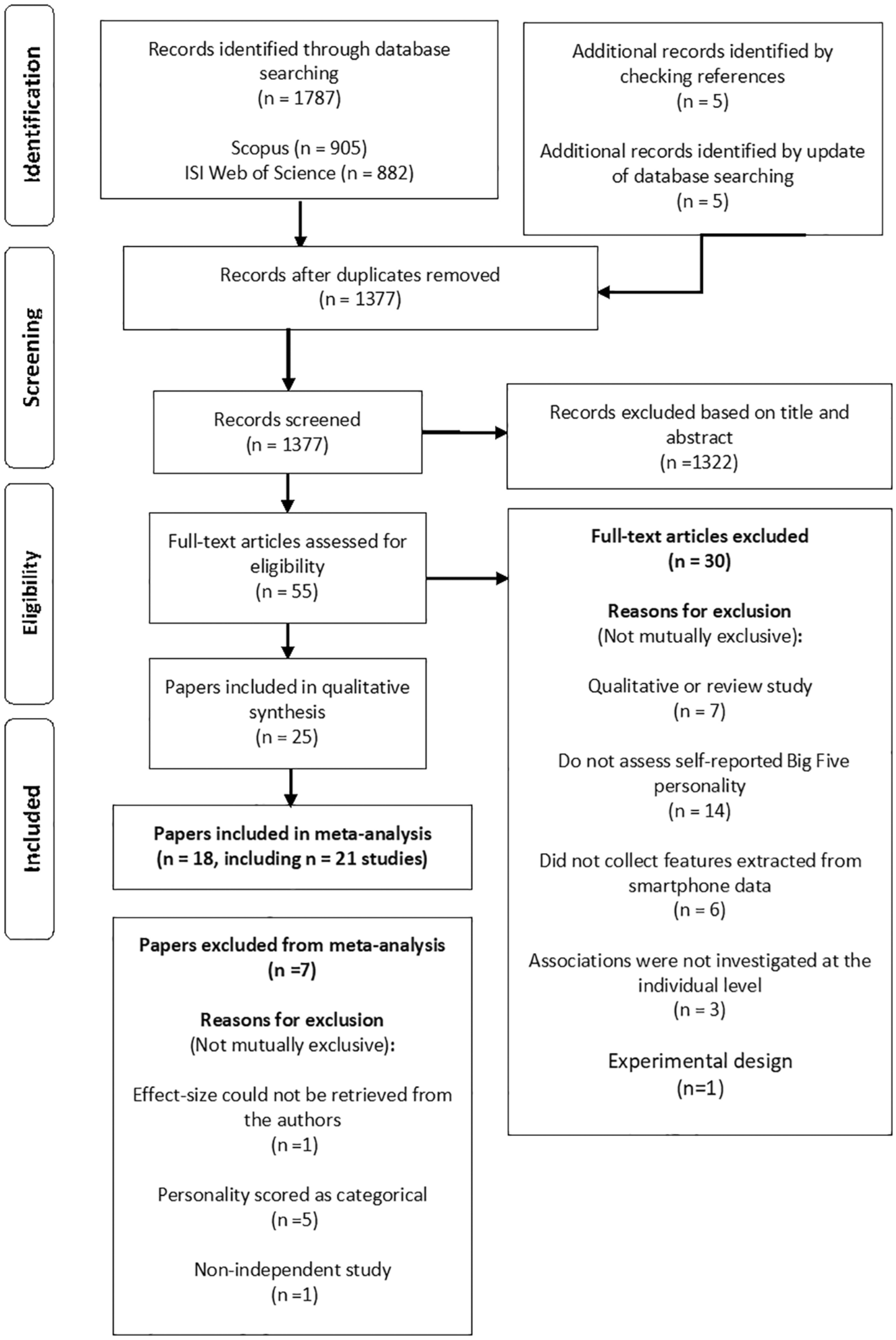


FIGURE 1 Flow diagram of study selection.

2.2 | Inclusion and exclusion criteria

Papers identified through database and reference searches were screened for the following inclusion criteria: retained papers had to: (1) Present the result of an original quantitative empirical study; (2) Include a self-report assessment of Big Five personality traits; (3) Collect quantitative features extracted from smartphone data; (4) Report effect size information on the association between Big Five personality traits and features extracted from smartphone data at the individual level.

Exclusion criteria were used to exclude studies that: (1) Did not provide effect size information, or this information could not be derived from the data available in the paper, or be obtained by contacting the authors; (2) Included effect sizes that were not computed at the individual level, but instead at an aggregated level (e.g., regional level); (3) Did not assess measures of smartphone activity data (e.g., they assessed smartphone activity via self-report, or examined data extracted from other types of devices, such as laptops or tablets); (4) Employed an experimental intervention research design (i.e., researchers did manipulate participants' smartphone activity).

Studies were considered non-independent based on the following criteria (1) Studies performed on samples including the same group of participants, and (2) Associations were investigated in the same smartphone data and by using the same Big Five measures. In case studies that had both the same sample and examined the same set of variables, studies including results based on model-based prediction of personality scores were selected over those only reporting bivariate associations between smartphone features and personality scores. In case studies reported the same results, the earliest study was selected.

2.3 | Research coding

Aside from retrieving effect size information, selected studies were coded for the following variables: study sample size, year of publication, distribution of gender (percentage by gender group) and age (mean and standard deviation), self-report personality assessment, the operationalization of personality, type of analytical approach used to analyze the data, and both the operative system and the type of collected smartphone data.

In coding the operationalization of personality, we distinguished between studies in which personality was scored using continuous scores as opposed to a categorical approach (e.g., dichotomous variables via median split). In coding statistical approaches, we distinguished between studies investigating bivariate associations (i.e., correlation between a personality variable and single feature extracted from smartphone data), and predictive (multivariate, adjusted) models aimed at prediction of personality based on a set of features extracted from smartphone data, either using a categorical classification approach (i.e., prediction performed on categorical transformation of personality scores as the dependent variable) or using a regression approach (i.e., personality traits were predicted in their original continuous form). Finally, when coding studies for the type of smartphone data, we distinguished between variables computed on the following data: (1) app usage data (i.e., indicators derived from the usage of or merely the installation of applications on the smartphone, including productivity, gaming, and social media app data); (2) sensors and system data (namely indicators derived from the use of sensors such as GPS, Bluetooth, WIFI, accelerometers, microphones, etc., and from system information, such as screen and battery usage; or information about smartphone components); and (3) call and texts log data (i.e., indicators pertaining calling and texting behaviors, as well as the managing of phone contacts). Please note that this coding strategy was in part driven by availability of studies.

2.4 | Strategy of analyses

For each selected study, we collected an effect size expressing the strength of association between Big Five personality traits and features extracted from smartphone log data. More specifically, we used the absolute value of correlation coefficients (Pearson's r or Spearman's ρ) as a metric for the strength of the association of smartphone data and Big Five personality scores. Effect sizes were only collected from studies in which personality traits were operationalized (i.e., scored) using their original continuous metric, as opposed to data-driven categorical coding (e.g., median split). We followed this approach because of the overall lack of theoretical and empirical support for the use of categorical typologies when scoring Big Five assessments, especially when brief assessments are used (Freudenstein et al., 2019; McCrae et al., 2006; Pittenger, 2004). Some of the studies investigated the association using a predictive approach, typically employing different machine-learning algorithms. For these studies, we selected the effect size for the best-performing predictive approach (i.e., the approach resulting in the highest absolute correlation between observed and predicted personality scores). When studies only reported associations between single features extracted from smartphone data and Big Five personality scores, we selected the highest effect size reported as the best available approximation of overall strength that would be achieved by a model including the entire set of features as predictors. In some cases, studies included both information about bivariate associations, and results of predictive models: if results of a model could not be expressed or transformed to correlations, we selected the strongest bivariate effect size. When studies did not report correlations, the reported effect sizes were converted to correlations (Rosenthal, 1994). In case we could not find correlations reported in the manuscript, or reported effect size information could not be converted to a correlation coefficient, we contacted the authors of the study to retrieve the missing information. If a study mentioned that an effect was not significant but failed to provide a correlation, the effect size was coded as zero.

Based on retrieved effect sizes, we conducted five separate meta-analyses, one for each Big Five trait. Meta-analyses were performed using a random-effect model because we expected true effect sizes to show significant between-study heterogeneity due to the diversity of characteristics of existing studies. Additionally, because of the lack of independence of many included studies, in estimating meta-analytical correlations, we used a multilevel approach. More specifically, analyses were performed using the rma.mv() function of the *metafor* package (Viechtbauer, 2010) for R by implementing a three-level meta-analytic model modeling three different variance components: at level 1, we modeled the sampling variance of the extracted effect sizes (i.e., the indeterminacy in effect sizes due to the use of samples, as opposed to population data to compute effect sizes); at level 2, we modeled variance at the study-level (i.e., between study variance); and at level 3, we accounted for variance related to data sources. Heterogeneity of effect sizes was investigated by computing the Q test of heterogeneity, the I^2 statistic representing the proportion of true variation in observed effects, and by determining the percentage of heterogeneity due to the different variance components (i.e., data source, study, and sampling variance). Next, following recommendations by Viechtbauer and Cheung (2010), in order to check the robustness of estimated meta-analytical correlations, a sensitivity analysis was performed by comparing the pooled effect size before and after omitting potential outliers. Effect sizes were labeled as outliers if their 95% confidence interval did not overlap with the 95% confidence interval of the previously estimated pooled effect.

Given the great variability in methodology and type of data examined, we expected a significant amount of heterogeneity in the effect sizes. Following suggestions by Viechtbauer (2007), we investigated a series of potential moderators of the effect sizes representing the association between smartphone data: year of publication (year), use of model-based prediction to infer personality from smartphone data (Yes = 1; No = 0); and type of smartphone data. The impact of the type of data was examined using three distinct dichotomous variables, one for each type of data (i.e., app usage; sensors and system data; call and text log data), indicating whether the type of data was used to generate the indicator on which the effect size was computed on (Yes = 1; No = 0). In meta-regression analyses, then, the effect of one type of data was tested against the other two (e.g., call and text log data vs. other types of data). We decided not to test more than one effect per meta-regression because of considerations related to statistical power. To ensure robustness of coefficient estimates, we followed the suggestion by Fu and colleagues and examined the effect of categorical moderators only if at least 4 studies per group were available (Fu et al., 2011). A critical value of $\alpha = .05$ was used for detecting effects in meta-regression analyses.

Finally, publication bias was investigated by inspecting the funnel plot of studies' effect sizes against their relative standard error. Symmetry of the funnel plot was determined by using a modified Egger's intercept test (Sterne & Egger, 2001). More specifically, we fitted a multilevel model predicting study effect sizes with sampling standard errors (i.e., the square root of sampling variance) as a moderator. The assumption of the test is that when publication bias is present, studies based on small sample sizes would have a greater chance of becoming published when reporting large effect sizes. Publication bias is thus detected when a positive association is found between the size of effects and their standard error. Classic fail-safe *N* was then used to evaluate the impact of a *file-drawer* problem (e.g., the number of unpublished studies reporting non-significant associations needed to nullify emerging meta-analytical associations). Here, we refer to Rosenthal's rule of thumb ($5 \times$ *number of effect sizes* $+ 10$; Rosenthal, 1979) to determine the cut-off value indicating the relevance of the file-drawer problem. A fail-safe number larger than the cut-off value would indicate that only a large amount of unpublished papers would be able to nullify the emerging effect size, ultimately downsizing the relevance of the "file-drawer" problem.

All analyses were performed using the *metafor* package for R (Viechtbauer, 2010). We include the relevant correlations and the program code (i.e., R script) to enable other researchers to reproduce our work as Supplementary material.

3 | RESULTS

3.1 | Overview of included studies

In total, we identified 25 eligible papers reporting on studies investigating associations between features extracted from smartphone data, and Big Five personality scores at the individual level. In depth information about characteristics of the identified studies are reported in Table S1 (Supplementary material).

Among the selected papers, we identified $n = 5$ papers that only used a categorical operationalization of personality when examining associations with smartphone data (de Montjoye et al., 2013; Lepri et al., 2016; Mønsted et al., 2018; Peltonen et al., 2020; Staiano et al., 2012). These studies were not included in the meta-analysis. However, in the discussion, we provide a qualitative comparison of the results of the meta-analysis and those emerging from these papers. Additionally, in one eligible study, correlations between personality and smartphone data were not reported in the manuscript, and could not be retrieved from the authors (Servia-Rodríguez et al., 2017).

We also found $n = 10$ papers presenting studies performed on non-independent samples. Specifically, we found $n = 2$ studies by Chittaranjan and colleagues (Chittaranjan et al., 2011, 2013), $n = 2$ studies by Stachl and colleagues (Stachl et al., 2017, 2020) and $n = 2$ studies by Harari and colleagues (Harari, Müller, et al., 2020;

Harari, Vaid, et al., 2020) whose samples showed a partial overlap in recruited participants and personality assessment. These studies were all retained in multilevel meta-analytic calculations by including a second-level indicator indicating they both shared the same source of data in the model. A similar situation could be observed when comparing $n = 2$ studies by Xu and colleagues (Xu, Frey, Fleisch, Ilic, 2016; Xu, Frey, & Ilic, 2016) which presented analyses performed on the same sample. In this case, only one study was included the meta-analysis.

After the removal of papers not including information about effect sizes compatible with correlations, and non-independent studies meeting the exclusion criteria, $n = 18$ papers remained including $n = 21$ studies performed on distinct samples/variables combinations. Note that some of the papers presented multiple studies perfomed on distinct datasets. These studies are reported in bold in Table S1 (Supplementary material), and listed in Figure 2 (Forest plot). The selected studies were performed on samples recruited in United States ($n = 9$), Germany ($n = 6$), China ($n = 2$), Switzerland ($n = 2$), Canada ($n = 1$), and Iran ($n = 1$), with a mean sample size of 386.69 participants (Range = 32–2418). Note that some of the selected studies failed to include detailed information about the distribution of age, gender, or both in the sample ($n = 8$). When information was reported, the mean percentage of male participants in the samples was 46% (Range: 23%–64%), while on average mean participant age was 25.9 years old (Range = 18.9–40.3 years old).

In these selected studies, Big Five traits were assessed using a variety of questionnaires: the Ten Item Personality Inventory (TIPI, Gosling et al., 2003, $n = 2$), NEO-Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992, $n = 2$), Big Five Inventory (BFI-44, John et al., 2008, $n = 9$), BFI-10 (BFI-10; Rammstedt & John, 2007, $n = 1$; or an alternative short BFI version, $n = 1$), Trait Self-Description Inventory (TSDI; Christal, 1994, $n = 1$), and Big Five Structure Inventory ($n = 3$; BFSI; Arendasy, 2009), and the Big Five Inventory 2-Short version ($n = 1$, BFI-2-S, Soto & John, 2017). Finally, one study mentioned using a 60-items Big Five assessment, but failed to provide bibliographic information for the questionnaire (Yu et al., 2019).

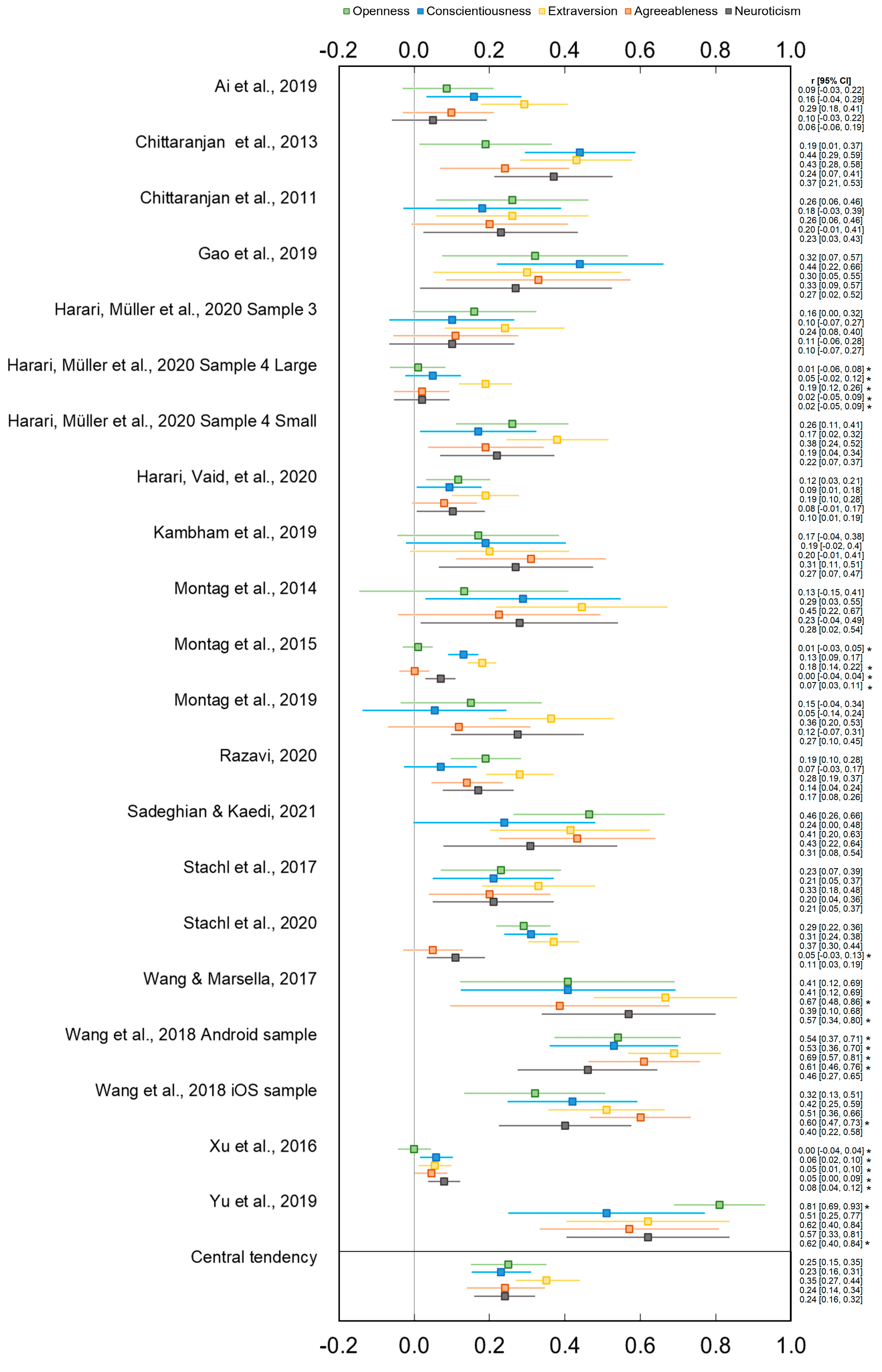
Based on information reported in the papers, $n = 15$ of the $n = 21$ selected studies were performed on participants using Android smartphones only, while remaining studies were performed on participants using a combination of Android and iOS ($n = 2$), iOS only ($n = 1$), or Symbian smartphones ($n = 2$); one study failed to report information about operative systems. Regarding the type of smartphone data, $n = 11$ studies analyzed logs of call and texting behaviors, $n = 12$ analyzed data consisting of usage patterns of smartphone apps, $n = 12$ studies analyzed data derived from use of one or more sensors (e.g., GPS, Wi-Fi, Bluetooth, accelerometer and gyroscope data); also note that some of the selected studies ($n = 9$) analyzed more than one data type. More detailed information about the type of smartphone data presented in each selected study, as well as information about studies sharing the same data source, are reported in the Supplementary material (Table S2).

Among the selected studies, $n = 10$ presented results emerging from both bivariate associations and model-based predictions, $n = 9$ only presented bivariate associations, and $n = 2$ only presented model-based predictions. Collected effect sizes either consisted of bivariate correlations between single features extracted from smartphone data and (continuous) personality scores ($n = 17$ studies), or correlations computed between observed personality scores and model-based (continuous) predictions obtained using a machine-learning approach ($n = 4$ studies, i.e., Stachl et al., 2020; $n = 2$ studies presented in Wang et al., 2018 performed on samples including respectively Android and iOS users; Yu et al., 2019).

3.2 | Meta-analytic computations

3.2.1 | Central tendency

Overall, we examined $n = 105$ distinct effect sizes ($n = 21$ effect sizes per trait), reported in $n = 21$ distinct studies, clustered in $n = 17$ data sources. A forest plot combining all effect sizes included in the meta-analyses and the meta-analytic correlations is presented in Figure 2. For each Big Five personality trait, the estimated meta-analytic correlations are also presented in Table S3 (Supplementary material), alongside Q tests for heterogeneity and the I^2 statistic. All traits showed a significant meta-analytical association with features extracted from smartphone data. Overall, extraversion ($r = .35$, 95% CI [0.27, 0.44]) showed a moderate association with features extracted from smartphone data, while the other traits showed smaller associations (openness: $r = .25$, 95% CI [0.15, 0.35]; conscientiousness: $r = .23$, 95% CI [0.16, 0.31]; agreeableness: ($r = .24$, 95% CI [0.14, 0.34]; and neuroticism: $r = .24$, 95% CI [0.16, 0.32]). Results of the Q tests for heterogeneity were significant for each trait, indicating the presence of non-negligible heterogeneity among the effect sizes. For all traits, observed dispersion of effect sizes was due to true heterogeneity ($I^2 \geq 87.50$) as opposed to sampling variance. In particular, based on model decomposition of effect size variance, we saw that for all traits, most of the heterogeneity was due to variance at the data source level (Range: 66.21% to 83.10%), while variance at the study level was typically lower (Range: 9.69% to 22.19%); the percentage of heterogeneity due to sampling variance was typically



Note. Outlier effect-size values are marked with an asterisk.

FIGURE 2 Forest plot combining effect sizes and central tendency estimates for each of the Big Five traits.

the lowest among the examined sources (Range: 6.39% to 12.49%). For more details about the decomposition of effect size variance see Table S3 (Supplementary material).

Note that Table S3 also reports information about number of outliers detected, and the pooled effect size and heterogeneity statistics computed after removing the outliers. Outliers are marked with an asterisk in Figure 2. In general, removal of outliers only resulted in a slight decrease in the size of meta-analytic correlations (change in $r \leq .03$) that did not affect the overall pattern of correlations.

3.2.2 | Moderator analyses

Because we found significant heterogeneity in effect sizes, we implemented a series of meta-regressions examining the role of potential moderators in explaining variability of the strength of associations between personality and smartphone data. Meta-regressions failed to show significant effects for the publication year of included studies on the association between smartphone data and Big Five traits (Openness: $\beta = .02$, $p = .20$; conscientiousness: $\beta = .00$, $p = .99$; extraversion: $\beta = .01$, $p = .68$; agreeableness: $\beta = .00$, $p = .96$; neuroticism: $\beta = .00$, $p = .75$). In turn, for all traits except for neuroticism, regressions showed an increase in the strength of association between smartphone data and personality traits when effect sizes were computed based on model-based personality predictions as opposed to simple bivariate associations between single features extracted from users' smartphone data and personality traits (Openness: $\beta = .32$, $p = .001$; conscientiousness: $\beta = .24$, $p = .002$; extraversion: $\beta = .24$, $p = .006$; agreeableness: $\beta = .27$, $p = .009$; neuroticism: $\beta = .16$, $p = .07$).

Finally, we examined the effect of the type of smartphone data, namely app usage, sensors and system data, and call and text log data on the association between smartphone data itself and personality and Big Five traits. Meta-regressions failed to showed significant moderating effects for sensors and system data (Openness: $\beta = .14$, $p = .11$; conscientiousness: $\beta = .08$, $p = .20$; extraversion: $\beta = .08$, $p = .26$; agreeableness: $\beta = .08$, $p = .34$; neuroticism: $\beta = .03$, $p = .65$) and app usage data (openness: $\beta = .01$, $p = .96$; conscientiousness: $\beta = .00$, $p = .96$; extraversion: $\beta = -.05$, $p = .55$; agreeableness: $\beta = -.06$, $p = .53$; neuroticism: $\beta = -.03$, $p = .73$). When investigating the impact of using call- and text-log-data, we found a small, significant effect indicating an increase in effect size for the association between smartphone data and extraversion scores, while non-significant effects were observed for the remaining traits (Openness: $\beta = .09$, $p = .17$; conscientiousness: $\beta = .06$, $p = .24$; extraversion: $\beta = .14$, $p = .03$; agreeableness: $\beta = -.01$, $p = .90$; neuroticism: $\beta = .04$, $p = .59$).

3.2.3 | Publication bias

For each Big Five trait, the funnel plot of standard errors versus the correlations was markedly asymmetric (see Supplementary material, Figures S1–S5), suggesting the existence of some form of publication bias. Egger's tests were significant for all traits except for openness (Openness: $p = .126$, conscientiousness: $p = .013$; extraversion: $p = .023$; agreeableness: $p = .002$; neuroticism: $p < .001$). After removing outliers, evidence of publication bias was confirmed only for agreeableness and neuroticism (openness: $p = .307$, conscientiousness: $p = .070$; extraversion: $p = .241$; agreeableness: $p = .001$; neuroticism: $p = .005$). In general, we found evidence that studies performed on small sample sizes reported larger effect sizes than studies employing large samples, resulting in non-negligible publication bias.

Finally, for each trait, the fail-safe N (Openness: $N = 1597$; conscientiousness: $N = 1616$; extraversion $= 4450$; agreeableness: $N = 1212$; neuroticism: $N = 1431$) value was significantly larger than the recommended rule-of-thumb limit ($5 \times$ *21 effect sizes* $+ 10 = 115$; Rosenthal, 1979). Removal of outliers did not alter these results (Openness: $N = 765$; conscientiousness: $N = 1159$; extraversion $= 2078$; agreeableness: $N = 491$; neuroticism: $N = 721$). These findings support the significance of the meta-analytic correlations, thus ruling out the existence of a relevant "file drawer" problem potentially nullifying the emerging correlations.

4 | DISCUSSION

The investigation of personality represents a timely topic, because it is well-known that personality represents an important variable predicting a large range of relevant life variables ranging from job-performance (Barrick & Mount, 1991) to health behavior (Bogg & Roberts, 2004; for newer literature see introduction). Mounting evidence suggests that (a) personality can be predicted from digital footprints and (b) that this kind of data might be of large interest for the marketing industry to conduct micro-targeting and thus improve their selling strategies (for problems of the data business model behind many online services see Montag et al., 2021; Montag, Thrul, et al., 2022). In this context, we also mentioned the term surveillance capitalism (Zuboff, 2019), and refer to newer work considering how to create healthier online platforms (Dhawan et al., 2022). Many studies focused in recent years on predicting personality from web-scraped social media data (Azucar et al., 2018; Peterka-Bonetta et al., 2021), while smartphone data to our knowledge has been investigated to a lesser extent. Neglect of smartphone data is on the one hand surprising, because humans carry their smartphones with them on a 24/7 basis, and therefore this source could provide better insights into a person than social media data, which is only accessed from time to time by people. On the other hand, research in this area is scarcer, because social media data are comparably more easily scraped, although—as mentioned—due to the *Cambridge Analytica data scandal* such data access is now much more restricted for some of the platforms (APIs from Facebook have been closed for independent scientists).

Against this background, the present study aimed to obtain insights into the actual effect sizes between each of the Big Five personality traits and smartphone data (and therefore the potential of digital phenotyping in this area). We observed that extraversion could be best predicted from the smartphone data (r = .35). This is not surprising, because extraverts have stronger urges to socially interact with each other, be it via directly calling or texting other persons via the smartphone (Montag et al., 2019) or using messenger/social media applications (Montag et al., 2015). Instead, the remaining four personality dimensions showed smaller associations with the Big Five, all in the range between 0.23 and 0.25.

Results are compatible with those emerging from studies studying associations between smartphone data and categorical operationalization of Big Five personality in indicating extraversion as the trait showing stronger associations with smartphone-derived features (e.g., de Montjoye et al., 2013; Mønsted et al., 2018). These findings are also coherent with recent studies exploring associations between personality and smartphone features not at the individual level (e.g., exploring association between personality and smartphone activity data using a within person approach, Beierle et al., 2020; Rüegger et al., 2020), highlighting the relevance of extraversion as factor explaining individual differences in indicators of smartphone activity. The stronger association between smartphone data and extraversion when compared with other traits might be related to the specific affordances allowed by smartphones that facilitate communication and sociality. Indeed, results from moderator analyses indicated that use of a specific type of data, namely call and text log data, appears to be instrumental in the prediction of individual differences in extraversion, more so than in predicting the other traits. Other types of data failed to reveal moderation effects. (e.g., sensors and system data; application usage). Note, however, that these limited findings may be due a general lack of diversity among studies, and availability of specific types of data. For the purpose of our study, for example, the relative scarcity of distinct studies documenting findings on data generated by different smartphone sensors, such as GPS, Wi-Fi, Bluetooth, accelerometers, guided us to group them together in a single category. Future meta-analyses including more studies and more fine-grained data types may find additional moderation by data type.

Overall, analysis of moderators also pointed out that the ability to predict personality traits based on smartphone data is improved when multiple features are combined in a single predictive model, as opposed to being based on bivariate associations. This is not unexpected and coherent with findings from meta-analyses exploring the feasibility to predict psychological constructs by mining other types of digital footprints, such as those derived from social media data (Azucar et al., 2018; Settanni et al., 2018). Findings on the overall effect size of associations between Big Five personality traits and smartphone data also show that associations are in a similar range to those emerging from studies based on social media data (i.e., $.2 \leq |r| \leq .4$; Azucar et al., 2018). In turn, it appears that studies predicting personality from smartphone data in general are performed on much smaller samples (mean sample size <1000) than studies relying on digital traces of social media data (mean sample size >10,000; for a review on studies relying on Facebook data see Marengo & Montag, 2020). A possible interpretation of this finding is related to the notion that smaller samples tend to produce larger, biased effect sizes (e.g., Sterne et al., 2000); on the other hand, one could argue that recordings of smartphone activity are inherently more diverse, and in some way richer than data retrievable via social media APIs, as they tap into both non-social (e.g., using Internet browser and setting up a calendar) and social behaviors (e.g., calling, texting, and sharing media) and involve the use of multiple applications and smartphone features, as opposed to the relatively limited set of behaviors one can perform on a single social media platform (also note that data collected via social media APIs typically do not include information about passive use, such as browsing and time spent on the platform). As such, fewer observations may be needed to predict personality from smartphone sensing data when compared to social media data. Still, the massive datasets of unstructured natural language and visual data generated by users' activity on social media remain intuitively more closely related to personality due to the unique view these data provide on users' image, emotions, and values. Future studies comparing smartphone-derived and social media data collected on the same users might help clarify their relative contribution in explaining individual differences in personality.

Finally, regarding publication bias, we found some evidence that pointed toward the existence of a "file drawer" problem affecting the data (i.e., small studies having higher chances of remaining unpublished when reporting small effect sizes). Note that this effect was more pronounced for the Neuroticism and Agreeableness traits, and could not be detected for the Openness. On the other hand, findings from the fail-safe *N* procedure also indicated that, for all traits, the "file drawer" problem detected for some of the traits did not severely affect the significance of emerging findings.

As one can see personality predictions being currently possible from smartphone data are far from perfect. For all traits, the average correlation between predicted and self-report personality scores is lower than what one would expect when correlating multiple assessments of the same individual (i.e., test–retest reliability, see also results by Kosinski et al., 2013), or convergence between personality instruments assessing the same latent construct (e.g., of $r \approx .75$ for short Big Five assessments, Pervin & John, 1999). In keeping with what emerged regarding the predictability of personality traits from digital traces of social media (e.g., Azucar et al., 2018), meta-analytic correlations are close to the expected strength of relationships between personality and behaviors (i.e., also known as "personality coefficient"), in most cases ranging from 0.2 to 0.4 (Back et al., 2009; Mischel, 1968; Roberts et al., 2007). As such, one could argue that their use for accurate individual assessment is currently limited.

On the other hand, smartphone-derived personality prediction data may nonetheless be used by smartphone- and marketing-companies, such as phone carriers and related businesses, to improve and customize user experience. Indeed, conditional on the phone carrier's privacy policy (and the users' privacy settings), phone carriers routinely track information about user activity on their smartphone, including location data, web activity, and application usage, and share these insights with third parties in exchange for monetary reimbursement. Both carriers and third parties may use collected data to personalize the user experience, including targeted advertisement (McAuliffe, 2022). As shown by Matz et al. (2017) in the realm of social media, personality predictions based on the mining of online activity may then improve the effectiveness of mobile advertising by allowing for a more accurate tailoring of the content and time of delivery of mobile advertisements, increasing the chances that the ad(s) will be acted upon by a user. For example, a user high in extraversion may respond positively to flashy, energetic advertisements, while introverted users may prefer more sober, subdued ads. Impulsive users may be more likely to either act on an advertisement or rapidly lose interest, while more conscientious users may require more time to consider an offer before deciding, thus suggesting the need for repeated exposure to an ad for a more extended period. Ideally, tailoring the presentation of ads based on users' personality might be beneficial to promoting products and services via mobile devices. Still, it is worthy to note that literature in this area is scarce and much is still speculative, at least for independent researchers without access to data from phone carriers, social media, and smartphone companies, which in turn have access to massive datasets. As can be easily understood, large privacy issues arise here and regulation is mandatory to protect individuals from data misuse. Aside from these negative aspects smartphone data also bear potential to improve the healthcare system. Sensing personality traits such as neuroticism or conscientiousness could be helpful to provide people with individualized programs to improve health behavior or reach a certain health goal (Alqahtani et al., 2022). In general, we believe that opportunities, as well as dilemmas requiring to be solved arise from digital phenotyping (Montag et al., 2020).

The present study comes with several limitations. As one can see, this research field is still young and although we believe that a meta-analysis is needed at this point to provide researchers with an overview, the available studies conducting personality sensing from smartphone data are still limited. Additionally, most of the study included in meta-analytical calculations were based on data collected on Android smartphones only. While smartphone users using different smartphone operation systems tend to show rather similar characteristics (e.g., Android vs. iPhone users; Götz et al., 2017), the type of data that can be collected from a smartphone depends in part on its operative system. In a few years, more data should be available in order to conduct a more comprehensive meta-analysis with richer literature, both in terms of number of studies, and their diversity.

The presented findings in our work are also limited by a publication bias. It appears that smaller sample sizes yielded larger correlation sizes, something to be thought of when examining our results. Finally, many sensors of the smartphone have not been used so far to predict personality traits and this might result in a growth of effect sizes in this area in the future. Future studies will also assess if facets of personality can be generally better predicted from smartphone data as suggested by Stachl et al. (2020). This could not be investigated in this meta-analysis, because other studies did not investigate facets of the Big Five. In sum, the present meta-analysis shows that personality can be robustly predicted from smartphone data (at least on group level) and this is in particular true for the personality trait extraversion.

AUTHOR CONTRIBUTIONS

CM and DM designed the present study. DM conducted the literature research and the data analysis. DM drafted the first version of the method and result section. CM drafted the first version of the abstract, introduction, and discussion section. JDE critically edited and revised the complete manuscript. All authors approved the final version of this article.

ACKNOWLEDGMENTS

Open Access Funding provided by Universita degli Studi di Torino within the CRUI-CARE Agreement.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ETHICAL APPROVAL

This study is a meta-analytic review of findings from previous studies and did not involve human subjects. No ethical approval was needed from an Institutional Review Board (IRB).

ENDNOTE

[1] Note that although we were able to identify 25 papers, only 18 papers presented studies that were eligible to be included in meta-analytical calculations. See methods.

REFERENCES

Ai, P., Liu, Y., & Zhao, X. (2019). Big Five personality traits predict daily spatial behavior: Evidence from smartphone data. *Personality and Individual Differences*, *147*, 285–291. https://doi.org/10.1016/j.paid.2019.04.027

Alqahtani, F., Meier, S., & Orji, R. (2022). Personality-based approach for tailoring persuasive mental health applications. *User Modeling and User-Adapted Interaction*, *32*(3), 253–295. https://doi.org/10.1007/s11257-021-09289-5

Arendasy, M. (2009). *BFSI: Big-Five Struktur-Inventar (Test & Manual)*. SCHUHFRIED.

Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, *124*, 150–159. https://doi.org/10.1016/j.paid.2017.12.018

Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology*, *97*(3), 533–548. https://doi.org/10.1037/a0016229

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A Meta-analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Baumeister, H., & Montag, C. (Eds.). (2019). *Digital phenotyping and mobile sensing: New developments in psychoinformatics*. Springer International Publishing. https://doi.org/10.1007/978-3-030-31620-4

Beierle, F., Probst, T., Allemand, M., Zimmermann, J., Pryss, R., Neff, P., Schlee, W., Stieger, S., & Budimir, S. (2020). Frequency and duration of daily smartphone usage in relation to personality traits. *Digital Psychology*, *1*(1), 20–28. https://doi.org/10.24989/dp.v1i1.1821

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, *130*(6), 887–919. https://doi.org/10.1037/0033-2909.130.6.887

Burtăverde, V., Vlăsceanu, S., & Avram, E. (2021). Exploring the relationship between personality structure and smartphone usage. *Current Psychology*, *40*(11), 5613–5625. https://doi.org/10.1007/s12144-019-00521-5

Chittaranjan, G., Blom, J., & Gatica-Perez, D. (2011). Who's who with Big-Five: Analyzing and classifying personality traits with smartphones. In *2011 15th Annual International Symposium on Wearable Computers (ISWC)* (pp. 29–36). IEEE. https://doi.org/10.1109/ISWC.2011.29

Chittaranjan, G., Blom, J., & Gatica-Perez, D. (2013). Mining large-scale smartphone data for personality studies. *Personal and Ubiquitous Computing*, *17*(3), 433–450. https://doi.org/10.1007/s00779-011-0490-1

Christal, R. E. (1994). *Non-cognitive research involving systems of testing and learning*. Final research and development status report (USAF Contract No. F33615-91-D-0010). Armstrong Laboratory, Brooks Air Force Base, United States Air Force.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Psychological Assessment Resources.

de Montjoye, Y.-A., Quoidbach, J., Robic, F., & Pentland, A. (Sandy). (2013). Predicting personality using novel mobile phone-based metrics. In A. M. Greenberg, W. G. Kennedy, & N. D. Bos (Eds.), *Social computing, behavioral-cultural modeling and prediction* (pp. 48–55). Springer. https://doi.org/10.1007/978-3-642-37210-0_6

Dhawan, S., Hegelich, S., Sindermann, C., & Montag, C. (2022). Re-start social media, but how? *Telematics and Informatics Reports*, *8*, 100017.

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(44), 11203–11208. https://doi.org/10.1073/pnas.1802331115

Elhai, J. D., Levine, J. C., Dvorak, R. D., & Hall, B. J. (2017). Non-social features of smartphone use are most related to depression, anxiety and problematic smartphone use. *Computers in Human Behavior*, *69*, 75–82. https://doi.org/10.1016/j.chb.2016.12.023

Freudenstein, J.-P., Strauch, C., Mussel, P., & Ziegler, M. (2019). Four personality types may be neither robust nor exhaustive. *Nature Human Behaviour*, *3*(10), 1045–1046. https://doi.org/10.1038/s41562-019-0721-4

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., Griffith, L., Oremus, M., Raina, P., Ismaila, A., Santaguida, P., Lau, J., & Trikalinos, T. A. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology*, *64*(11), 1187–1197. https://doi.org/10.1016/j.jclinepi.2010.08.010

Gao, N., Shao, W., & Salim, F. D. (2019). Predicting personality traits from physical activity intensity. *Computer*, *52*(7), 47–56. https://doi.org/10.1109/MC.2019.2913751

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Götz, F. M., Stieger, S., & Reips, U. D. (2017). Users of the main smartphone operating systems (iOS, Android) differ only little in personality. *PLoS One*, *12*(5), e0176921. https://doi.org/10.1371/journal.pone.0176921

Harari, G. M., Müller, S. R., Stachl, C., Wang, R., Wang, W., Bühner, M., Rentfrow, P. J., Campbell, A. T., & Gosling, S. D. (2020). Sensing sociability: Individual differences in young adults' conversation, calling, texting, and app use behaviors in daily life. *Journal of Personality and Social Psychology*, *119*(1), 204–228. https://doi.org/10.1037/pspp0000245

Harari, G. M., Vaid, S. S., Müller, S. R., Stachl, C., Marrero, Z., Schoedel, R., Bühner, M., & Gosling, S. D. (2020). Personality sensing for theory development and assessment in the digital age. *European Journal of Personality*, *34*(5), 649–669. https://doi.org/10.1002/per.2273

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and Research* (3rd ed.). Guilford Press.

Kambham, N. K., Stanley, K. G., & Bell, S. (2019). *Predicting personality traits using smartphone sensor data and app usage data*. Presented at the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2018. https://doi.org/10.1109/IEMCON.2018.8614854

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Lepri, B., Staiano, J., Shmueli, E., Pianesi, F., & Pentland, A. (2016). The role of personality in shaping social networks and mediating behavioral change. *User Modeling and User-Adapted Interaction*, *26*(2), 143–175. https://doi.org/10.1007/s11257-016-9173-y

Marengo, D., & Montag, C. (2020). Digital phenotyping of Big Five personality via Facebook data mining: a meta-analysis. *Digital Psychology*, *1*(1), 52–64. https://doi.org/10.24989/dp.v1i1.1823

Markowetz, A., Błaszkiewicz, K., Montag, C., Switala, C., & Schlaepfer, T. E. (2014). Psycho-informatics: Big Data shaping modern psychometrics. *Medical Hypotheses*, *82*(4), 405–411. https://doi.org/10.1016/j.mehy.2013.11.030

Matz, S. C., Appel, R. E., & Kosinski, M. (2020). Privacy in the age of psychological targeting. *Current Opinion in Psychology*, *31*, 116–121. https://doi.org/10.1016/j.copsyc.2019.08.010

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(48), 12714–12719. https://doi.org/10.1073/pnas.1710966114

McAuliffe, Z. (2022, April 19). *Don't want your phone carrier tracking your personal data? You can tell it to stop*. CNET. https://www.cnet.com/tech/services-and-software/dont-want-your-phone-carrier-tracking-your-personal-data-you-can-tell-it-to-stop/

McCrae, R. R., Terracciano, A., Costa, P. T., & Ozer, D. J. (2006). Person-factors in the California Adult Q-Set: Closing the door on personality trait types? *European Journal of Personality*, *20*(1), 29–44. https://doi.org/10.1002/per.553

Miller, G. (2012). The Smartphone Psychology Manifesto. *Perspectives on Psychological Science*, *7*(3), 221–237. https://doi.org/10.1177/1745691612441215

Mischel, W. (1968). *Personality and assessment*. Wiley.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1–9. https://doi.org/10.1186/2046-4053-4-1

Mønsted, B., Mollgaard, A., & Mathiesen, J. (2018). Phone-based metric as a predictor for basic personality traits. *Journal of Research in Personality*, *74*, 16–22. https://doi.org/10.1016/j.jrp.2017.12.004

Montag, C., Baumeister, H., Kannen, C., Sariyska, R., Meßner, E.-M., & Brand, M. (2019). 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. *Journal of Multidisciplinary Scientific Journal*, *2*(2), 102–115. https://doi.org/10.3390/j2020008

Montag, C., Błaszkiewicz, K., Lachmann, B., Andone, I., Sariyska, R., Trendafilov, B., Reuter, M., & Markowetz, A. (2014). Correlating personality and actual phone usage. *Journal of Individual Differences*, *35*(3), 158–165. https://doi.org/10.1027/1614-0001/a000139

Montag, C., Błaszkiewicz, K., Sariyska, R., Lachmann, B., Andone, I., Trendafilov, B., Eibes, M., & Markowetz, A. (2015). Smartphone usage in the 21st century: Who is active on WhatsApp? *BMC Research Notes*, *8*(1), 331. https://doi.org/10.1186/s13104-015-1280-z

Montag, C., Dagum, P., Hall, B. J., & Elhai, J. D. (2022). Do we still need psychological self-report questionnaires in the age of the Internet of Things? *Discover Psychology*, *2*(1).

Montag, C., Duke, É., & Markowetz, A. (2016). Toward psychoinformatics: Computer science meets psychology. *Computational and Mathematical Methods in Medicine*, *2016*, 1–10. https://doi.org/10.1155/2016/2983685

Montag, C., & Elhai, J. D. (2019). A new agenda for personality psychology in the digital age? *Personality and Individual Differences*, *147*, 128–134. https://doi.org/10.1016/j.paid.2019.03.045

Montag, C., Hegelich, S., Sindermann, C., Rozgonjuk, D., Marengo, D., & Elhai, J. D. (2021). On corporate responsibility when studying social media use and well-being. *Trends in Cognitive Sciences*, *25*(4), 268–270. https://doi.org/10.1016/j.tics.2021.01.002

Montag, C., Sindermann, C., & Baumeister, H. (2020). Digital phenotyping in psychological and medical sciences: A reflection about necessary prerequisites to reduce harm and increase benefits. *Current Opinion in Psychology*, *36*, 19–24. https://doi.org/10.1016/j.copsyc.2020.03.013

Montag, C., Thrul, J., & van Rooij, A. J. (2022). Social media companies or their users—Which party needs to change to reduce online time? *Addiction*, *117*(8), 2363–2364. https://doi.org/10.1111/add.15946

Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421. https://doi.org/10.1146/annurev.psych.57.102904.190127

Peltonen, E., Sharmila, P., Opoku Asare, K., Visuri, A., Lagerspetz, E., & Ferreira, D. (2020). When phones get personal: Predicting Big Five personality traits from application usage. *Pervasive* and Mobile Computing, 69, 101269. https://doi.org/10.1016/j.pmcj.2020.101269

Pervin, L. A., & John, O. P. (Eds.). (1999). *Handbook of personality: Theory and research*. Elsevier.

Peterka-Bonetta, J., Sindermann, C., Elhai, J. D., & Montag, C. (2021). How objectively measured Twitter and Instagram use relate to self-reported personality and tendencies toward Internet/Smartphone Use Disorder. *Human Behavior and Emerging Technologies*, *3*(5), 898–911. https://doi.org/10.1002/hbe2.243

Pittenger, D. J. (2004). The limitations of extracting typologies from trait measures of personality. *Personality and Individual Differences*, *37*(4), 779–787. https://doi.org/10.1016/j.paid.2003.10.006

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Razavi, R. (2020). Personality segmentation of users through mining their mobile usage patterns. *International Journal of Human-Computer Studies*, *143*, 102470. https://doi.org/10.1016/j.ijhcs.2020.102470

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rosenthal, R. (1994). Parametric measures of effect size. In *The handbook of research synthesis* (pp. 231–244). Russell Sage Foundation.

Rüegger, D., Stieger, M., Nißen, M., Allemand, M., Fleisch, E., & Kowatsch, T. (2020). How are personality states associated with smartphone data? *European Journal of Personality*, *34*(5), 687–713. https://doi.org/10.1002/per.2309

Sariyska, R., & Montag, C. (2019). An overview on doing psychodiagnostics in personality psychology and tracking physical activity via smartphones (pp. 45–63). https://doi.org/10.1007/978-3-030-31620-4_4

Servia-Rodríguez, S., Rachuri, K. K., Mascolo, C., Rentfrow, P. J., Lathia, N., & Sandstrom, G. M. (2017). Mobile sensing at the service of mental well-being: A large-scale longitudinal study. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 103–112). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/3038912.3052618

Settanni, M., Azucar, D., & Marengo, D. (2018). Predicting individual characteristics from digital traces on social media: A meta-analysis. *Cyberpsychology, Behavior and Social Networking*, *21*(4), 217–228. https://doi.org/10.1089/cyber.2017.0384

Soto, C. J. (2019). How replicable are links between personality traits and consequential life outcomes? The life outcomes of personality replication project. *Psychological Science*, *30*(5), 711–727. https://doi.org/10.1177/0956797619831612

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81. https://doi.org/10.1016/j.jrp.2017.02.004

Stachl, C., Au, Q., Schoedel, R., Gosling, S. D., Harari, G. M., Buschek, D., Völkel, S. T., Schuwerk, T., Oldemeier, M., Ullmann, T., Hussmann, H., Bischl, B., & Bühner, M. (2020). Predicting personality from patterns of behavior collected with smartphones. *Proceedings of the National Academy of Sciences*, *117*(30), 17680–17687. https://doi.org/10.1073/pnas.1920484117

Stachl, C., Hilbert, S., Au, J.-Q., Buschek, D., Luca, A. D., Bischl, B., Hussmann, H., & Bühner, M. (2017). Personality traits predict smartphone usage. *European Journal of Personality*, *31*(6), 701–722. https://doi.org/10.1002/per.2113

Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N., & Pentland, A. (2012). Friends don't lie: Inferring personality traits from social network structure. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp. 321–330). https://doi.org/10.1145/2370216.2370266

Sterne, J. A., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, *53*(11), 1119–1129. https://doi.org/10.1016/S0895-4356(00)00242-0

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*, *54*(10), 1046–1055. https://doi.org/10.1016/S0895-4356(01)00377-8

Strickhouser, J. E., Zell, E., & Krizan, Z. (2017). Does personality predict health and well-being? A metasynthesis. *Health Psychology*, *36*(8), 797–810. https://doi.org/10.1037/hea0000475

Sutin, A. R., Stephan, Y., Luchetti, M., Artese, A., Oshio, A., & Terracciano, A. (2016). The five-factor model of personality and physical inactivity: A meta-analysis of 16 samples. *Journal of Research in Personality*, *63*, 22–28. https://doi.org/10.1037/hea0000475

Viechtbauer, W. (2007). Accounting for heterogeneity via random-effects models and moderator analyses in meta-analysis. *Zeitschrift für Psychologie/Journal of Psychology*, *215*(2), 104–121. https://doi.org/10.1027/0044-3409.215.2.104

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Wang, H., & Marsella, S. (2017). *Assessing personality through objective behavioral sensing*. Presented at the 2017 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017. https://doi.org/10.1109/ACII.2017.8273590

Wang, W., Harari, G. M., Wang, R., Müller, S. R., Mirjafari, S., Masaba, K., & Campbell, A. T. (2018). Sensing behavioral change over time: Using within-person variability features from mobile sensing to predict personality traits. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 1–21. https://doi.org/10.1145/3264951

Xu, R., Frey, R. M., Fleisch, E., & Ilic, A. (2016). Understanding the impact of personality traits on mobile app adoption - Insights from a large-scale field study. *Computers in Human Behavior*, *62*, 244–256. https://doi.org/10.1016/j.chb.2016.04.011

Xu, R., Frey, R. M., & Ilic, A. (2016). Individual differences and mobile service adoption: An empirical analysis. In *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)* (pp. 234–243). https://doi.org/10.1109/BigDataService.2016.15

Yarkoni, T. (2012). Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science*, *21*(6), 391–397. https://doi.org/10.1177/0963721412457362

Yu, Z., Xu, E., Du, H., Guo, B., & Yao, L. (2019). Inferring user profile attributes from multidimensional mobile phone sensory data. *IEEE Internet of Things Journal*, *6*(3), 5152–5162. https://doi.org/10.1109/JIOT.2019.2897334

Zell, E., & Lesick, T. L. (2022). Big Five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality*, *90*(4), 559–573. https://doi.org/10.1111/jopy.12683

Zuboff, S. (2019, January). Surveillance capitalism and the challenge of collective action. In *New labor forum* (Vol. *28*, pp. 10–29). SAGE Publications.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Marengo, D., Elhai, J. D., & Montag, C. (2023). Predicting Big Five personality traits from smartphone data: A meta-analysis on the potential of digital phenotyping. *Journal of Personality*, *00*, 1–15. https://doi.org/10.1111/jopy.12817