Research Article

Development and Evaluation of Psychometric Properties of the Online Profile-Based Personality Questionnaire

Iman Mohammadi[1], Hadi Farhadi[*2] [*], Gholamreza Manshaei[3]

Abstract
There has been a significant growth in people's use of social media in recent years. It has been determined that the digital footprints are a reflection of the personality. It is possible to use this data to predict characteristics and use it in various fields, such as job Assessment Centers. The purpose of this research was to develop and examine the psychometric properties of the online profile-based personality questionnaire (OPBPQ). This study was conducted based on a quantitative positivist paradigm, utilizing a cross-sectional design. A total of 300 participants from different organizations completed the self-reported questionnaires. The data analysis procedure included item screening, confirmatory factor analysis (CFA), test for common method bias (CMB), determination of convergent validity, and invariance analysis (configural, metric, and scalar). The results showed that the OPBPQ can appropriately predict personality features and is a reliable and valid instrument. The results showed that the personality prediction questionnaire based on an online profile can appropriately predict personality features and is a reliable and valid instrument. The questionnaire can predict personality features using the data that people share on social media. Also, it can assist psychologists in using a new and reliable instrument for evaluating one's characteristics. The development of a valid and reliable questionnaire can be useful for psychotherapists, employees, and researchers for the purposes of helping clients and organizations.

Keywords: Online Profile-Based Personality Questionnaire (OPBPQ), Characteristics, Psychometric Properties

Introduction
Almost half of the people spend more than two hours of their daily time in social medial. It can be assumed that a person spends about 6 years and 8 months of their entire life in virtual spaces (Vaid & Harari, 2021). In recent years, there has been an increase in people's use of virtual space. This dramatic growth has led to a significant increase in interpersonal communication as well as the expression of personal daily events and topics (Settanni et al., 2018). Also, Instagram has a special place with more than 800 million active users who share a lot of data daily (Ferwerda & Tkalcic, 2018b). Since it has been determined that the social media and the profiles created in them are a true reflection of the personal expressions of the users, this space has been used as an ideal platform for studying human behavior (Golbeck et al., 2011). Among other reasons, people use online platforms to communicate with friends and family members. Users of virtual space usually

[1] . Ph.D. Student of Psychology, Research Center for Behavioral and Psychological Science, Department of Psychology, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.
Email: imanmoha.71@gmail.com
[2] Associate Professor, Research Center for Behavioral and Psychological Science, Department of Psychology, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.
* Corresponding author:
Email: h.farhadi@khuisf.ac.ir
[3] . Associate Professor, Department of Psychology, Faculty of Psychology and Educational Sciences, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.
Email: smanshaee@khuisf.ac.ir

follow their friends and colleagues to share their posts, comments, and opinions with them. Using social media as an external source of information to predict personality has many advantages, including the fact that people share a lot of information on social media (Golbeck et al., 2011). Facebook, Twitter, and Instagram contain a lot of information that can be used to evaluate personality traits. This information can be easily extracted (Golbeck et al., 2011). Having information about people's personalities can help us provide appropriate reactions when needed. Also, awareness of personality traits provides us with reliable information about people's capabilities, strengths, and weaknesses (Souri et al., 2018).

Personality refers to a set of behaviors and emotions that includes unique characteristics caused by the environment and heredity varying from person to person based on different cognitive and emotional conditions. Personality traits are formed based on activities, cognitions, emotions, and external behaviors. Therefore, any activity of a person is caused by the interaction of nature and the environment. Recently, people have used virtual space to share their emotions, feelings, thoughts, and activities. This information can be classified (Valanarasu, 2021). Most of the research so far has analyzed the information available on social media with the help of machine learning to predict personality.

The use of machine learning (artificial intelligence to study statistical algorithms) has its difficulties and is time-consuming and expensive. It also requires a large sample and volume of input data. On the other hand, the studies carried out with the help of *Machine learning* have provided researchers with information about existing connections and correlations (Singh and Singh, 2023). The researcher of this study came to the conclusion that with the help of the findings in previous research, an instrument can be designed that can predict personality traits by evaluating the information available in social media and people's digital footprints. It is assumed that personality characteristics can be predicted from the information that users share on social media (Valanarasu, 2021). With the correct and scientific application of this information, we have developed a tool to reduce the existing limitations and provided a useful instrument for predicting personality traits for researchers, psychologists, and members of the Occupational Assessment Center group.

In a study, Shipankar et al. (2022) examined personality prediction capability by the use of data from social media. In their research, Twitter and Myers-Briggs personality assessment tests were used. The results revealed that a user's MBTI personality traits can be predicted from the public information they share on Twitter. The presented methodology significantly improved the accuracy of recognizing the Intuition (I), Sensing (S), Introversion (I), and Extroversion (E) personality categories, as well as slightly better accuracy for the Judging (J), Perceiving (P) of personality category. The results showed that the model proposed in this study can predict personality traits with 76% accuracy.

In a study, Valanarasu (2021) did a comparative analysis to predict personality through the digital footprints of people in virtual networks. The proposed algorithm of this study used multiple data as well as the data available in several virtual network platforms including Twitter, YouTube, and Facebook. Using multiple platforms increased measurement accuracy. The Naïve Bayes machine learning is probabilistic and it provides a better prediction rate than all other machine learning algorithms. They can predict the personality of any job applicant with higher accuracy by using the proposed algorithm. The proposed algorithm can be incorporated with traditional questionnaire-based personality predictors, which can provide high accuracy and better prediction. The results showed that this method can predict the Big Five characteristics better than other learning machine tools. Also, the results clearly showed that the use of multiple platforms significantly improves the accuracy of personality prediction through virtual networks.

Ock and An (2021) conducted a study called, "The Machine Learning-based Approach to Personality Assessment and Its Application in Personnel Selection". This research has conducted a summary review of studies that have measured the validity of machine learning in personality assessment. These researchers concluded that the studies of tools related to machine learning should be continuously validated. Besides, one should not simply rely on the results of current studies and new solutions should be developed for the basic understanding of personality based on the data available in social networks.

In a study, Kern et al. (2019) predicted personality traits and values based on social networks and their ability to choose an ideal job. Jobs have their special profiles, which can be well predicted through linguistic data shared in social networks. To automatically determine each user's personality digital fingerprint, they used the IBM Watson Personality Insights system. In this study, using a relatively large sample, users' personalities and profiles of different jobs were drawn automatically. Similar jobs were grouped in separate groups. Using a large dataset, information unobtrusively available online (i.e., Twitter language), and a combination of Big 5 traits and 5 basic values, our study suggests that personality digital fingerprints relate to distinctive occupations. Results demonstrate the potential to create an atlas of career aptitude, based on non-cognitive personality traits and values. We anticipate that this could have significant applications in career guidance for new graduates, disengaged employees, career changers, and the unemployed. The results showed that the data available in social networks can be used to choose the most suitable job for each person.

In a study, Subramani (2018) predicted and analyzed personal characteristics using Instagram. Images, content of posts, and main performance indicators for each post have been used. In this study, the appearance of each image was examined through the analysis of filters and colors. The linguistic data of words were evaluated and finally, the amount of activity and social communication was used to measure personality traits. Research suggests that all 3 Instagram features help to predict the personality traits with considerable accuracy level. Importantly, image features provide outstanding results when we deal with the mentioned classification results. Conscientious persons share a bunch of posts that have plenty of likes & comments from their circle. Moreover, they have a large number of followers and followings. Also, these people view and follow the latest trends such as fashion and style. Furthermore, they pretend that they are an active user on Instagram. People with agreeableness traits don't share a lot of posts and stories. Also, they don't have many followers. These people often share and follow posts about photography, and electronic products, especially electronic cigarettes and mobile. People high in Extraversion like to share information about arts and music. Also, they report that they are a regular user of social media. They don't receive a lot of interaction from their social circle. From the correlation, we can understand that neurotic people are also not active users of social media. These people don't receive likes and comments from other users. They follow and post the latest fashion styles and photos. They are also the regular user of Instagram. People with openness to a new experience personality trait, follow and post more information about the news, research, cities, festivals, and game content on their social media. We found a significant negative correlation in the "Number of the post" which means that people with an openness to new experiences trait, don't share many posts on Instagram.

Azucar et al. (2018) in a meta-analytical study reviewed the articles related to the prediction of 5 main personality factors by means of digital footprints in social media. The results can be used for various purposes, including the development of online services to improve user experiences, improve guidance systems, and possible screenings for the public health of the community. In this study, a series of meta-analyses were conducted to determine the predictive power of digital footprints in social networks. The results of meta-analyses have shown that this correlation is higher than normal. The results showed that the accuracy of prediction based on five main personality factors is suitable and this prediction is improved by adding demographic characteristics and using several forms of digital footprints.

Ferwerda and Tkalcic (2018a) studied personality prediction through Instagram images. In their research, they studied two different features that could be extracted from images. The first one is visual characteristics (color, capacity, and saturation) whereas the second is content characteristics (content of images). These researchers developed an online survey to collect data and asked the participants to fill in the personality questionnaire. On the other hand, by providing researchers with the possibility of accessing the Instagram profile, the random forest classifier shows slight improvements on half of the personality traits: openness to experience, agreeableness, and neuroticism. In their research, 54962 images from 193 Instagram users have been collected. The results showed that visual and content characteristics can be used to predict personality well. In this work, they showed that the visual features as well as the content features consist of information for personality prediction that attain similar results. The results showed that visual and content characteristics can be used to predict personality well. In this work, they showed that the visual features as

well as the content features consist of information for personality prediction that attain similar results. Neuroticism personality trait has a negative correlation with sharing photos with cyan color. It means that people with this trait also tend to upload photos with less cyan color. Also, we found a positive correlation between Saturation Mean, Median, and Standard deviation which means photos have vivid and bleak colors. Furthermore, we noticed a positive correlation between the Value median and Standard deviation; however, the Value minimum has a negative correlation. It indicates that participants are not uploading a photo with low brightness. In other words, the Brightness level must be intermediate. They found a significant positive correlation with Saturation standard deviation which means people with agreeable traits tend to have both vivid and bleak colors in their photos. Also, these personalities tend to upload a photo that has a higher number of people. Extraversion personality has a positive correlation with orange color and a negative correlation with Cyan. This indicates that people with this type of personality, share photos that have a huge orange portion and fewer cyan tones.

In this study, we aimed to identify personality dimensions based on an online profile (Instagram), develop a personality questionnaire based on that, check its validity and reliability, and check the capability of this questionnaire in Assessment Centers.

Methodology

This sequential exploratory mixed-method study was conducted in a qualitative and a quantitative phase which is designed used when a researcher does not know the most important concepts for the study and when there are no appropriate tools for the measurement of an intended concept. In the qualitative phase, the concept of personality prediction by digital footprints and its dimensions have been explored based on the literature review. Then, the primary items were developed. After that, the psychometric properties of the questionnaire were assessed in the quantitative phase (Nikpour et al., 2018). We report how we determined our sample size, data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, measures in the study, and analyses including tested models. If we use inferential tests, we report exact p values, effect sizes, and 95% confidence or credible intervals. The data is not publicly available to ensure the anonymity and privacy of the participants. However, the data is available upon request from the corresponding author.

The qualitative phase
Data collection
For data collection, the deductive method was used. The study in the literature review was performed to clarify the concept and dimensions of personality traits based on online profiles. The researcher used Google Scholar and reliable search engines to search the literature related to this topic.

Development of the online profile-based personality prediction questionnaire
An online profile-based personality prediction questionnaire is designed to predict the five main personality factors. While recognizing the importance of predicting personality features in various situations, the current literature lacks a precise instrument that can use the data mined from social networks to predict characteristics.

The process involved in scale development includes the following:
- clearly defining the construct under investigation
- specifying the dimensionality and purpose of the construct
- item generation, informed either deductively or inductively such as the focus groups or interviews
- determining the most suitable response format to employ
- inspecting the validity and reliability of items (Mabitsela et al., 2024)

The source questionnaire was developed in Persian. So, we applied a reliable method to translate it into English. Documentation of the Translation Method includes several steps:

- Develop the questionnaire in the source language
- Translation: Documenting the number of independent translations to indicate the range of options going into the next phase. The double (or parallel) translation is regarded as the best practice since it shows divergent interpretations and offers variants to choose from. The adaptation uses the version from another country and tailors it to the cultural and linguistic particularities of the target country.
- Review: Document reviewing procedures to give indications of objectivity, sharing of expertise, and overall quality.
- Harmonization: Document harmonization procedures to indicate consideration of comparability needs on a larger scale.
- Pretesting: Document pretesting activities to indicate validation procedures among the target population.
- Quality control: Document further procedures to cover all steps implemented in a translation endeavor, such as back translation(s), back translation comparison, external quality control, and proof-reading (Behr & Zabal, 2020)

All the processes above have been held precisely and the questionnaire has been formulated, revised, and Finalized.

Data analysis

After completing the research and extracting related items, the researcher began to analyze data using conventional content analysis using the Grameim and Lundman model. The units have been interpreted and coded. Then, the codes were compared and categorized into subcategories according to their conceptual similarities and differences (Graneheim and Lundman, 2004). Subcategories are also grouped based on their interrelationships to form comprehensive, mutually exclusive categories. Finally, the main categories and subcategories of personality traits were identified. These main categories and subcategories have been used to generate the primary items for the Online Profile-Based Personality Questionnaire (OPBPQ). The existing literature and instruments will also be used for item generation

Population, sampling, and the sample

The population of this study is the Isfahan workforce employed in organizations across the city without any focus on a specific province or industry. Respondents of this study were conveniently selected. Two organizations were selected through the convenience sampling method from Isfahan. The participants were selected randomly regarding several inclusion and exclusion criteria. Inclusion criteria were consent for participation, continuous activity on Instagram for at least 2 years, having at least 10 posts on the Instagram page, having an active personal social media page, and an age of 20–50. Exclusion criteria were lack of full cooperation in the implementation of questionnaires, complete commitment to the accurate implementation of questionnaires, and having a working and advertising page. A number of 1746 persons have been specified and 300 participants were selected randomly based on this method. The results of demographic findings showed that the average age of the sample is 33.42 and its standard deviation is 4.93 years. A total of 159 persons (53 percent) were male and 141 persons (47 percent) were female; 165 persons (55 percent) of the sample were married and 135 persons (45 percent) were unmarried. Education in 9 persons (3 percent) was less than a diploma, 18 persons (6 percent) had a diploma, 15 persons (5 percent) had a post-graduate diploma, 162 persons (54 percent) had a bachelor's degree, 57 (19 percent) had a master's degree, and 39 persons (13 percent) had a doctorate degree.

The quantitative phase

The focus of this phase is on assessing the psychometric properties of OPBPQ. The face, content, construct, divergent, and convergent validity as well as reliability have been examined.

Face and content validity assessment

After designing the questionnaire, during qualitative content validity assessment, the tool was provided to 10 experts, and each of them rated the necessity of each item in the questionnaire in a 3-degree spectrum ("the item is necessary", "the item is useful but not necessary" and "the item is not necessary") and the degree of relevance of each item to the whole questionnaire was specified in a spectrum of 4 degrees ("not relevant", "relatively relevant", "relevant", "completely relevant"). Then the Content Validity Index (CVI) and Content Validity Ratio (CVR) coefficients were calculated. The calculation results showed that the CVR and CVI value of 5 items was less than 0.62 and therefore those questions were removed from the questionnaire. The final questionnaire concluded with 50 questions. Questions 1 to 10 evaluate neuroticism, questions 11 to 20 evaluate Extraversion, questions 21 to 30 evaluate openness to new experience, questions 31 to 40 evaluate agreeableness, and questions 41 to 50 evaluate conscientiousness.

Construct validity assessment

To check the validity of the questionnaire, construct validity was assessed by confirmatory factor analysis.

Sample size and sampling strategy

One method for calculating sample size in confirmatory factor analysis is the rule of thumb which states that samples greater than 300 are appropriate for most confirmatory factor analyses (Harrington, 2008). Therefore, in this study, the sample size is considered to be 300. To perform confirmatory factor analysis, first, the quality of the correlation matrix of the scale propositions and the sampling capability were examined. The results showed that the value of Bartlett's sphericity test was equal to 8740.99, which was significant at the 0.001 level. The KMO coefficient for this analysis was equal to 0.875. In this way, the information in the data matrix is meaningful and the sample size is favorable for factor analysis. Also, the results of confirmatory factor analysis including the factor loadings of the eigenvalue questionnaire, the percentage of variance, and the percentage of cumulative variance showed that the factor loadings of the questions in all five factors are higher than 0.4 and, therefore, favorable.

Convergent and divergent validity assessments

Convergent validity was calculated using the average variance criterion and divergent validity was calculated using the maximum common square and mean square common variance indices. Convergent validity was calculated using the average variance criterion and divergent validity was calculated using the maximum common square and mean square common variance indices. The results showed that the final validity of the personality questionnaire based on the online profile of the AVE index or the average variance extracted in all dimensions of the questionnaire was higher than 0.5. Also, the MSV and ASV indices are lower than the AVE index. Based on this, it can be said that the personality questionnaire based on an online profile has good reliability. Convergent and divergent validity of the questionnaire is also acceptable. Then, the correlation of OPBPQ with the NEO personality dimensions was evaluated to achieve concurrent convergent validity.

Reliability assessment

OPBPQ reliability is evaluated via the internal consistency, stability, and construct reliability assessment methods. During the internal consistency assessment, Cronbach's alpha values were calculated for OPBPQ and its subscales. Values greater than 0.7 are considered acceptable (Mayers, 2013). In the final application of the questionnaire, Cronbach's alpha coefficients and halving coefficients in all dimensions and components were obtained above 0.7, which is acceptable. Also, the combined reliability is higher than 0.7 and is acceptable.

Developing OPBPQ scoring system

Items have been scored on a five-point Likert-type scale from five to one. There is no negatively worded item to score reversely.

Statistical data analysis

The SPSS26 and the AMOS24 were used for data analysis. Data presentation has been done via the measures of descriptive statistics including mean, standard deviation, and frequencies. Moreover, statistical analyses will be performed through running confirmatory factor analysis, Pearson correlation analysis, paired- and independent sample T-tests, Friedman test, Cronbach's alpha model, intra-class correlation coefficient, confirmatory factor analysis, and standard error measurement.

Ethical considerations

Ethical clearance for the study was obtained from the Research Ethics Committee of Islamic Azad University, Isfahan (Khorasgan) Branch. All participants applied to the research willingly and were informed completely of the necessary details. All data have been coded to protect participant identity and fully compliance with the principle of confidentiality.

Results

Based on the results obtained in Table 1, the relative chi-square index is equal to 2.608, which shows that this model has an acceptable condition. The value of TLI and CFI comparative indices is higher than 0.9. The value of PCFI is also higher than 0.5 and desirable as a goodness of fit index. The value of RMSEA as the most important index of overall fit is equal to 0.07 and shows that the model has a good fit in general.

TABLE 1. Structural equation model fit indices for confirmatory factor analysis of personality questionnaire based on online profile

Index name		Limit required	Index value in the model	Situation in the proposed model
Absolute	(χ^2 or CMIN) Chi-Square	The significance level is above 0.05	3004.923 (Significance level 0.001)	optimal according to other indicators
	Degree of freedom	-	1152	-
Comparative	(TLI)	Above 0.9	0.916	Optimal
	(CFI)		0.905	
	(PCFI)	Above 0.5	0.760	Optimal
Parsimonious	(RMSEA)	Less than 0.1	0.07	Optimal
	(CMIN/DF)	Less than 5	2.608	Optimal

The results in Table 2 have shown that factor loadings in all questions in the confirmatory factor analysis are higher than 0.3 and significant ($p < 0.001$). Also, as can be seen, the t-statistic in all items of the questionnaire is higher than 1.96. Based on this, it can be said that the construct validity using confirmatory factor analysis in the online profile-based personality questionnaire is acceptable.

The results in Table 3 have shown that the correlation between the dimensions is between 0.243 and 0.633 and significant ($p < 0.001$). Furthermore, a significant correlation ($p<0.001$) between dimensions of neuroticism in online and NEO questionnaires (r = 0.814), extroversion in two questionnaires (r = 0.692), experientiality (r=0.669), agreeableness (r=0.686) and being conscientious (r = 0.663) was obtained. Based

on the above-mentioned results, it can be concluded that concurrent validity of the convergent type has been obtained in the personality questionnaire based on an acceptable online profile.

The reliability coefficients of Cronbach's alpha and combined reliability to check reliability and average variance extracted to check convergent validity are presented in Table 4. The average variance extracted (AVE) shows the degree of correlation of a structure with its indicators, which should be higher than 0.5. Moreover, the reliability coefficients of internal consistency using Cronbach's alpha coefficient and composite reliability should be higher than 0.7.

The results in Table 4 indicate that in all five dimensions of the online profile-based personality questionnaire, Cronbach's alpha coefficients and combined reliability in all dimensions are higher than 0.7, which is acceptable. Also, the AVE index or average variance extracted in all aspects of the questionnaire is higher than 0

Table 2: Results of estimations and general specifications of the model

Dimension	Question	Non-standard estimate	Standard estimate	T statistic	Meaningful	Dimension	Question	Non-standard estimate	Standard estimate	T statistic	Meaningful
Neuroticism	1	0.508	1.000	-	-	Openness	26	0.732	1.64	6.703	0.001
	2	0.418	0.583	5.71	0.001		27	0.553	1.09	6.03	0.001
	3	0.476	0.755	6.27	0.001		28	0.518	1.13	5.85	0.001
	4	0.435	0.51	5.88	0.001		29	0.546	1.12	5.98	0.001
	5	0.576	0.83	7.08	0.001		30	0.579	1.11	6.15	0.001
	6	0.722	1.38	7.98	0.001		31	0.718	1.000	-	-
	7	0.535	0.955	6.77	0.001		32	0.719	1.06	11.36	0.001
	8	0.614	1.38	7.65	0.001		33	0.540	0.853	8.601	0.001
	9	0.583	1.11	7.14	0.001		34	0.583	0.877	9.29	0.001
	10	0.516	1.302	6.61	0.001		35	0.457	0.673	7.31	0.001
	11	0.382	1.000	-	-		36	0.505	0.526	8.08	0.001
	12	0.553	1.74	5.91	0.001		37	0.597	0.781	9.52	0.001
	13	0.537	1.67	5.84	0.001		38	0.408	0.826	6.55	0.001
	14	0.650	2.19	6.26	0.001		39	0.494	0.897	7.907	0.001
Extraversion	15	0.788	2.42	6.62	0.001	Agreeableness	40	0.438	0.906	7.02	0.001
	16	0.902	3.06	6.82	0.001		41	0.771	1.000	-	-
	17	0.934	3.15	6.87	0.001		42	0.789	1.16	15.67	0.001
	18	0.739	2.23	6.51	0.001		43	0.819	1.01	16.18	0.001
	19	0.662	1.96	6.307	0.001		44	0.830	1.02	17.11	0.001
	20	0.549	1.64	5.89	0.001		45	0.902	1.002	17.47	0.001
	21	0.411	1.000	-	-		46	0.924	1.13	19.95	0.001
	22	0.537	1.26	7.46	0.001		47	0.906	1.14	20.8	0.001
Openness	23	0.463	1.26	5.52	0.001	Conscientiousness	48	0.825	1.15	20.08	0.001
	24	0.754	1.74	6.76	0.001		49	0.862	0.933	17.37	0.001
	25	0.7	1.41	6.607	0.001		50	0.820	1.025	26.33	0.001

Table 3. Descriptive findings and correlation between personality dimensions based on online profile

	Variables	M	SD	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness	Neuroticism	Extraversion	Openness	Agreeableness
Personality based on online profile	Neuroticism	21.8	5.55									
	Extraversion	28.7	6.47	-0.633**								
	Openness	27.7	4.41	-0.258**	0.243**							
	Agreeableness	32.31	4.66	-0.514**	0.547**	0.356**						
	Conscientiousness	34.56	5.55	-0.445**	0.355**	0.243**	0.462**					
Personality based on NEO	Neuroticism	23.84	4.28	0.814**	-0.486**	-0.176*	0.473**	-0.318**				
	Extraversion	32.57	4.75	-0.334**	0.692**	0.251**	0.515**	0.315**	-0.322**			
	Openness	33.34	3.76	-0.278**	0.202**	0.669**	0.369**	0.329**	-0.556**	0.285**		
	Agreeableness	35.58	3.32	-0.318**	0.396**	0.348**	0.686**	0.466**	-0.239**	0.375**	0.335**	
	Conscientiousness	36.35	4.43	-0.560**	0.404**	0.222**	0.624**	0.663**	-0.491**	0.394**	0.526**	0.478**

Note. $*p < .05$; $**p < .01$, M = Mean, SD = Standard Deviation

Table 4. Cronbach's alpha reliability coefficients, split coefficient and combined reliability and AVE coefficient to check the validity of the dimensions of the personality scale based on the online profile

Dimension	Cronbach's alpha	Combined reliability	AVE index
Neuroticism	0.795	0.824	0.513
Extraversion	0.902	0.917	0.529
Openness	0.840	0.876	0.516
Agreeableness	0.812	0.862	0.533
Conscientiousness	0.955	0.969	0.756

Discussion and Conclusion

Due to the growing use of virtual networks and the significant use of it by people, as well as the large amount of data available in the virtual space, the vacancy of a handy tool to measure personality based on the digital footprint of people has been quite evident. Many studies have investigated and predicted personality traits through people's digital footprints. but they mostly used machine learning to achieve this goal. In this mixed-method study, a questionnaire has been developed and Then, the psychometric properties of the questionnaire assessed. So far the lack of standardized questionnaires for personality prediction by the use of digital footprints, OPBPQ can be a critical step in assessing characteristics. Also other scholars can cross-culturally adapt and use OPBPQ according to their immediate contexts. The strengths of this study were its sequential exploratory mixed-method design, sampling carefully and precisely, and its relatively large sample size. OPBPQ create the possibility of collecting data about big 5 personality traits in a new and innovative way. The results of structural equation model fit of the second-order confirmatory factor analysis have shown that this model has an acceptable condition, and RMSEA and PCFI show that the model generally has a good fit. Also, the T-statistics and factor loadings in all the questions in the confirmatory factor analysis were significant. Based on this, it can be said that the construct validity using confirmatory factor analysis in the OPBPQ is acceptable. Correlations between significant dimensions have been obtained. so concurrent validity of the convergent type has been obtained in the personality questionnaire based on an acceptable online profile The reliability coefficients of Cronbach's alpha and combined reliability to check the reliability and average variance extracted to check the convergent validity were also significant. In all five dimensions of the online profile-based personality questionnaire, Cronbach's alpha coefficients and composite reliability have been obtained in all significant dimensions, which are acceptable. Also, the extracted average variance index was acceptable in all aspects of the questionnaire. The results showed that OPBPQ is valid and reliable and it is practical to predict the personality traits The results obtained in this literature were in line with the researches results of Valanarasu (2021), Ock & An (2021) and Azucar et al. (2018) that digital footprints can predict personality traits. Therefor the OPBPQ can predict personality traits with a simple, easy and accessible tool instead of machine learnings which are hard to access and even hard to apply. Also, Kern et al. (2019) predicted personality traits and values based on social networks and their ability to choose an ideal job. Jobs have their own special profiles, which can be well predicted through linguistic data shared in social networks. They used machine learning, a new way but hard to use in clinical and psychological situations. The similar results have been obtaining in this literature, so that we could predict personality traits and apply it in occupational psychology. Ferwerda & Tkalcic (2018a) and Subramani (2018) predicted and analyzed characteristics using Instagram. Almost all of items such as colors, brightness, user activity and contents were found relevant in previous literatures and current.

Limitation

The inclusion of a cross-sectional design could be viewed as a limitation in the study because it could lead to meager information concerning the process over time or causality. Nevertheless, this constraint can be addressed using several methods and research designs like qualitative approaches. The difficulties with sampling and lack of cooperation were the study limitations too.

Implications & Originality

This study's contribution to science, practice, and the community is based on the importance of personality prediction ability. This study further addresses the need for a reliable and valid questionnaire that is developed for different uses, thus being context-specific and relevant. The researchers propose that the survey be conducted in various settings such as psychotherapy rooms, organizations, and industries; in Iran and globally. Likewise, the researchers also recommend that the OPBPQ can predict other variants such as job performance, mental well-being, and other psychological variants. Therefore, this study could open a new era to explore and examine the potential of OPBPQ and social media data to be used in psychological matters.

References

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

Behr, D., & Zabal, A. (2020). Documenting Survey Translation. Mannheim, GESIS - Leibniz Institute for the Social Sciences (GESIS- Survey Guidelines).

Ferwerda, B., & Tkalcic, M. (2018a). Predicting Users' Personality from Instagram Pictures. *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*, 157–161. https://doi.org/10.1145/3209219.3209248

Ferwerda, B., & Tkalcic, M. (2018b). You are what you post: What the content of instagram pictures tells about users' personality. *CEUR Workshop Proceedings*, *2068*.

Golbeck, J., Robles, C., & Turner, K. (2011). Predicting personality with social media. *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, 253–262. https://doi.org/10.1145/1979742.1979614

Graneheim, U. ., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112.

Harrington, D. (2008). *Confirmatory Factor Analysis* (1st ed.). Oxford University Press.

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. https://doi.org/10.1073/pnas.1917942116

Khanlari, F., & Farhadi, H. (2023). Relationship between psychological variables and psychological symptoms with Internet addiction (a review study). *Journal of Social Psychology*, *10*(66), 1–29.

Lawshe, C. H. (1975). A quantitive approach to content validity. *Personnel Psychology*, *28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393

Mabitsela, T., Geldenhuys, M., & Łaba, K. (2024). Development and validation: Fairness perceptions of broad-based black economic empowerment. *SA Journal of Human Resource Management*, *22*. https://doi.org/10.4102/sajhrm.v22i0.2357

Mayers, A. (2013). *Introduction to Statistics and SPSS in Psychology* (1st ed.). Pearson.

Nikpour, M., Tirgar, A., Ebadi, A., Ghaffari, F., Firouzbakht, M., & Hajiahmadi, M. (2018). Development and psychometric evaluation of a women shift workers' reproductive health questionnaire: study protocol for a sequential exploratory mixed-method study. *Reproductive Health*, *15*(1), 22. https://doi.org/10.1186/s12978-018-0456-0

Ock, J., & An, H. (2021). Machine Learning Approach to Personality Assessment and Its Application to Personnel Selection. *Korean Journal of Industrial and Organizational Psychology*, *34*(2), 213–236. https://doi.org/10.24230/kjiop.v34i2.213-236

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

Shipankar, S., Sawale, G., Shelke, R., & Khairkar, A. (2022). Personality Prediction Using Social Media Platform. *International Research Journal of Engineering and Technology*, *09*(05), 2298–2301.

Singh, S., & Singh, W. (2023). AI-based personality prediction for human well-being from text data: a systematic review. *Multimedia Tools and Applications*, *83*(15), 46325–46368. https://doi.org/10.1007/s11042-023-17282-w

Souri, A., Hosseinpour, S., & Rahmani, A. M. (2018). Personality classification based on profiles of social networks' users and the five-factor model of personality. *Human-Centric Computing and Information Sciences*, *8*(1), 24-39. https://doi.org/10.1186/s13673-018-0147-4

Subramani, B. (2018). Predicting and analyzing personality traits from Instagram features. *Social Media and Business Analytics*, 1–9.

Vaid, S. S., & Harari, G. M. (2021). Who uses what and how often?: Personality predictors of multiplatform social media use among young adults. *Journal of Research in Personality*, *91*, 104005. https://doi.org/10.1016/j.jrp.2020.104005

Valanarasu, M. R. (2021). Comparative Analysis for Personality Prediction by Digital Footprints in Social Media. *Journal of Information Technology and Digital World*, *3*(2), 77–91. https://doi.org/10.36548/jitdw.2021.2.002

Appendix 1
Online Profile-Based Personality Questionnaire (OPBPQ) (EN)

		Questions	Very low	Low	Ave.	High	Very high
N1		I use images with blue and turquoise colors					
	2	I use black and white images in the pictures					
	3	To what extent have you used filters with high brightness in the images?					
	4	To what extent do you consider yourself an active Instagram user?					
	5	How much do you like other people's posts?					
	6	How many fashion and style-related posts do you follow or share?					
	7	How many photography-related posts have you followed or shared?					
	8	To what extent have you followed or shared posts related to personality traits?					
	9	How many lifestyle posts have you followed or shared?					
	10	To what extent have you followed or shared posts related to special personal or cultural, religious, and social occasions?					
E11		To what extent have you used the yellow spectrum in the pictures?					
	12	To what extent have you used the spectrum of blue colors in the pictures?					
	13	How many posts related to music and art have you followed or shared?					
	14	How much have you followed or shared Advertising posts?					
	15	How much have you shared posts related to your personal life?					
	16	How much have you followed or shared posts related to negative social events?					
	17	How much do you consider yourself an active Instagram user?					
	18	How many followers do you have?					
	19	How many accounts do you follow?					
	20	How much do you post and story on Instagram?					
O21		How much have you shared your face in the pictures?					
	22	How much have you shared other people's faces in the pictures?					
	23	How much green color spectrum have you used in the pictures?					
	24	How often do you follow or share posts related to daily news?					
	25	To what extent have you followed or shared posts related to research, articles, and scientific materials?					
	26	To what extent have you followed or shared posts related to the culture and social events of different cities and countries?					
	27	How many posts have you followed or shared about contests, sweepstakes, and prizes?					
	28	To what extent have you followed or shared posts related to human values?					
	29	To what extent have you followed or shared posts related to entertainment and tourist attractions?					
	30	To what extent have you followed or shared food-related posts?					
A31		To what extent have you used black-and-white images?					
	32	To what extent have you used pictures of human faces in the pictures?					
	33	What is the average monthly number of posts you have shared?					
	34	How much would you like to have a user page with a large number of followers?					
	35	How willing are you to follow other people's profiles?					
	36	How much do you want your profile to be personal?					
	37	How many photography-related posts have you followed or shared?					
	38	How many posts related to technology and electronics have you followed or shared?					
	39	How often do you follow or share posts related to global affairs?					
	40	To what extent have you followed or shared posts related to different occupations and professions?					
C41		How much do you feel you need to use Instagram?					
	42	How much do you like other people's posts?					
	43	To what extent have you commented on other people's posts?					
	44	It is pleasant for you to have a lot of followers؟					
	45	It is pleasant for you to follow other people's personal user pages.					
	46	On average, how much time do you spend on Instagram each time?					

47	On average, how many times do you visit Instagram in a day?					
48	To what extent have you followed or shared related to Social communication posts?					
49	How many fashion and style-related posts do you follow or share?					
50	How much do you follow or share posts related to various industries and companies?					

Appendix 2
Online Profile-Based Personality Questionnaire (OPBPQ) (FA)

خیلی زیاد	زیاد	متوسط	کم	خیلی کم	سوالات	
					از تصاویر با رنگ های آبی و فیروزه ای استفاده میکنم.	۱N
					در تصاویر از تصاویر سیاه و سفید استفاده می کنم.	۲
					تا چه میزان در تصاویر از فیلترهایی با روشنایی زیاد استفاده کرده‌اید؟	۳
					تا چه میزان خود را کاربر فعال اینستاگرام می‌دانید؟	٤
					تا چه میزان پست های سایرین را لایک می‌کنید؟	٥
					تا چه میزان پست‌های مرتبط با مد و استایل را دنبال کرده یا به اشتراک گذاشته‌اید؟	٦
					تا چه میزان پست‌های مرتبط با عکاسی را دنبال کرده یا به اشتراک گذاشته‌اید ؟	٧
					تا چه میزان پست‌های مرتبط با ویژگی‌های شخصیتی را دنبال کرده یا به اشتراک گذاشته‌اید ؟	٨
					تا چه میزان پست‌های مرتبط با سبک زندگی را دنبال کرده یا به اشتراک گذاشته‌اید ؟	٩
					تا چه میزان پست‌های مرتبط با مناسبت‌های ویژه شخصی یا فرهنگی، مذهبی و اجتماعی را دنبال کرده یا به اشتراک گذاشته‌اید ؟	۱۰
					در تصاویر تا چه میزان از طیف رنگ زرد استفاده کرده‌اید؟	۱۱E
					در تصاویر تا چه میزان از طیف رنگ های آبی استفاده کرده اید؟	۱۲
					تا چه میزان پست های مرتبط با موسیقی و هنر را دنبال کرده یا به اشتراک گذاشته‌اید؟	۱۳
					تا چه میزان پست هایی که جنبه تبلیغاتی دارند را دنبال کرده یا به اشتراک گذاشته‌اید؟	۱٤
					تا چه میزان پست های مرتبط با زندگی شخصی خود را به اشتراک گذاشته‌اید؟	۱٥
					تا چه میزان پست های مرتبط با وقایع منفی اجتماعی را دنبال کرده یا به اشتراک گذاشته‌اید؟	۱٦
					تا چه میزان خود را کاربر فعال اینستاگرام می‌دانید؟	۱٧
					تعداد فالوور (دنبال‌کننده) های شما چه میزان هستند؟	۱٨
					تعداد فالووینگ (دنبال شونده) های شما چه میزان هستند؟	۱٩
					تا چه میزان با گذاشتن پست و استوری در اینستاگرام فعالیت می کنید؟	۲۰
					در تصاویر تا چه میزان عکس چهره خودتان را به اشتراک گذاشته‌اید؟	۲۱O
					در تصاویر تا چه میزان عکس چهره سایر افراد را اشتراک گذاشته‌اید؟	۲۲
					در تصاویر تا چه میز ان از طیف رنگ سبز استفاده کرده اید؟	۲۳
					تا چه میزان پست های مرتبط به اخبار روزمره را دنبال کرده یا به اشتراک گذاشته‌اید؟	۲٤
					تا چه میزان پست های مرتبط با پژوهش ها و مقالات و مطالب علمی را دنبال کرده یا به اشتراک گذاشته‌اید؟	۲٥
					تا چه میزان پست های مرتبط با فرهنگ و مراسمات اجتماعی شهرها و کشورهای مختلف را دنبال کرده یا به اشتراک گذاشته‌اید؟	۲٦
					تا چه میزان پست های مرتبط با مسابقات و قرعه کشی ها و جایزه ها را دنبال کرده یا به	۲٧

					اشتراک گذاشته‌اید؟	
					تا چه میزان پست های مرتبط با ارزش‌های انسانی را دنبال کرده یا به اشتراک گذاشته‌اید؟	۲۸
					تا چه میزان پست های مرتبط با تفریح و سرگرمی ها و جاذبه های توریستی را دنبال کرده یا به اشتراک گذاشته‌اید؟	۲۹
					تا چه میزان پست های مرتبط با غذاها و خوراکی ها را دنبال یا به اشتراک گذاشته‌اید؟	۳۰
					تا چه میزان از تصاویر سیاه و سفید استفاده کرده‌اید؟	۳۱A
					در تصاویر تا چه میزان از عکس چهره انسان استفاده کرده‌اید؟	۳۲
					میانگین ماهانه تعداد پست هایی که به اشتراک گذاشته اید چه میزان است؟	۳۳
					تا چه میزان تمایل دارید تا صفحه کاربری با تعداد فالوور (دنبال کننده) زیاد داشته باشید؟	۳۴
					تا چه میزان تمایل به دنبال کردن صفحه کاربری شخصی دیگران دارید؟	۳۵
					تا چه میزان تمایل دارید صفحه کاربریتان شخصی باشد؟	۳۶
					تا چه میزان پست‌های مرتبط با عکاسی را دنبال کرده یا به اشتراک گذاشته‌اید؟	۳۷
					تا چه میزان پست‌های مرتبط با تکنولوژی و الکترونیک را دنبال کرده یا به اشتراک گذاشته‌اید؟	۳۸
					تا چه میزان پست‌های مرتبط با شرایط جهانی را دنبال کرده یا به اشتراک گذاشته‌اید؟	۳۹
					تا چه میزان پست‌های مرتبط با مشاغل و حرفه‌های مخالف را دنبال کرده یا به اشتراک گذاشته‌اید؟	٤٠
					تا چه میزان برای استفاده از اینستاگرام احساس نیاز می کنید؟	٤۱C
					تا چه میزان پست های سایرین را لایک (تایید) می کنید؟	٤۲
					تا چه میزان برای پست‌های سایرین پیام (کامنت) گذاشته‌اید؟	٤۳
					داشتن فالوور (دنبال کننده) زیاد برای شما خوشایند است.	٤٥
					دنبال کردن صفحه کاربری شخصی دیگران برای شما خوشایند است.	٤٥
					به طور میانگین، در هر بار مراجعه به اینستاگرام، چه میزان وقت در آن می‌گذرانید؟	٤٦
					به طور میانگین، در یک شبانه روز، چند مرتبه به اینستاگرام مراجعه می‌کنید؟	٤۷
					تا چه میزان پست های مرتبط با ارتباطات اجتماعی را دنبال کرده یا به اشتراک گذاشته‌اید؟	٤۸
					تا چه میزان پست‌های مرتبط با مد و استایل را دنبال یا به اشتراک گذاشته‌اید؟	٤۹
					تا چه میزان پستهای مرتبط با صنایع و شرکت‌های مخالف را دنبال کرده یا به اشتراک گذاشته‌اید؟	٥٠