



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *UMAP '20: 28th ACM Conference on User Modeling, Adaptation and Personalization, Genoa, Italy, 12-18 July, 2020*.

Citation for the original published paper:

Ferwerda, B., Tkalčič, M. (2020)

Exploring the Prediction of Personality Traits from Drug Consumption Profiles

In: *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 2-5). Association for Computing Machinery (ACM)

<https://doi.org/10.1145/3386392.3397589>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-50067>

# Exploring the Prediction of Personality Traits from Drug Consumption Profiles

Bruce Ferwerda

Department of Computer Science and Informatics  
School of Engineering  
Jönköping University  
Jönköping, Sweden  
bruce.ferwerda@ju.se

Marko Tkalčič

Faculty of Mathematics, Natural Sciences and Information  
Technologies (FAMNIT)  
University of Primorska  
Koper, Slovenia  
marko.tkalcic@famnit.upr.si

## ABSTRACT

The number of people that have been in touch with drugs is continuously increasing. Excessive intake of drugs becomes problematic when it turns into disorderly behaviors, such as addictions. In order to treat these disorderly behaviors, treatment plans often adhere to a one-size-fits-all approach with fixed and standardized steps. However, for effective treatment of disorderly behaviors it has been acknowledged that personalized treatment programs are necessary. The personality of people has been argued to be a factor that plays an important role in setting up effective treatment plans. In this work we explored the predictability of people's personality traits based on their drug consumption profile. Based on self-reported consumption frequencies of "abusable psychoactive drugs," we found among 1878 respondents that drug consumption profiles can be used to predict people's personality traits. The prediction of personality traits can be used to circumvent intruding questionnaires and to implicitly create personalized treatment programs.

## CCS CONCEPTS

• **Information systems** → *Recommender systems*; • **Human-centered computing** → *User models*; *User studies*.

## KEYWORDS

Personality, Drug Consumption, User Modeling, Personalized Treatment Plans, Disorderly Behaviors, Addiction

## ACM Reference Format:

Bruce Ferwerda and Marko Tkalčič. 2020. Exploring the Prediction of Personality Traits from Drug Consumption Profiles. In *Adjunct Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20 Adjunct)*, July 14–17, 2020, Genoa, Italy, Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, Article 4, 4 pages. <https://doi.org/10.1145/3386392.3397589>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*UMAP '20 Adjunct*, July 14–17, 2020, Genoa, Italy

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-7950-2/20/07.  
<https://doi.org/10.1145/3386392.3397589>

## 1 INTRODUCTION

Global trends in drug usage show that the number of people engaging in some kind of drug<sup>1</sup> intake is steadily increasing. Estimates from 2016 show that globally over a quarter of a billion people have used some kind drug in their life. This is an increase of 20% compared to estimates from 2015. Adolescents have often experimented with drugs through their membership to a subculture or peer group in which the usage is part of their identity search [20]. Especially a prolonged use of drugs is a risk behavior that can turn into a disorder,<sup>2</sup> which consequentially has negative effects on health, mortality, and morbidity [15]. The 2016 estimates show that one out of nine drug users suffer from a drug disorder (roughly 30.5 million people in 2016) [16]. Given the increase of global trends, the drug disorder estimate is currently probably higher as well.

The treatment of (addictive) disorders are supported through a variety of facilities, such as the Alcoholics Anonymous (AA). These facilities often provide treatment in the combined form of "moral pedagogy" (i.e., well-meant advice) and social work. These treatments often follow a predefined number of steps to recovery through standardized treatment strategies. However, in recent years the acknowledgment of personalized treatment became increasingly apparent. It has been acknowledged by fundamental research that humans show significant differences on every discernable level [20]. Hence, for treatments to succeed, they need to be personalized to individual needs. Although personalized treatments plans have been argued for, to personalize for each individual separately is costly to develop and maintain. Opposed to a one-size-fits-all approach, a "higher level" of personalization (or a "lower level" of generalization) may balance between cost efficiency and personalized treatments plans for each individual separately.

A way to achieve such a higher level of personalization is through the use of personality traits. The incorporation of personality traits have been argued to be of importance when personalizing treatments; especially given its relation with initiation, maintenance and treatment of drug misuse [19]. On a more general level, personality traits have shown to be related to individual preferences, behaviors, and needs (e.g., [5]). Hence, personality traits have already

<sup>1</sup>A drug is defined as a chemical intake that influences biological functions other than through nutrition and hydration [12]. Where drugs can be divided by their effect on different biological functions, the dataset that is used in the current study comprises only "abusable psychoactive drugs," which relate to drugs that have a pleasant or interesting effects on the mental state. Hence, are often taken for other reasons than for a specific malady relief.

<sup>2</sup>A disorder is identified as an excessive intake of drugs to a point that users reach dependence and require treatment.

been increasingly proposed for personalizing recommender systems in diverse domains (e.g., [4, 7, 9, 13]), and has proven to gain a personalization advantage over non-personality based systems [9].

In order to support personality-based personalization, attempts have been made to implicitly acquire personality traits to circumvent the use of extensive and intruding questionnaires (e.g., [4, 6]). For example, data that becomes available through the increased interconnectedness of services (e.g., single sign-on mechanisms with social networking services) has been exploited to predict personality traits from. In this paper we explore the prediction of personality from drug-consumption data. We show that drug consumption profiles consist of information that is useful for personality prediction. Hence, personality can be acquired given a drug consumption profile and thereby a personalized treatment plan can be provided.

## 2 RELATED WORK

Relying on psychological theories and models has gained an increased interest in recommender systems for personalization purposes. The use of psychological theories/models allows for a better and deeper understanding of the root cause of users' behaviors, preferences, and needs [8], which in turn allows for findings to easier being applied cross-domains [1].

As personality is considered to be reflected in human behaviors, it has been used as a common psychological model for personalization. For example, Ferwerda & Schedl [4] are using personality traits to personalize music recommendations. Lee & Ferwerda [13] proposes the use of personality traits to tailor online educational tools to better suit the learning style. When it comes to drug usage, research has shown the importance of personalizing treatment programs [20]. For example, related to personality, it has been shown that the treatment of people with a high reward-sensitivity has shown a link with tangible rewards (e.g., money). Another example is the treatment of people with rash impulsivity, which includes training in problem-solving skills with a focus on identifying strategies that allow time for reflection and thereby enabling opportunities for alternative responses [19].

There are several ways to acquire personality traits of users [3]: 1) explicitly through personality quizzes and questionnaires, or 2) implicitly by inferring personality traits through behavioral data. With the former an increased accuracy can be achieved, but is more intrusive and time consuming as users need to participate in the quiz/questionnaire. The latter is more unobtrusive, but accuracy is often compromised as it depends on the quality of the source (e.g., frequency of capturing behavioral data). Recently, research has been focusing on improving explicit personality acquisition; especially with the rise of external data sources (e.g., social networking services) that are increasingly connected with applications through single-on mechanisms [4]. These mechanisms are retrieving user information from an external data source for the user to conveniently register or sign in to the application. The data that becomes available can be exploited to implicitly acquire personality traits. For example, Quercia et al. [17] looked at the characteristics of Twitter messages to make personality trait estimations. Skowron et al. [18] used a combination of Twitter and Instagram information to increase the accuracy of personality trait predictions. By looking at personality predictions from drug consumption profile,

| Drugs          | Drugs       | Drugs                          |
|----------------|-------------|--------------------------------|
| Alcohol        | Cocaine     | LSD                            |
| Amphetamines   | Caffeine    | Methadone                      |
| Amyl           | Crack       | Mushrooms                      |
| Nitrite        | Ecstasy     | Nicotine                       |
| Benzodiazepine | Heroin      | Semerone (fictitious drug)     |
| Cannabis       | Ketamine    | Volatile substance abuse (VSA) |
| Chocolate      | Legal highs |                                |

**Table 1: Drugs included in the dataset**

| General dimensions     | Primary factors   |
|------------------------|---|
| Openness to experience | Artistic, curious, imaginative, insightful, original, wide interest |
| Conscientiousness      | Efficient, organized, planful, reliable, responsible, thorough      |
| Extraversion           | Active, assertive, energetic, enthusiastic, outgoing, talkative     |
| Agreeableness          | Appreciative, forgiving, generous, kind, sympathetic, trusting      |
| Neuroticism            | Anxious, self-pitying, tense, touchy, unstable, worrying            |

**Table 2: Five-factor model**

personality traits can be implicitly acquired. Hence, time consuming questionnaires can be avoided to provide personalized treatment plans to treat disorderly behaviors.

## 3 METHOD

To explore the predictability of personality traits from drug consumption profiles, we used the dataset of Fehrman et al. [2]. The dataset was collected between March 2011 and March 2012 with a total of 1885 participants. Participants were asked to respond to 20 legal and illegal drug (see Table 1 for the drugs that were included in the dataset) consumption questions ("Never used", "Used over a decade ago", "Used in last decade", "Used in last year", "Used in last month", "Used in last week", and "Used in last day") and additionally to fill in the NEO-FFI-R questionnaire, which measures the five personality traits (openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism) through 60-items [14] (see Table 2 for a description of each of the five personality factors).

The original dataset included a fictitious drug (Semeron) to identify over-claimers. We used this variable to identify invalid responses. By removing positive responses to the intake of Semeron, we ended with a dataset that consists of 1879 valid responses for training our personality prediction model.

## 4 RESULTS

We used the learner-based feature selection to select the best features (i.e. drugs) to train our model for predicting personality traits from the self-reported drug intake. All the 19 drugs (see Table 1) that were used in the dataset showed to be relevant features for the prediction of each personality trait. Hence, no further selection of the features was needed.

|   | ZeroR | M5' rules   | Random forest | RBF network |
|---|-------|-------------|---------------|-------------|
| O | 0.98  | <b>0.85</b> | <b>0.88</b>   | <b>0.83</b> |
| C | 1.00  | <b>0.92</b> | <b>0.92</b>   | <b>0.88</b> |
| E | 0.99  | <b>0.95</b> | <b>0.97</b>   | <b>0.94</b> |
| A | 0.99  | <b>0.94</b> | <b>0.95</b>   | <b>0.89</b> |
| N | 0.98  | <b>0.95</b> | <b>0.96</b>   | <b>0.93</b> |

**Table 3: Different predictive models for each personality trait. ZeroR classifier represents the baseline. The boldfaced numbers indicate an out-performance of the baseline. Root-mean-square error (RMSE) is reported ( $r \in [1,5]$ ) to indicate prediction performance of the personality traits: (O)penness to experience, (C)onscientiousness, (E)xtraversion, (A)greeableness, and (N)euroticism.**

The ZeroR majority classifier was used as the baseline model. Three different classifiers, chosen on the base of previous works on personality prediction, were used and compared against the baseline: M5' rules, random forest, and radial basis function (RBF) network (see Table 3 for the results). These classifiers are commonly used and have shown to perform well in personality prediction in previous works (e.g., [4, 6, 17]).

We trained our models with the aforementioned classifiers in Weka using 10-fold cross-validation with 10 iterations. For each classifier, we report the root-mean-squared error (RMSE) in Table 3 to indicate the difference between the predicted and observed values. The RMSE of each personality trait relates to a [1,5] score scale. We first started to train our predictive model with the M5' rules. This is a classifier that has shown to be an effective classifier in previous works on personality prediction (e.g., [4, 6, 17]). The M5' rules outperform the baseline model in predicting all of the personality traits. To further explore possible improvements by other classifiers, we tried out the random forests classifier. Random forests are known to have a reasonable performance when the features consist of high amounts of noise [10]. Unfortunately, no further improvements were able to be made. The results show that some personality traits were able to be predicted better than the M5' rules, whereas others were more challenging. In an attempt to further improve the prediction model, we used the RBF network classifier. The RBF network is a neural network that has shown to work well on smaller datasets [11].

All of the predictive models outperform the ZeroR baseline model in predicting all of the personality traits, with the M5' rules performing slightly better than the random forest, and the RBF outperforming both M5' rules and the random forest.

#### 4.1 Comparison with other predictive models

To be able to put the results into perspective we compare our results with other personality predictive models of prior work in other domains that are using RMSE to measure performance (see Table 4 for the comparison).

Although the overall performance between different datasets differ, similar trends can be observed on the predictability of each personality trait. The performance of the current dataset does not reach the levels of other datasets, but similar trends are observable

| Personality traits | RMSE            |      |      |      |
|--------------------|-----------------|------|------|------|
|                    | Current results | [6]  | [4]  | [17] |
| O                  | 0.83            | 0.71 | 0.73 | 0.69 |
| C                  | 0.88            | 0.62 | 0.73 | 0.73 |
| E                  | 0.94            | 0.98 | 0.99 | 0.96 |
| A                  | 0.89            | 0.61 | 0.73 | 0.78 |
| N                  | 0.93            | 0.89 | 0.83 | 0.97 |

**Table 4: Comparison of personality prediction compared to prior work on Instagram [6], Facebook [4], and Twitter [17] that are using a similar performance measurement. Root-mean-square error (RMSE) is reported ( $r \in [1,5]$ ) to indicate prediction performance of the personality traits: (O)penness to experience, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism.**

concerning the predictability of each personality trait. Given that the other datasets are much richer in captured behaviors of people then the dataset in this work, the slight lack in performance does not come as a surprise. Despite of that, similar trends in the results are observable: the most difficult traits to predict are extraversion and neuroticism. With a less rich dataset, we show that personality prediction is possible.

## 5 CONCLUSION & DISCUSSION

In this work we explored the prediction of personality traits from drug consumption profiles. By testing drug profiles based on the consumption of 19 legal and illegal abusable psychoactive drugs, we found that personality traits can be predicted from this self-reported data. For the prediction of all personality traits, we were able to outperform the baseline model by using three different classifiers: M5' rules, random forest, and RBF network. Out of the three tested classifiers, we found that the RBF network performs best across the classifiers that were used. This seems to be in line with findings of prior work that tested multiple classifiers as well (e.g., [4, 6]). The RBF network has shown to perform well on smaller datasets that are similar in size as used in the current work. Of all personality traits, the prediction of the openness to experience trait gained the highest improvement over the baseline model, while agreeableness gained very little. These results are in line with the performances of prior works in personality prediction (e.g., [6, 17]).

## 6 LIMITATIONS & FUTURE WORK

Although we did not conduct any extensive feature engineering, we were already able to create predictive models that could outperform the baseline model. For future work we will try to improve the current models by creating better features. We see three different situations that need to be addressed: 1) The current dataset does not take into account the consumption frequency related to the drug characteristic (e.g., the consequences of excessive intake of such a drug). For example, caffeine consumption may occur way more frequent than the frequency of consumption of cocaine, 2) We did not look at different age groups (e.g., some drugs may be more "popular" than others), and 3) The nature of the drug consumption question may have introduced a confound for the participants.

Participants were asked to provide responses to a 7-point consumption frequency measurement ("Never used", "Used over a decade ago", "Used in last decade", "Used in last year", "Used in last month", "Used in last week", and "Used in last day"). The recollection of some events may have been too far in the past for participants to recall them accurately and assign the correct time span to it. The aforementioned situations can partly be addressed through applying different coding schemes to the consumption frequency responses (e.g., combining scales) and/or giving weights by taking into account the drug characteristic. Furthermore, the current dataset only takes into account abusable psychoactive drugs. The inclusion of other kind of drugs could contribute to a richer drug consumption profile, and may further improve the personality prediction possibilities.

Although some research has investigated treatment plans based on personality traits, research should further explore specific personality-based treatment opportunities given the possibilities that this work has shown in implicit personality acquisition.

## REFERENCES

- [1] Iván Cantador, Ignacio Fernández-Tobías, and Alejandro Bellogin. 2013. Relating personality types with user preferences in multiple entertainment domains. In *CEUR workshop proceedings*. Shlomo Berkovsky.
- [2] Elaine Fehrman, Awaz K Muhammad, Evgeny M Mirkes, Vincent Egan, and Alexander N Gorban. 2017. The Five Factor Model of personality and evaluation of drug consumption risk. In *Data Science*. Springer, 231–242.
- [3] Bruce Ferwerda and Markus Schedl. 2014. Enhancing Music Recommender Systems with Personality Information and Emotional States: A Proposal. In *Umap workshops*.
- [4] Bruce Ferwerda and Markus Schedl. 2016. Personality-based user modeling for music recommender systems. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 254–257.
- [5] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. 2015. Personality & emotional states: Understanding users' music listening needs. CEUR-WS. org.
- [6] Bruce Ferwerda and Marko Tkalcic. 2018. Predicting Users' Personality from Instagram Pictures: Using Visual and/or Content Features?. In *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*. ACM, 157–161.
- [7] Bruce Ferwerda, Emily Yang, Markus Schedl, and Marko Tkalcic. 2019. Personality and taxonomy preferences, and the influence of category choice on the user experience for music streaming services. *Multimedia Tools and Applications* (2019), 1–34.
- [8] Mark Graus and Bruce Ferwerda. 2019. *Theory-Grounded User Modeling for Personalized HCI*. DeGruyter Oldenbourg. In Press.
- [9] Rong Hu and Pearl Pu. 2011. Enhancing collaborative filtering systems with personality information. In *Proceedings of the fifth ACM conference on Recommender systems*. ACM, 197–204.
- [10] Elizabeth M Humston, Joshua D Knowles, Andrew McShea, and Robert E Synovec. 2010. Quantitative assessment of moisture damage for cacao bean quality using two-dimensional gas chromatography combined with time-of-flight mass spectrometry and chemometrics. *Journal of Chromatography A* 1217, 12 (2010), 1963–1970.
- [11] Lav R Khot, Suranjan Panigrahi, Curt Doetkott, Young Chang, Jacob Glower, Jayendra Amamcharla, Catherine Logue, and Julie Sherwood. 2012. Evaluation of technique to overcome small dataset problems during neural-network based contamination classification of packaged beef using integrated olfactory sensor system. *LWT-Food Science and Technology* (2012).
- [12] Mark AR Kleiman, Jonathan P Caulkins, and Angela Hawken. 2011. *Drugs and Drug Policy: What Everyone Needs to Know*. Oxford University Press.
- [13] Michael J Lee and Bruce Ferwerda. 2017. Personalizing online educational tools. In *Proceedings of the 2017 ACM Workshop on Theory-Informed User Modeling for Tailoring and Personalizing Interfaces*. ACM, 27–30.
- [14] Robert R McCrae and Paul T Costa Jr. 2004. A contemplated revision of the NEO Five-Factor Inventory. *Personality and individual differences* 36, 3 (2004), 587–596.
- [15] J Michael McGinnis and William H Foege. 1993. Actual causes of death in the United States. *Jama* 270, 18 (1993), 2207–2212.
- [16] United Nations Office on Drugs and Crime. 2019. World Drug Report. (2019). [https://www.unodc.org/wdr2018/prelaunch/WDR18\\_Booklet\\_2\\_GLOBAL.pdf](https://www.unodc.org/wdr2018/prelaunch/WDR18_Booklet_2_GLOBAL.pdf) accessed: 03 Apr 2019.
- [17] Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. 2011. Our twitter profiles, our selves: Predicting personality with twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE, 180–185.
- [18] Marcin Skowron, Marko Tkalcic, Bruce Ferwerda, and Markus Schedl. 2016. Fusing social media cues: personality prediction from twitter and instagram. In *Proceedings of the 25th international conference companion on world wide web*. International World Wide Web Conferences Steering Committee, 107–108.
- [19] Petra K Staiger, Nicolas Kambouropoulos, and Sharon Dawe. 2007. Should personality traits be considered when refining substance misuse treatment programs? *Drug and alcohol review* 26, 1 (2007), 17–23.
- [20] Jaap van der Stel. 2015. Focus: Personalized Medicine: Precision in Addiction Care: Does It Make a Difference? *The Yale journal of biology and medicine* 88, 4 (2015), 415.