Digital Phenotyping of Big Five Personality Traits via Facebook Data Mining: A Meta-Analysis

Davide Marengo[1*] & Christian Montag[2]

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

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

Abstract

Background: About 2.5 billion people around the world currently have an active account on Facebook. By interacting with Facebook, users generate a vast dataset of information with potential links to psychological and behavioral characteristics. In particular, several researchers have already demonstrated that it is feasible to predict personality from activity logs, posted text, or "Like" behaviors on Facebook.
Objectives: In this study, we carried out a meta-analysis of the available literature on predicting personality from Facebook data.
Methods: Meta-analysis computations were performed using a multilevel approach.
Results: Results showed that, on average, the accuracy of prediction of user personality scores through the mining of Facebook data is moderate (r = .34). However, prediction accuracy was improved when models included demographic variables, and multiple types of digital footprints.
Discussions: Currently, generating personality predictions from Facebook data is feasible, but accuracy is at best moderate. Therefore, current predictions cannot be used for assessment purposes at the individual level, but may provide useful information when conducting group-level assessments. However, prediction accuracy is expected to improve as larger datasets and new types of data are mined for prediction purposes.

Keywords: social media, personality, Facebook, digital phenotyping, psychoinformatics

1 Introduction

Use of social media platforms is widespread, particularly amongst young people (Perrin, & Anderson, 2019). Among existing platforms, Facebook remains the leading social media platform in terms of active users (2.45 billion monthly active users as of third quarter of 2019 (Rabe, 2019)). Every day online users come to Facebook and share content, such as text, pictures and videos, which can be liked, commented upon, or shared by other users. This interactive process produces a massive dataset of user-generated data, also referred to as "digital footprints" "digital records", or "digital traces", with significant connections to users' behavioral and psychosocial characteristics (e.g. Settanni & Marengo, 2015; Marengo, Azucar, Longobardi & Settanni, 2020; Marengo, Azucar, Giannotta, Basile, & Settanni, 2019), including personality (Azucar, Marengo & Settanni, 2018). These digital footprints can be downloaded and mined to gain insight about users' characteristics, interests, and online and offline behaviors (Kosinski, Matz, Gosling, Popov, & Stillwell, 2015). Research in this field of study, typically referred to as *Psychoinformatics*, uses methods derived from both

psychology and computer science (Montag, Duke, Markowetz, 2016; Yarkoni, 2012) to improve the collection and analysis of psychosocial data, including datasets from mobile devices and online social networks.

Concerning in particular Facebook, based on the analysis of digital traces left by users on the platform it has been possible to develop predictive models detecting demographic variables and psychological characteristics, sometimes with remarkable accuracy (Kosinski et al., 2013; Montag, Duke, & Markowetz, 2016). In this study, when referring to such predictive models we mean to establish links between Facebook activity logs, text, pictures and individual traits, as opposed to using predictive models for explanatory purposes (e.g. theory building and testing, Yarkoni, & Westfall, 2017). Mining Facebook data has been shown to be especially beneficial for the purpose of personality prediction (Azucar, Marengo & Settanni, 2018), to the extent that computer-based personality predictions have been shown to be more accurate than those made by close acquaintances of the users (e.g. friends, and relatives, Youyou, Kosinski, Stillwell, 2015). Overall, findings from meta-analyses have shown that the overall predictive power of social media data for users' personality

is moderate, with correlations between observed and predicted personality scores ranging from .30 to .40 (Azucar, Marengo & Settanni, 2018).

Typically, studies investigating the use of Facebook data for personality prediction employ a common methodological approach. First, researchers collect information about users' personality scores by administering validated self-report personality questionnaires, with the large majority of studies focusing on personality traits drawn from the Big Five/Five Factor model (McCrae & Costa, 1987; McCrae & John, 1992). Next, having obtained authorization from users and from Facebook (Facebook for Developers, 2019) to access user data, digital footprints are collected and processed to extract predictive features based on a variety of approaches, depending on the nature of the data collected (e.g. demographic data, activity information, Likes, texts, or pictures). Such predictive features include *categorical variables* giving insights into socio-demographics (e.g. age group, gender, education level), and *count variables* representing frequency of online activities (e.g. number of posts, pictures, and videos posted in a specific time frame). Beyond this, *count variables* such as the number of Likes expressed to specific online pages, as well features representing topics, words, and phrases naturally occurring in posted text (e.g. open-vocabulary features, Schwartz et al., 2013), or visual features in posted pictures (e.g. facial expressions, style of make-up, hair style, etc., Torfarson, Agustsson, Rothe Timofte, 2016) are also studied. Next, predictive analyses are performed using a machine-learning approach to study the feasibility of using the features extracted from digital footprints to predict users' personality scores as derived from self-report questionnaires. Different models, varying in relation to the number and type of features examined, are compared based on the accuracy of predictor scores compared to self-report scores, typically by examining the correlation between predicted and observed scores, and/or by looking at absolute measures such as the mean absolute error (MAE). Based on these metrics, the best performing models are retained. Online services based on predictive models developed using this approach are now available for both research and commercial purposes (e.g. Apply Magic Sauce, https://applymagicsauce.com, IBM Watson Personality Insight, https://www.ibm.com/watson/services/personality-insights/). These services can be used to generate unobtrusive personality predictions for individual users uniquely based on their digital footprints on Facebook (and other social media platforms).

As noted above, a recent meta-analytic study has established the overall strength of association between social media data and Big Five personality traits (Azucar, Marengo & Settanni, 2018). However, the meta-analysis by Azucar and colleagues (2018) included data from studies presenting only correlation coefficients describing associations between single indicators of social media activity and personality scores (e.g. Gosling, Augustine, Vazire, Holtzman, Gaddis, 2011; Kern, et al., 2014). For this reason, the results do not strictly apply to studies developing predictive models based on more than one digital variable for

personality prediction. Another limitation of the study by Azucar and colleagues (2018) is that, in order to deal with the non-independence of studies sharing a common source of data (e.g. MyPersonality dataset, Kosinski, Stillwell, & Graepel, 2013), results from many studies were not included in the meta-analysis. Finally, the majority of the included studies were published before 2017, whereas many new papers which have been published since now also have to be considered.

Based on these considerations, in this article we present an update of this meta-analytic study which aims to determine the overall predictive accuracy of model-based personality predictions performed using digital footprints on Facebook. Building on previous findings by Azucar and colleagues (2018), we focus our analysis on studies performing predictions of Big Five personality traits because they represent the large majority of existing studies. Further, we only include studies analyzing Facebook data. In limiting our scope to these studies, we aim to provide a clearer view of the potential of mining Facebook data for the prediction of Big Five personality traits. Analyses are performed using a multilevel meta-analytic approach, allowing for non-independent studies (i.e. studies sharing the same data source) to be included in a single analysis, therefore retaining important information which would be excluded using traditional meta-analytic approaches.

2 Methods

2.1 Selection of literature

We started by searching for research papers examining the relationship between Big Five personality traits and digital footprints. A two-step procedure was followed, building on previous work by Azucar and colleagues (2018). First, all the records (n = 789) screened by Azucar and colleagues (2018) were obtained. Next, we applied the same literature search strategy employed by Azucar and colleagues (2018) to identify newly published papers. More specifically, we used the same keyword search strategy used by the authors to investigate the Scopus, ISI Web of Science, Pubmed, and Proquest databases. Combined, the searches performed on the databases resulted in a total of n = 935 unique papers. After removing records overlapping with those screened by Azucar and colleagues (2018), this approach resulted in 146 new papers which were eligible for selection. The original literature search was performed in July 2018. Papers selected from Azucar and colleagues (2018) (n = 24), and those identified through the new search (n = 146), were screened according to the following inclusion criteria: 1. studies must analyze digital footprints collected on Facebook; 2. studies must present results of models predicting Big Five personality traits at the individual level based on digital footprints; 3; studies must include personality scores based on self-report measures of Big Five personality traits (i.e. openness to new experiences, conscientiousness, extraversion, agreeableness, and neuroticism;

OCEAN model); 4. studies must report information about the accuracy of prediction of Big Five personality traits using correlations, or provide information that could be used to compute correlations.

Ultimately, based on the aforementioned criteria, n = 14 papers out of the n = 24 identified by Azucar and colleagues (2018) were selected. In selecting papers from this source, we excluded papers that do not focus on Facebook data (n = 7), and that do not present model-based predictions (n = 3; i.e. Gosling et al. 2011; Kern, et al., 2014; Quercia, Lambiotte, Stillwell, Kosinski & Crowcroft, 2012). The n = 14 papers selected from Azucar and colleagues (2018) included n = 8 papers which, although deemed eligible for inclusion by the authors, ultimately were not included in the analyses presented in Azucar and colleagues (2018) because they were based on non-independent samples derived from a common data source (i.e. the dataset by Golbeck et al. 2011, and the MyPersonality dataset). For the purpose of the present study, the use of a multivelel analytical approach allowed us to retain these papers in the analysis.

With regard to the papers gathered through literature search (n = 146), after removing review papers (n = 4; Azucar, Marengo & Settanni, 2018; Ihsan & Furnham, 2018; Hinds & Joinson,

2019; Settanni, Azucar, Marengo, 2018), papers that do not investigate the link between digital footprints and personality (n = 100), and papers that did not focus on Facebook data (n = 30), n = 12 eligible papers were identified from the literature search. Among those removed were n = 5 papers that did not include effect-sizes which could be transformed into a correlation coefficient (i.e. papers reporting results using mean absolute error (MAE) and root mean square error (RMSE) statistics, Al Marouf, Hasan, Mahmud, 2019; Tadesse, Lin, Xu, Yang, 2018; Tandera, Suhartono, Wongso,Prasetio, 2017; Yulianto, Girsang, Rumagit, 2018; Zhong, Guo, Gao, Shan, Xue, 2018). This approach produced a set of 21 unique papers, of which n = 14 overlap with those selected in Azucar and colleagues (2018), and n = 7 newly selected papers. A flow diagram representing the study selection process is presented in Figure 1. Because in some cases papers included more than one study (i.e. predictions are performed on different datasets within the same paper), effect-sizes were extracted from 23 distinct studies, of which 16 studies were previously analyzed by Azucar and colleagues (2018) and 7 were identified by the new literature search. The characteristics of the selected studies are presented in Table 1, along with collected effect-sizes.
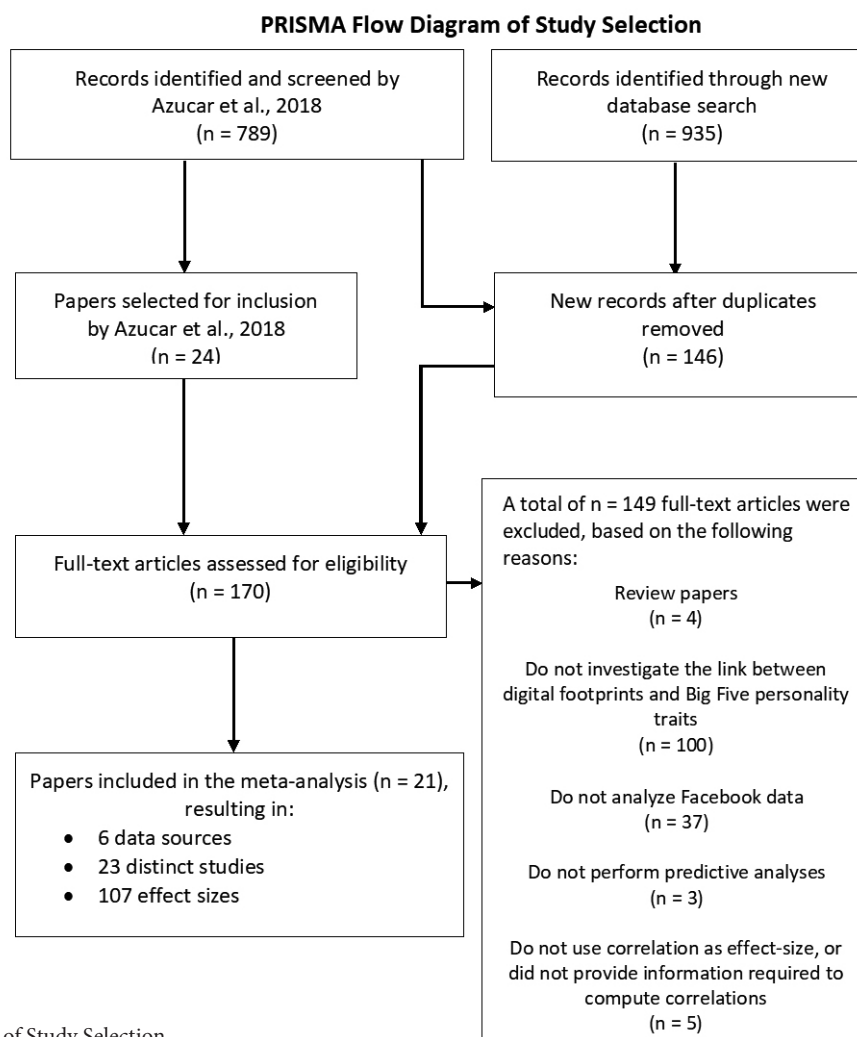

Figure 1. Flow Diagram of Study Selection.

2.2 Coding of studies

Because the studies varied considerably in relation to the type of the digital footprints collected from Facebook and mined for prediction, as well as in the approach used to validate predictions, they were coded using the strategy employed in Azucar and colleagues (2018). Specifically, concerning the type of examined digital footprints, studies were coded based on inclusion (1 = yes, 0 = no) in the analyses of specific types of digital footprints, defined based on their content: 1) user demographics (typically extracted from Facebook personal information section, including gender, age, education, etc.); 2) activity statistics (e.g. number of posts, number of friends or network density, number of received Likes, comments, and user tags); 3) Likes (e.g. Likes expressed to specific Facebook pages); 4) features derived from the analysis of language in text (e.g. features extracted using closed- and/or open-vocabulary approaches); 5) features derived from pictures (e.g. features extracted from uploaded pictures); 6) use of multiple vs. a single type of digital footprints. Additionally, we coded the selected studies based on the approach used to validate the results of predictive models. In this context, model validation refers to the step taken by researchers to determine the accuracy of trained models on new, unseen observations. Different validation approaches exist in this field (for a review, see Marengo & Settanni, 2020), including the *holdout* method and the *k*-fold validation method. Using the *holdout* method, a random split is performed on the data so that two datasets – a larger training set and a smaller test set – are obtained. Then, models are first applied to the training set, and later trained on the smaller test set to evaluate their accuracy. Similarly, the *k*-fold method also involves randomly splitting the data in a training set and a test set, but this process is repeated *k* times resulting in *k* random train/test splits. Analyses are then performed on each of the split, resulting in *k* sets of results which are combined to produce a single accuracy estimate (Hastie, Tibshirani, R., & Friedman, 2009). Here, in coding studies based on the validation approach, we distinguished between studies reporting effect-sizes as computed on the same dataset used to train the model (no validation condition = 0), and studies performing some form of cross-validation of trained models (cross-validation of results = 1, i.e. *holdout* method, or use of *k*-fold cross-validation).

2.3 Strategy of analysis

For each study, we collected the effect-sizes expressing the accuracy of prediction of Big Five personality traits based on the tested predictive models, selecting only a single effect-size per trait. For the purpose of performing the meta-analysis, Pearson's correlation coefficient was used as the effect-size of choice. In the event that a study did not report correlations but other types of effect-size, we used available information to compute correlations using the same approach (for details, see Azucar Marengo, & Settanni, 2018). In the event that a study reported results for more than one predictive model for a single trait (e.g. studies in which models with different set of predictors are compared), the effect size of the best performing model was included in the analysis. Based on this approach, there were 107 distinct effect-sizes for 23 studies (see Table 1). All studies reported effect-sizes for each of the five traits, except for one study which investigated only extraversion (Baik, Lee, Lee, Kim, Choi, 2016) and one study reporting only the average effect size across all Big Five personality traits (Torfason et al., 2016).

Next, meta-analysis computations were performed using a multilevel approach (Cheung, 2014; Van den Noortgate, López-López, Marín-Martínez, Sánchez-Meca, 2015). This approach was used because of the presence of non-independence in our data due to many studies demonstrating more than one effect-size, and sharing the same data source. Indeed, as shown in Table 1, n =17 studies were performed on data sourced from the MyPersonality dataset (Kosinski, Stillwell, & Graepel, 2013), n = 2 studies shared the same dataset used in Golbeck, Robles and Turner (2011), while n = 4 used independent datasets. To give an example, non-independent studies may show a certain degree of overlap in observations (i.e. use samples derived from the same data source, e.g. MyPersonality dataset), but use different types of digital footprints (e.g. language data vs. Likes) to perform predictions.

Using a multilevel approach, variability in effect-sizes due to different variance components is modeled using random effects. For the purpose of the present study, we employ a four-level meta-analytic model modeling four different variance components: at level 1, we model 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 estimate effect-sizes); at level 2, we model the variance existing between effect sizes extracted from the same study (within-study variance); at level 3, we model the variance at the study-level (between-study variance); and at level 4, we account for the variance related to data sources. This model is computed in order to estimate the overall meta-analytic correlation between Big Five personality scores and scores generated by predictive models based on Facebook data, while controlling for different sources of variability. In our dataset, we distinguish between 107 unique effect sizes (level 2) clustered in 23 distinct studies (level 3), and 6 data sources (level 4). In keeping with Schmidt and Hunter (2014), for the purpose of estimating the overall meta-analytic effect-size, correlations were not transformed using *Fisher's z* transformation. Such conversion is not indicated for meta-analytic random-effects models, because they yield an upward bias in the estimation of the average correlation. The distribution of variance over the four levels of the model was examined using the approach described by Assink & Wibbelink (2016), which takes advantage of the formula for estimating study sampling variance proposed by Cheung (2014, p. 215, formula 14). Overall heterogeneity of effect-sizes was examined by using the Q test. The significance of within-study variance (level 2), between-study variance (level 3), and variance due to the specific data source (level 4) was de-

Table 1. Characteristics of studies included in the meta-analysis.

Data source study	Effect-size					Sample size	Cross-validation	Type of digital footprints
	O	C	E	A	N			
Independent datasets								
Baik et al., 2016	–	–	0.42	–	–	565	k-fold	Demographics, Usage stats, Likes
Celli et al., 2014	0.07	0.06	0.18	0.26	0.19	89	Holdout	Pictures
Kleanthous et al., 2016	0.26	0.03	0.28	-0.16	-0.01	62	No cross-validation	Usage Stats
Wald et al., 2012	0.77	0.61	0.68	0.70	0.61	537	No cross-validation	Demographics, Usage Stats, Language
Golbeck et al., 2011 Dataset								
Golbeck et al., 2011	0.65	0.60	0.55	0.48	0.53	167	k-fold	Demographics, Usage Stats , Language
Golbeck, 2016 Study 3	-0.35	-0.07	0.24	-0.35	-0.18	69	No cross-validation	Language
MyPersonality dataset								
Bachrach et al., 2012	0.33	0.41	0.57	0.10	0.51	5000	k-fold	Usage Stats
Cutler & Culis, 2018	0.41	0.35	0.38	0.30	0.32	84451	Holdout	Language
Farnadi et al., 2016 Study 1	0.19	0.24	0.27	0.16	0.24	3731	k-fold	Demographics, Usage Stats , Language
Farnadi et al., 2018	0.26	0.19	0.16	0.11	0.14	5670	k-fold	Likes, Language, Pictures
Golbeck, 2016 Study 1	0.36	0.25	0.37	0.41	0.38	127	No cross-validation	Language
Golbeck, 2016 Study 2	0.20	0.20	0.22	0.24	0.18	8569	No cross-validation	Language
Kosinski et al., 2013	0.43	0.29	0.40	0.30	0.30	54373	k-fold	Likes
Kosinski et al., 2014	0.11	0.16	0.31	0.05	0.23	9515 – 45565	k-fold	Usage Stats
Laleh & Shahram, 2017	0.38	0.29	0.34	0.22	0.27	92225	Holdout	Likes
Markovikj et al., 2013	0.71	0.71	0.70	0.60	0.59	250	No cross-validation	Demographics, Usage Stats , Language
Nave et al., 2018	0.30	0.19	0.21	0.17	0.18	21929	k-fold	Likes
Park et al., 2015	0.43	0.37	0.42	0.35	0.35	4824	Holdout	Language
Schwartz et al., 2013	0.42	0.35	0.38	0.31	0.31	18177	Holdout	Language
Thilakaratne et al., 2016	0.36	0.40	0.44	0.30	0.39	344 – 387	k-fold	Language
Torfason et al., 2016*	–	–	–	–	–	51617	k-fold	Likes, Pictures
Youyou et al., 2015	0.51	0.42	0.45	0.38	0.40	1919	k-fold	Likes
Zhang et al., 2018	0.40	0.35	0.36	0.29	0.32	55835	Holdout	Language

Note. Studies in plain text were selected from Azucar et al., 2018 (n =16). Studies in bold were selected through literature search (n = 7). O = Openness, C = Conscientiousness, E = Extraversion, A = Agreeableness, N = Neuroticism. * The study only reported an average effect-size.

termined using log-likelihood-ratio tests. Using these tests, we compared the model in which the variance at each level (2, 3, and 4) is freely estimated, with an additional model in which the variance for each level was iteratively fixed at zero, while letting the variance for the other levels be freely estimated.

It is worth noting that, by using this multilevel approach, we can provide a single correlation representing the overall predictive power of Facebook data to estimate personality as assessed using the Big Five model. However, potential differences in prediction accuracy across Big Five personality traits can be investigated by way of moderation, i.e. by including a categorical indicator grouping effect-sizes based on the relative Big Five personality trait as fixed effect in the multilevel model, and performing (Bonferroni corrected) pairwise contrasts between estimated correlations for each trait. Next, we examined the

following moderating effects by using dichotomous indicators (1 = yes; 0 = no): (1) use of demographic data; (2) use of activity statistics; (3) use of Likes; (4) use of language features; and (5) use of multiple vs. single type of digital footprints. Finally, we looked at possible differences in estimated effect size based on (6) cross-validation of model results. Moderators are tested separately by including the above mentioned indicators in the model as fixed effects, while accounting for all sources of non-independence with random effects. As only n = 3 of included studies explored use of pictures as a data source, a moderator for this type of data was not included as we did not expect to reach an adequate level of statistical power. For each moderator, an assessment was made of how much incremental variance could be explained by its inclusion in the model.

Finally, we looked at possible publication bias in reported ef-

fect-sizes. More specifically, we examined: 1) asymmetry of the funnel plot visualizing the association between collected effect sizes and their associated standard errors; and 2) significance of a modified Egger's regression test (Egger, Smith, Schneider & Minder, 1997) computed by including the standard error as a predictor of effect sizes in the multilevel model. In this context, the funnel plot was generated as a scatterplot of the correlations between observed and predicted scores collected from each study plotted against their standard error, while Egger's regression test provided an estimate of the asymmetry of the scatterplot. Because the standard error of a study is a measure of (lack of) precision in estimating effects (i.e. lower standard errors indicate higher precision of the effect size estimate), publication bias might be present if less precise studies tend to show higher effect sizes than more precise studies (i.e. standard error is found to positively predict effect-size).

All analyses were performed in *R* using the *metafor* package (Viechtbauer, 2010) by adapting the code provided by Assink and Wibbelink (2016) to a four-level multilevel meta-analytic model (code is provided as Supplementary Material).

3 Results

3.1 Central tendency of effect-sizes

Information about study effect-sizes, as well characteristics of selected studies are reported in Table 1. Figure 2 shows the forest plot of collected effect-sizes. The estimated overall meta-analytic correlation emerging from selected studies for digital footprints predicting Big Five personality traits was 0.34 (SE = 0.043; 95% CI: 0.26–0.43). The result of the Q test for heterogeneity was meaningful (QE (106) = 185879.73, p < .001), indicating significant heterogeneity existed among the effect-sizes. However, based on the estimated proportion of sampling variance per level of the model, it emerged that only 0.08 percent of the total variance can be traced back to variance at level 1 (i.e. sampling variance). Rather, 10.78 percent of the total variance can be attributed to differences between effect sizes from the same study at level 2 (i.e. within-study variance). Further, at 89.14 percent, the largest portion of variance can be traced back to between-study differences at level 3 (i.e. between-study variance). Finally, the portion of variance that could be attributed

to level 4 (i.e. variance due to different data sources) was < 0.01 percent. Accordingly, based on significance of log-likelihood tests, it emerged that both within-study variance (i.e. variability in effect-sizes extracted from the same study; (χ^2 (3) = 6828.71, p < .001) and between-study variance (χ^2 (3) = 97.2939, p<.001) represent a significant source of effect-size heterogeneity, while variance due to data source is not (χ^2 (3) < 0.01, p > .99).

Next, we take a more detailed look at differences among personality traits in estimated effect-size. Results indicate significant differences exist in the prediction accuracy of different personality traits (F (4, 102) = 9.34, p < .001). Based on the estimated effect-sizes and relative 95% confidence interval for each Big Five trait, extraversion (0.39 [0.30, 0.48]) shows the highest overall prediction accuracy, followed by openness (r = 0.38, [0.29, 0.47]), conscientiousness (r = .34 [0.24, 0.43]), neuroticism (r = 0.33 [.23, 0.42]), and agreeableness (r = 0.28 [0.19, 0.38]). However, when looking at (Bonferroni corrected) pairwise contrasts, we identify few significant contrasts: agreeableness can be less accurately predicted from the digital footprints on Facebook than extraversion (contrast = 0.11 [0.07, 0.15], p<.05), and openness (contrast = 0.10 [0.06, 0.14], p<.05).

3.2 Moderator analyses

Table 2 presents the results of moderator analyses concerning the type of digital footprints used for prediction and the approach used to validate results. Only two moderators showed a significant effect. Use of multiple types of digital footprints, as opposed to a single type, was linked to a significant increase in the predictive power of models. Use of demographic variables also showed a positive effect on predictive power. The remaining moderators did not show significant effects.

3.3 Publication bias

The investigation of publication bias via visualization of funnel plot and Egger's test provided interesting results. The funnel plot is presented in Figure 3. It is easy to see that the distribution of effect sizes is asymmetrical, with a clear pattern showing that, at least for a subgroup of estimates, the standard error of effect sizes is negatively related to the magnitude of the effect-size. Simi-

Table 2. Result of moderation analyses: effect of type of digital footprints and validation approach on prediction accuracy

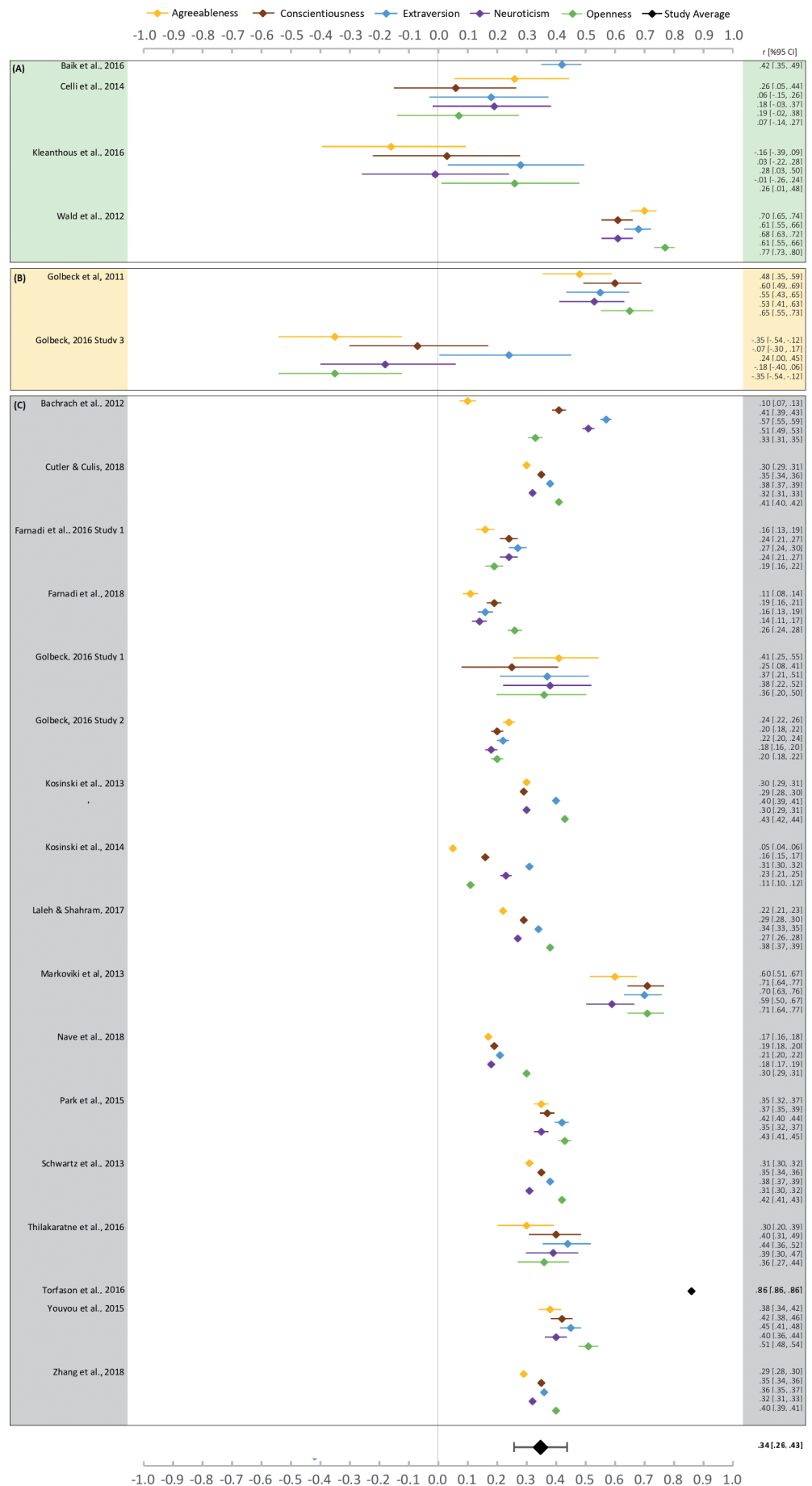| Moderator	B [95% CI]	SE	t	p	Explained variance
Use of demographics	0.24 [0.05, 0.44]	0.10	2.45	0.02	.07
Use of activity statistics	0.10 [-0.09, 0.28]	0.09	1.05	0.30	.00
Use of Facebook Likes	0.06 [-0.13, 0.25]	0.10	0.63	0.53	.00
Use of language features	0.02 [-0.16, 0.19]	0.09	0.18	0.86	.00
Use of multiple types of digital footprints	0.24 [0.08, 0.41]	0.08	2.94	<.01	.17
Cross-validation of model results	0.04 [-0.16, 0.24]	0.10	0.40	0.69	.00

Figure 2.

Forest-plot of study effect-sizes.

Note. Studies in panel A (green box) are independent studies. Studies in panel B (yellow box) are based on data from Golbeck et al, 2011. Studies in panel C (grey box) are based on data from the myPersonality project.
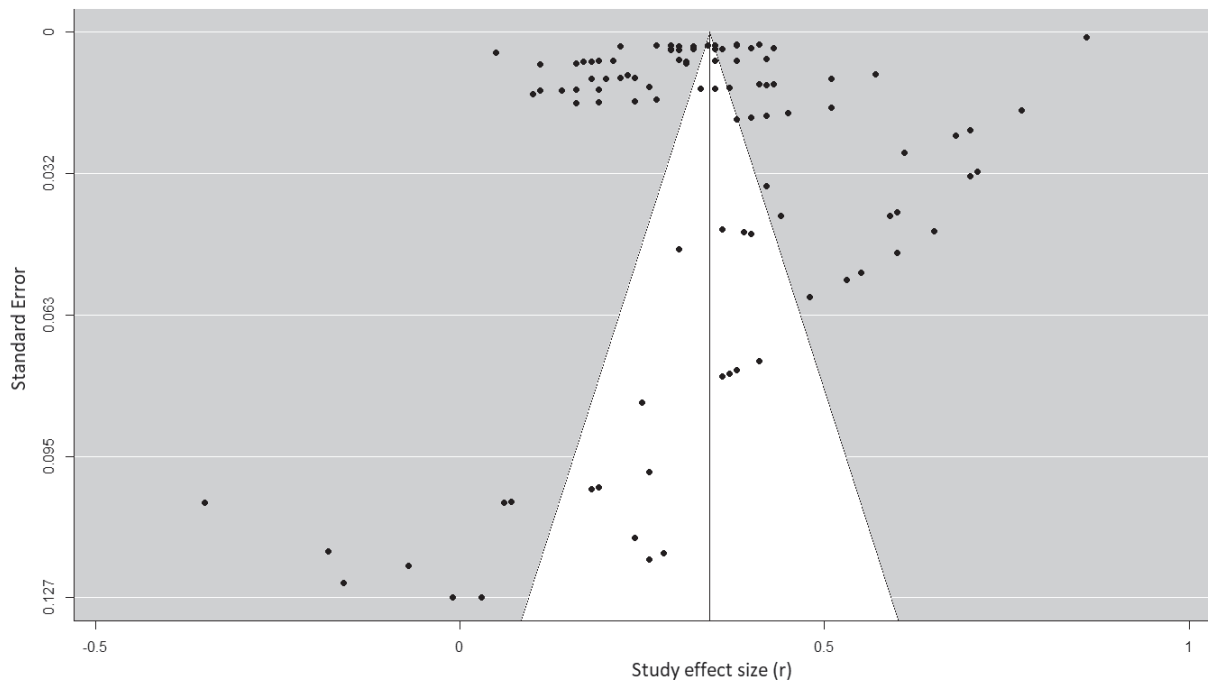
Figure 3. Funnel plot of study effect sizes by relative standard errors.

Note. Studies at the top of the funnel plot (Standard error ≤ .02) are based on sample size ≥ 1000.

larly, Egger's test was coherent in highlighting a small, negative association between magnitude of effect size estimate for digital footprints predicting personality traits and the standard error of the estimate (B = -2.60 [-4.82, -0.38], SE = 1.19, t = 2.33, p = 0.02, explained variance = .04). It is worth noting that the direction of this emerging effect is the opposite of what would be expected based on the hypothesis of publication bias (i.e. a positive association between standard error, and magnitude effect-size). Instead, the emerging effect seems to indicate that, in published studies, the accuracy of personality predictions tends to increase with precision of estimates (i.e. the inverse of standard error).

4 Discussion

In this study, we presented a meta-analysis of research exploring the feasibility of mining digital footprints of Facebook users for the prediction of Big Five personality traits. We built on a previous meta-analytic study by Azucar and colleagues (2018), including newer studies and employing a multilevel approach that allowed us to retain important information which would have been discarded using traditional meta-analytic procedures. However, in an effort to provide a clearer view on the feasibility of using Facebook data to predict personality, we limited the scope of this paper to the analysis of studies mining Facebook data using predictive modeling techniques, discarding strictly correlational studies. Results showed that on average, the accuracy of prediction of individual Big Five personality scores based on predictive models is moderate (r = .34), and most of the variability existing among included effect-sizes is linked to

study-level differences (89.14%), while only a relatively small proportion is related to within-study differences among effect-sizes (10.78%). Among the traits, extraversion is associated with the highest prediction accuracy (r = .39), while agreeableness shows the lowest prediction accuracy (r = .28). However, when comparing effects across personality traits, pairwise contrasts were generally non-significant (with the exception of the contrasts comparing the agreeableness trait with extraversion, and openness), indicating a general overlap in prediction accuracy among traits. This indicates that the performance of predictive models tends to be quite stable across personality traits, while most of the differences in predictive power can be traced back to differences among individual studies, possibly due to methodological differences in the specific analytical approach used to mine collected data, as well as the amount and type of data collected. The specific data source used in the different studies does not seem to have a significant impact on accuracy of predictions: on the contrary, even among studies using the same data source (e.g. MyPersonality data), there remains a significant amount of variability in prediction performance, which is possibly related to methodological differences across studies. Accordingly, moderator analyses revealed that existing differences among studies in the use of multiple types of digital footprints (as opposed to a single type), and use of demographic information among the predictor set, significantly contribute in explaining differences in the accuracy of personality predictions. Concerning demographics, findings confirm the importance of demographic information, including age and gender, as factors in explaining individual differences in Big Five traits (e.g. Lehmann, Denissen, Allemand, & Penke, 2013; Soto, John, Gosling, & Potter, 2011).

Further, the investigation of publication bias revealed a theoretically interpretable effect showing a negative link between the standard error of estimates (i.e. the inverse of precision), and overall accuracy of personality prediction. Because precision error is directly related to study sample size (Kirkwood & Sterne, 2010), this result highlights the importance of recruiting large samples of users for the purpose of improving accuracy of prediction (Kosinski, Wang, Lakkaraju, & Leskovec, 2016). It is important to note that, looking at the funnel plot of studies effect-sizes plotted against their standard error, it is apparent that this effect is most prominent in studies using small- to moderately-sized samples (n < 1000). In turn, among studies performed on larger samples (n ≥ 1000), there remains a relevant heterogeneity in effect-sizes between studies, possibly related to methodological differences between them.

Overall, this study has demonstrated that Big Five personality variables can be inferred with moderate accuracy using currently available social media data. Because the overall meta-analytic effect size presented here is moderate, it appears that the analysis of digital footprints still falls short in predicting such characteristics with accuracy allowing for assessment at the individual level. For example, for each trait, the average correlation between predicted and self-report personality scores is much lower than the correlation one would expect between consecutive self-report personality assessments of the same individual (i.e. test-retest reliability, see Kosinski et al., 2013). Similarly, the strength of the correlation between predicted and self-report personality scores is far below that expected for personality instruments that are intended to assess the same latent construct (e.g. convergent validity, r ≈ .75 for short Big Five assessments, Pervin & John, 1999). However, it is reasonable to expect that prediction accuracy might become more precise in the future, as larger datasets become available, and new types of data are collected and mined for prediction purposes (e.g. features extracted from visual data or location data). Overall, the existing findings seem to indicate that demographic and behavioral variables may be more easily predicted than unobservable – and hence latent – personality traits (Kosinski, Stillwell, & Graepel, 2013). Still, personality remains an important topic to study, because it is associated with important life variables such as longevity (via health behaviors), (Bogg & Roberts, 2004; Bogg & Roberts, 2012; Jackson, Connolly, Garrison, Leveille, Connolly, 2015), job performance (Barrick & Mount, 1991), or vulnerability to psychiatric disorders such as depression (Lahey, 2009). Further, personality has been linked to variables such as burnout (Alarcon, Eschleman, Bowling, 2009), and personality information could be of use in adjusting work processes based on individual characteristics, such as taking into account if a person is easily stressed. Therefore, predicting who might be vulnerable to stress might be particularly useful to target work place interventions aimed at restructuring the digital work flow (e.g., such as introducing limits to e-mail checking, Kushlev & Dunn, 2015). The study of Facebook posts appears also to be suitable method for providing an initial screening of individuals for depression (Eichstaedt et

al. 2017), thereby helping to potentially reduce individual suffering by enabling the provision of pre-emptive support. Further, the *digital phenotyping* scene aims not only to predict psychological traits and states from the study of human-machine interaction, but ultimately also the neurobiology underlying these traits/states (Montag et al., 2017; Sariyska, Rathner, Baumeister, Montag, 2018).

However, given the feasibility of using Facebook data to infer individual characteristics unobtrusively, there is an emerging need for a more careful consideration of ethical challenges, and related sociopolitical consequences, of the use of extracted data (Montag, Sindermann & Baumeister, 2020). As highlighted by Matz and colleagues (2017), psychological targeting procedures leveraging predictive models might be used to target and manipulate the behavior of large groups of people, without the individuals being aware of it (see also problems around the filter-bubble: Sindermann et al., 2020). Predicted traits could be used to make financial or job-related decisions without users knowing it, or without explicitly stating to users that their characteristics have been determined through their social media usage patterns (Kern et al., 2019). Indeed, Facebook data could be used for purposes that go beyond what users intended when they consented to the collection of their digital footprints, revealing information that they may wish to keep private (Wang & Kosinski, 2018). As recently noted in a *Nature* editorial (2018, March 27) concerning the Cambridge Analytica scandal, the simple availability of social media data is not a sufficient reason to conduct research bound to have putative negative consequences for individual or a group of users. For a practitioner's view on ethics in digital phenotyping, see the work by Dagum & Montag (2019).

4.1 Limitations and Future Directions

The findings of the present study should be understood in light of a number of limitations. First, existing differences in data extraction and analytical procedures across the studies as a source of variability in effect-size of personality prediction were not investigated. Second, the impact of cultural differences on the accuracy of personality predictions was not examined, as most of the included studies focused on samples of English-speaking users, and only a small number involved samples derived from samples of non-English speakers. Hence, there remains a need for more culturally diverse samples in order to determine the cultural invariance of emerging findings. An additional limitation relates to the decision to include only studies assessing prediction accuracy using Pearson's correlation, and excluding those reporting only MAE and RMSE statistics, which may have introduced bias in the selection of the studies for inclusion in the meta-analytical computations. This decision was related to the potential incomparability across studies of the metric of both MAE and RMSE statistics, which in turn is dependent on the metric of the specific questionnaire used in the study to assess personality (e.g. the number of items, and the procedure used to generate scores).

60

Because model-based predictions aim to provide an assessment of personality, it is important to establish their convergent validity with self-report scores. However, MAE and RMSE statistics (as opposed to correlation) do not provide information about the strength of the linear relationship between observed and predicted scores, which in turn represents an important factor in determining the convergent validity between self-report personality scores and model-based predictions. For this reason, we decided to focus on correlation as the effect-size for the meta-analysis. As noted above, in doing this, some studies were excluded from the analysis. Although the number of excluded studies was limited, the results should be understood in light of this potential bias. A final limitation concerns the examination of the use of features extracted from pictures and videos for personality prediction. Sharing of visual content has increased dramatically over the last few years, and highly visual social media platforms such as Instagram and Snapchat are now outgrowing Facebook in popularity especially among younger people (Marengo, Longobardi, Fabris, & Settanni, 2018, Marengo, Sindermann, Elhai & Montag, in press). Because only a minority of selected studies included in the meta-analysis also used picture-information as a predictor, we could not fully investigate the impact of the inclusion of features derived from visual data in influencing accuracy of personality predictions. Given the increasing importance of this data source, future studies should consider taking such information into account when detecting personality differences.

References

Al Marouf, A., Hasan, M.K., Mahmud, H. (2019). Identifying Neuroticism from User Generated Content of Social Media based on Psycholinguistic Cues. In: 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). Cox'sBazar: IEEE, 1–5. Available at https://ieeexplore.ieee.org/abstract/document/8679505

Alarcon, G., Eschleman, K.J., Bowling, N.A. (2009). Relationships between personality variables and burnout: A meta-analysis. Work & stress, 23(3), 244–263.

Assink, M. & Wibbelink, C.J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. The Quantitative Methods for Psychology, 12(3), 154–174.

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.

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012, June). Personality and patterns of Facebook usage. In Proceedings of the 4th annual ACM Web Science Conference (pp. 24–32).

Baik, J., Lee, K., Lee, S., Kim, Y., & Choi, J. (2016). Predicting personality traits related to consumer behavior using SNS analysis. New Review of Hypermedia and Multimedia, 22(3), 189–206. https://doi.org/10.1080/13614568.2016.1152313

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. Personnel psychology, 44(1), 1–26.

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.

Bogg, T. & Roberts, B.W. (2012). The case for conscientiousness: Evidence and implications for a personality trait marker of health and longevity. Annals of Behavioral Medicine, 45(3), 278–288.

Celli, F., Bruni, E., & Lepri, B. (2014, November). Automatic personality and interaction style recognition from facebook profile pictures. In Proceedings of the 22nd ACM International Conference on Multimedia (pp. 1101–1104).

Cambridge Analytica controversy must spur researchers to update data ethics. (2018, March 27). Nature, 555, 559–560. doi: 10.1038/d41586-018-03856-4

Cheung, M.W. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. Psychological Methods, 19(2), 229.

Cutler, A., & Kulis, B. (2018, September). Inferring human traits from Facebook statuses. In International Conference on Social Informatics (pp. 167–195). Springer, Cham.

Dagum, P., & Montag, C. (2019). Ethical Considerations of Digital Phenotyping from the Perspective of a Healthcare Practitioner. In Digital Phenotyping and Mobile Sensing (pp. 13–28). Springer, Cham.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. Bmj, 315(7109), 629–634.

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

Farnadi, G., Sitaraman, G., Sushmita, S., Celli, F., Kosinski, M., Stillwell, D., Davalos, S., Moens, M.-F., & De Cock, M. (2016). Computational personality recognition in social media. User Modeling and User-Adapted Interaction, 26(2), 109–142.

Farnadi, G., Tang, J., De Cock, M., & Moens, M. F. (2018, February). User profiling through deep multimodal fusion. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (pp. 171–179).

Facebook for developers (2019) Graph API. https://developers.facebook.com/docs/graph-api.

Golbeck, J. (2016). Predicting personality from social media text. AIS Transactions on Replication Research, 2(2), 1–10.

Golbeck, J., Robles, C., & Turner, K. (2011). Predicting personality with social media. In CHI'11 Extended Abstracts on Human Factors in Computing Systems (pp. 253–262).

Gosling, S.D., Augustine, A.A., Vazire, S., Holtzman, N., Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information. Cyberpsychology, Behavior, and Social Networking, 14(9), 483–488.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.

Hinds, J., & Joinson, A. (2019). Human and computer personality prediction from digital footprints. Current Directions in Psychological Science, 28(2), 204–211.

Ihsan, Z., & Furnham, A. (2018). The new technologies in personality assessment: A review. Consulting Psychology Journal: Practice and Research, 70(2), 147–166.

Jackson, J.J., Connolly, J.J., Garrison, S.M., Leveille, M.M., Connolly, S.L. (2015). Your friends know how long you will live: A 75-year study of peer-rated personality traits. Psychological Science, 26(3), 335–340.

Kern, M.L., Eichstaedt, J.C., Schwartz, H.A., Dziurzynski, L., Ungar, L.H., Stillwell, D.J., Kosinski, M., Ramones, S.M., Seligman, M.E. (2014). The online social self: An open vocabulary approach to personality. Assessment, 21(2), 158–169.

Kern, M. L., McCarthy, P. X., Chakrabarty, D., & Rizoiu, M. A. (2019). Social media-predicted personality traits and values can help match people to their ideal jobs. Proceedings of the National Academy of Sciences, 116(52), 26459–26464.

Kirkwood, B. R., & Sterne, J. A. (2010). Essential medical statistics. John Wiley & Sons.

Kleanthous, S., Herodotou, C., Samaras, G., & Germanakos, P. (2016, July). Detecting personality traces in users' social activity. In International Conference on Social Computing and Social Media (pp. 287–297). Springer, Cham.

Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. Machine Learning, 95(3), 357–380.

Kosinski, M., Matz, S.C., Gosling, S.D., Popov, V., Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. American Psychologist, 70(6), 543.

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, 110(15), 5802–5805.

Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes. Psychological Methods, 21(4), 493–506.

Kushlev, K. & Dunn, E.W. (2015). Checking email less frequently reduces stress. Computers in Human Behavior, 43, 220–228.

Lahey, B.B. (2009). Public health significance of neuroticism. American Psychologist, 64(4), 256.

Laleh, A., & Shahram, R. (2017, December). Analyzing Facebook activities for personality recognition. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 960–964). IEEE.

Lehmann, R., Denissen, J. J., Allemand, M., & Penke, L. (2013). Age and gender differences in motivational manifestations of the Big Five from age 16 to 60. Developmental Psychology, 49(2), 365–383.

Marengo, D., Azucar, D., Giannotta, F., Basile, V., & Settanni, M. (2019). Exploring the association between problem drinking and language use on Facebook in young adults. Heliyon, 5(10), e02523.

Marengo, D., Azucar, D., Longobardi, C., & Settanni, M. (2020). Mining Facebook data for Quality of Life assessment. Behaviour & Information Technology, 1–11.

Marengo, D., Longobardi, C., Fabris, M. A., & Settanni, M. (2018). Highly-visual social media and internalizing symptoms in adolescence: The mediating role of body image concerns. Computers in Human Behavior, 82, 63–69.

Marengo, D., Sindermann, C., Elhai, J. D. & Montag, C. (in press). One Social Media Company to Rule Them All: Associations between Use of Facebook-Owned Social Media Platforms, Sociodemographic Characteristics, and the Big Five of Personality. Frontiers in Psychology.

Marengo, D., & Settanni, M. (2019). Mining Facebook Data for Personality Prediction: An Overview. In Digital Phenotyping and Mobile Sensing (pp. 109–124). Springer, Cham.

Markovikj, D., Gievska, S., Kosinski, M., & Stillwell, D. J. (2013, June). Mining facebook data for predictive personality modeling. In Seventh International AAAI Conference on Weblogs and Social Media.

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, 114(48), 12714–12719.

McCrae, R.R. & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. Journal of Personality and Social Psychology, 52(1), 81–90.

McCrae, R.R. & John, O.P. (1992). An introduction to the five-factor model and its applications. Journal of Personality, 60(2), 175–215.

Montag, C., Duke, É., Markowetz, A. (2016). Toward Psychoinformatics: Computer science meets psychology. Computational and Mathematical Methods in Medicine, 2016. doi: 10.1155/2016/2983685.

Montag, C., Markowetz, A., Blaszkiewicz, K., Andone, I., Lachmann, B., Sariyska, R., Trendafilov, B., Eibes, M., Kolb, J., Reuter, M. (2017). Facebook usage on smartphones and gray matter volume of the nucleus accumbens. Behavioural Brain Research, 329, 221–228.

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

Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., & Rentfrow, J. (2018). Musical preferences predict personality: evidence from active listening and Facebook likes. Psychological Science, 29(7), 1145–1158.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., & Seligman, M. E. P. (2015). Automatic personality assessment through social media language. Journal of Personality and Social Psychology, 108(6), 934–952.

Perrin, A. & Anderson, M. (2019, April 10). Share of US adults using social media, including Facebook, is mostly unchanged since 2018. Pew Resarch Center. Retrieved April 18, 2019, form https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/

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

Quercia, D., Lambiotte, R., Stillwell, D., Kosinski, M., & Crowcroft, J. (2012, February). The personality of popular facebook users. In Proceedings of the ACM 2012 conference on computer supported cooperative work (pp. 955–964).

Rabe, L. (2019, July 25). Anzahl der monatlich aktiven Facebook Nutzer nach Regionen weltweit vom 1. Quartal 2013 bis zum 2. Quartal 2019 (in Millionen). Statista. Retrieved from https://de.statista.com/statistik/daten/studie/885734/umfrage/anzahl-der-monatlich-aktiven-nutzer-von-facebook-nach-regionen/

Sariyska, R., Rathner, E., Baumeister, H., Montag, C. (2018). Feasibility of linking molecular genetic markers to real-world social network size tracked on smartphones. Frontiers in neuroscience, 12.

Settanni, M. & Marengo, D. (2015). Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. Frontiers in Psychology, 6, 1045.

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.

Sindermann, C., Elhai, J. D., Moshagen, M., & Montag, C. (2020). Age, gender, personality, ideological attitudes and individual differences in a person's news spectrum: how many and who might be prone to "filter bubbles" and "echo chambers" online?. Heliyon, 6(1), e03214.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. Journal of Personality and Social Psychology, 100(2), 348.

Tadesse, M.M., Lin, H., Xu, B., Yang, L. (2018). Personality predictions based on user behavior on the Facebook social media platform. IEEE Access, 6, 61959–61969.

Tandera, T., Suhartono, D., Wongso, R., Prasetio, Y.L. (2017). Personality prediction system from Facebook users. Procedia Computer Science, 116, 604–611.

Thilakaratne, M., Weerasinghe, R., & Perera, S. (2016, October). Knowledge-driven approach to predict personality traits by leveraging social media data. In 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI) (pp. 288–295). IEEE.

Torfason, R., Agustsson, E., Rothe, R., & Timofte, R. (2016). From Face Images and Attributes to Attributes. In S.-H. Lai, V. Lepetit, K. Nishino, & Y. Sato (A c. Di), Computer Vision – ACCV 2016 – 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20–24, 2016, Revised Selected Papers, Part III (Vol. 10113, pagg. 313–329). Springer.

Van den Noortgate, W., López-López, J.A., Marín-Martínez, F., Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. Behavior Research Methods, 47(4), 1274–1294.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of statistical software, 36(3), 1–48.

Wald, R., Khoshgoftaar, T., & Sumner, C. (2012, August). Machine prediction of personality from Facebook profiles. In 2012 IEEE 13th International Conference on Information Reuse & Integration (IRI) (pp. 109–115). IEEE.

Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of Personality and Social Psychology, 114(2), 246–257.

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

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

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. Proceedings of the National Academy of Sciences, 112(4), 1036–1040.

Yulianto, M., Girsang, A.S., Rumagit, R.Y. (2018). Business intelligence for social media interaction in the travel industry in Indonesia. Journal of Intelligence Studies in Business, 8(2).

Zhang, L., Zhao, L., Zhang, X., Kong, W., Sheng, Z., & Lu, C. T. (2018, December). Situation-Based Interpretable Learning for Personality Prediction in Social Media. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 1554–1562). IEEE.

Potential conflicts of interests

The authors report no conflicts of interest with this paper.

Nevertheless, for reasons of transparency, Dr. Montag mentions that he has received (to Ulm University and earlier University of Bonn) grants from agencies such as the German Research Foundation (DFG). Dr. Montag 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 the gaming or social media industry. Dr. Montag 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 functions as independent scientist on the scientific advisory board of the Nymphenburg group. This activity is financially compensated.

Author contributions

DM and CM designed the present study. DM analyzed the data and wrote the method/result sections. CM drafted the introduction and discussion sections, which were later edited and revised by DM. Both authors worked over the manuscript and critically revised it.

Funding

*Corresponding author

Davide Marengo, PhD. Department of Psychology, University of Turin, Via Verdi 10, 10124, Turin, Italy.
Email: davide.marengo@unito.it
Telephone: +39 011 6702793

Supplementary Material

eTable 1: Data Description

Variable	Explanation
study	study identifier
correlation	effect-size data
id	effect-size identifier
dataset	data source identifier
samplevar	sampling variance estimate
stander	standard error of the correlation
n_sample	sample size for correlation
multiple	Study used multiple types of digital footprints to perform prediction (1) vs. a single type of digital footprints (0)
validation	Study used a cross-validation method (holdout or k-fold) (1) vs. no cross-validation (0)
demos	Study used demographic data to perform prediction (1) vs. no use of demographic data (0)
stats	Study used activity statistics to perform prediction (1) vs. no use of activity statistics (0)
language	Study used language features to perform prediction (1) vs. no use of language features (0)
likes	Study used Facebook Likes to perform prediction (1) vs. no use of Facebook Likes (0)
trait	Personality trait on which prediction was performed: 1 = Agreeableness; 2 = Conscientiousness; 3 = Extraversion; 4 = Neuroticism; 5 = Openness