Predicting Openness to Experience via a Multiplex Cognitive Network Approach

Gal Samuel[1], Massimo Stella[2], Roger E. Beaty[3], & Yoed N. Kenett[1]

[1] Faculty of Data and Decision Sciences, Technion – Israel Institute of Technology, Haifa, Israel

[2] Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy

[3] Department of Psychology, Pennsylvania State University, University Park, PA, USA

Corresponding author:

Yoed N. Kenett

Technion – Israel Institute of Technology

Faculty of industrial Engineering & Management

Haifa, Israel 3200003

Email: yoedk@technion.ac.il

Phone number: +972-4-48294430

Abstract

Openness to Experience is most strongly related to aspects of high-level cognition, such as creativity. Yet, the role of cognitive capacities in Openness is still far from understood. We examine how individuals search their memory predicts levels of Openness. Participants (N = 163) had one minute to generate synonyms to the word *hot*, operationalized as mental navigation over a multidimensional representation of the mental lexicon – a cognitive multiplex network. We find high accuracy in low- and high- Openness group classification, and good prediction of individual differences in Openness. These results support the use of computational cognitive modelling for the study of personality traits. Further, our results suggest that people high in Openness engage in a distinct style of cognitive search.

Keywords: Openness to Experience, multiplex networks, prediction, classification

Predicting Openness to Experience via a Multiplex Cognitive Network Approach

1. Introduction

Being open to new experiences is an important personality trait, related to curiosity, creativity, a drive to learn new things, and having diverse hobbies. Among the Big-5 personality traits, Openness to Experience is the personality trait that is most strongly related to cognitive capacities (for additional references, see Zillig et al., 2002), including intelligence, working memory, semantic memory, and creativity (Christensen, Kenett, et al., 2018; DeYoung et al., 2012; Kaufman et al., 2010; Kaufman et al., 2016). As such, this personality trait is complex in nature, being related to different aspects of higher-level cognition. It is commonly measured via several self-report personality questionnaires (Christensen, Cotter, et al., 2018). Given the significance of this personality trait, is it possible to predict it from simpler, behavioral tasks, that do not rely on self-report? In the current study, we explore whether Openness can be predicted from performance on a brief verbal fluency task, via a computational cognitive multiplex network model. The success of such a prediction model will further strengthen the relation between personality and cognition, as well as the role of semantic memory in Openness.

1.1. Openness to Experience and Semantic Memory Networks

Openness to Experience has been most strongly linked to creativity (Oleynick et al., 2017), to the point that it has been considered the "creativity personality trait" (Johnson, 1994). In fact, Openness to Experience is the most consistent predictor of creative achievements across the arts and sciences (Feist, 1998; Kaufman et al., 2016). Another defining characteristic of people high in Openness to Experience is engaging in a variety of experiences that lead to the acquisition of broader general knowledge. In general, such people tend to be curious and are motivated to learn and acquire new knowledge (Kashdan et al., 2004; Silvia & Sanders, 2010; von Stumm, 2018). These characteristics impact their general knowledge, stored in semantic memory—the cognitive system that stores facts and concepts (Kumar, 2021). Since measuring Openness is based on various types of subjective self-report questionnaires (Christensen, Cotter, et al., 2018), aspects of semantic memory may provide a valuable, objective measure to predict individual differences of this personality trait.

Recent computational advances have paved the way to study semantic memory structure and processes that operate over it (Hills & Kenett, 2022). One such approach that has been gaining popularity is cognitive network science (Baronchelli et al., 2013; Borge-Holthoefer & Arenas, 2010; Castro & Siew, 2020; Siew et al., 2019). Cognitive network science applies computational network science methodologies, that are based on mathematical graph theory, to represent and investigate the complexity of cognitive systems (such as language and memory). These computational tools have been applied to study broad cognitive domains, such as language, memory, learning, aging, and creativity (Siew et al., 2019). In relation to creativity, several studies have shown how creativity is related to a more flexible, richly connected semantic memory structure, both at the group (Kenett et al., 2014; Kenett et al., 2016) and individual (Benedek et al., 2017; He et al., 2021; Ovando-Tellez, Kenett, et al., 2022) levels. Such a memory structure is theorized to facilitate creative search behavior (Kenett, forthcoming) as well as cognitive flexibility (Kenett et al., 2018). In relation to Openness to Experience, Christensen, Kenett et al. (2018) have shown how people high on Openness to Experience exhibit a flexible, richly connected semantic memory network, similarly found in highly creative people.

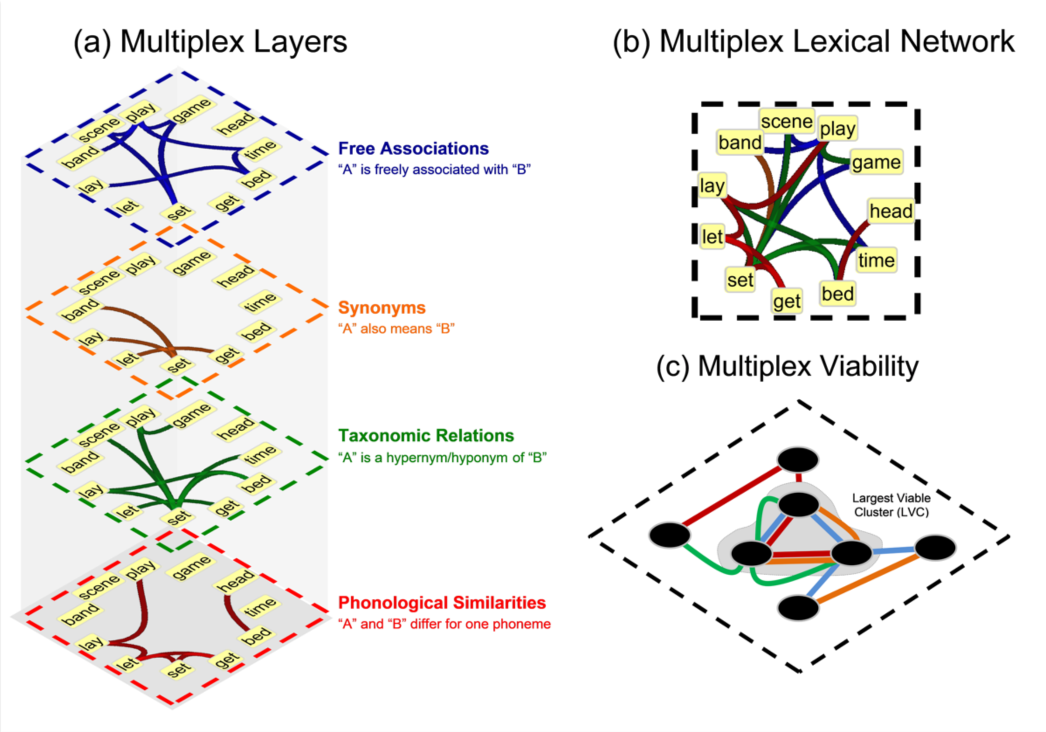
1.2. Cognitive Multiplex Networks

However, cognitive network science research largely studies single types of networks, e.g., semantic networks or phonological networks (Siew et al., 2019). Yet, the mental lexicon is considered to be a multidimensional structure with different layers, that reflects different types of features (e.g., phonology, semantics, etc.). Thus, studying multidimensional, or multilayer, cognitive networks is needed to advance our understanding of the complexity of the human mind (Hills & Kenett, 2022). A multiplex network is a mathematical structure composed of a number of independent networks, or layers, and the overlapping nodes between them. The multiplex network preserves the links from all independent layers and also merges the independent layers into one multiplex network (Levy et al., 2021; Stella, 2019; Stella et al., 2017; Stella et al., 2018; **Figure 1**).

Research using a more complex multidimensional cognitive structure recently demonstrated its ability to classify low- and high- creative individuals, based on a simple semantic fluency task (Stella & Kenett, 2019). To achieve such a creativity classification, the authors used a multiplex network to represent lexical memory in a broad fashion. In Stella & Kenett (2019), the layers of their multiplex network represented lexical information, consisting of a synonyms layer, a phonological layer, an associative layer, and a hypernym/hyponym layer (**Figure 1).** Using all these layers together computationally allowed examining the way people exploit their memory, and classify participants into low- and high- creative individuals, based on the way they "walked" on the multiplex network (Stella & Kenett, 2019). Importantly, the authors focused on the Largest Viable Cluster (LVC) to measure participants' performance on a verbal fluency task. The LVC is a component of the multiplex network made of the largest collection of nodes which are connected between themselves across all independent layers of the multiplex (Stella et al., 2018). In short, Stella and Kenett (2019) found that low- and high- creative individuals significantly differ in the way they rely on the LVC when generating animal category members, and in the number of responses they are able to generate: The authors show that high-creative individuals generate more responses which are less retrieved from the LVC.

Figure 1.

Example visualization of the cognitive multiplex network



Note. (A) In the multiplex structure, nodes represent concepts replicated across four layers, namely free associations, synonyms, taxonomic relations, and phonological similarities. (B) All layers can be condensed in one edge-colored network, where links of multiple colors co-exist. Each color represents one layer (e.g., red for phonological similarities). (C) In these edge-colored networks, the largest viable cluster (LVC) is the largest set of nodes that are simultaneously connected across all layers. There must always be at least one sequence of links connecting any two nodes in an LVC, for every layer in the network.

1.3. The Present Research

The study by Stella and Kenett (2019) was largely a proof of concept, analyzing a small dataset to examine the overall success of a cognitive multiplex network to predict such complex behavior. In the current study, we aim to replicate their method and extend it by predicting Openness, using cognitive multiplex network's properties and a machine learning model. Specifically, we aim to predict a complex cognitive trait (Openness) using a simple behavioral task (verbal fluency). While Stella and Kenett (2019) focused on creativity, this study focuses on predicting Openness, since both Openness and creativity are complex cognitive traits, and Openness is specifically known for its relation to creativity (Christensen, Cotter, et al., 2018; Lee & Ashton, 2004).

Similar to Stella and Kenett (2019), we analyze performance in a semantic fluency task, as an operationalization of a mental navigation task that operates over memory when searching internally (Benigni et al., 2021; Todd & Hills, 2020). In this task, participants are required to generate as many category members as possible, in a given amount of time. Computational methods allow examining how people search through their memory (Abbott et al., 2015; Hills et al., 2012; Hills et al., 2015), tracing the paths they traverse over representations of their mental lexicon (Benigni et al., 2021). Often, this task is based on the *animal* category (Hills et al., 2012), but other categories are also used such as fruits and vegetables (Borodkin et al., 2016), as well as synonyms to words as *cold* or *hot* (Beaty et al., 2014).

In this study, we analyze data collected by Beaty et al. (2018) where participants completed a fluency task and a personality assessment. Similar to Stella and Kenett (2019), we apply a supervised machine learning model based on a cognitive multiplex network. The use of supervised learning models for predictions is becoming increasingly popular in the clinical and computational sciences. For example, supervised learning models have been used to predict hospitalization due to heart diseases (Dai et al., 2015), and geological phenomena (Lin et al., 2018). However, to the best of our knowledge, limited cognitive research has been conducted with similar machine learning approaches. As described above, ample evidence links Openness with creativity as well as each of these constructs to semantic memory (Christensen, Kenett, et al., 2018; Kenett & Faust, 2019). Thus, based on Stella and Kenett (2019), we expect that a cognitive multiplex network may be similarly used to predict Openness scores. We expect to find differences in how low- and high-Openness groups rely on the LVC on the fluency task, thus solving a classification task. Furthermore, we expect our machine learning model to predict individual-based differences in Openness, thus solving a regression task. Together, these two approaches can shed light on the cognitive correlates of Openness, revealing how highly Open people engage in memory search and retrieval. Based on Stella and Kenett (2019), we predict that the cognitive multiplex network model will allow to accurately classify between low- and high- Open people, and that high-Open people will generate more responses in a semantic fluency task, responses that less rely on the LVC. In a more exploratory fashion, we predict that our model will be able to capture individual-based differences in Openness scores.

2. Methods

2.1. Participants

We reanalyze data collected by Beaty et al. (2018). The total sample consisted of 163 participants recruited from the University of North Carolina at Greensboro (UNCG) and the surrounding community (113 women, mean age = 22.50 years, SD = 5.79) and specifically over-sampled art, music, and science majors to increase the sample's population of creative domains. A recent meta-analysis shows that this sample size corresponds to the sample sizes of previous studies that attempted to predict Big-5 personality traits (Azucar et al., 2018). Participants were recruited as part of a larger study on individual differences in creativity (which involved numerous laboratory and ecological measures and procedures not discussed here) and were paid up to $100 for their time. All participants were right-handed with normal or corrected-to-normal vision and reported no history of neurological disorder, cognitive disability, or medication that affects the central nervous system. Participants provided written informed consent. The study was approved by the UNCG Institutional Review Board (Beaty et al., 2018).

2.2. Materials

2.2.1. Behavioral Assessments

2.2.1.1. Associational fluency (*hot* synonyms). Participants were required to list (type) as many synonyms to the word *hot* as they could, in one minute. Based on Silvia et al. (2013), the instructions participants were given is as follows: "On the next screen, you'll be asked to write as many different synonyms for HOT as you can. You will have 1 minute. Please press ENTER after typing each word.". No explicit definition of what a synonym is was given. Responses that were not synonyms to *hot* were preprocessed via the SemNA package, including automatic spelling corrections and exclusion of irrelevant responses (Christensen & Kenett, 2021b). All valid responses were saved, and responses were preprocessed to fix typos, nonsensical responses, and remove repetitions, using the SemNA pipeline (Christensen & Kenett, 2021b). Synonyms to the word hot were based on synonyms retrieved from online searches in Google (e.g., www.google.com) and thesaurus (e.g., https://www.thesaurus.com/browse/hot) websites.

2.2.1.2. Personality assessmen*t*. The Big Five personality traits (and corresponding facets) were assessed using the 240 item NEO-PI-3 questionnaire (McCrae et al., 2005). Of the NEO-PI-3 items that measure Openness, we preformed confirmatory factor analysis via the WSLMV estimator to estimate factor scores. This analysis was conducted using the *lavaan* package in R (version 0.6.10; Rosseel, 2012); these factors scores were used for subsequent prediction analyses.

2.2.2. Group assignment

Based on Christensen et al. (2018) and in an attempt to replicate Stella and Kenett (2019), we first divide participants into low- and high- Openness classes based on their Openness factor score percentile. The percentile threshold (X), separating across the classes, was fixed through a parameter sweep that examined the difference between the low- and high- groups, using a non-parametric Mann-Whitney U-test, starting with X equal to .025 and proceeding with steps of .025 up until .5. The value of X maximizing these statistics was selected as a candidate for partitioning the data. This value was relative to X = .30, so that the “low” class became the cluster of individuals falling in the lower 30th percentile of the distribution of Openness (N = 46) and the “high” class became the cluster of individuals with the highest 30% of scores in the same distribution (N = 46).

In addition to this “extreme groups” classification analysis, we also conducted a regression analysis with the full sample. Such analysis circumvents issues of splitting individuals who vary on a continuous score into groups, and allows us to analyze the full data. Thus, we first follow Stella and Kenett (2019) and conduct a classification analysis. We then extend this line of research by conducting a prediction analysis in relation to the full dataset.

2.2.3. Cognitive multiplex analysis

2.2.3.1. Multiplex construction. Our multiplex network consisted of four networks, that constituted its layers: Free associations, synonyms, phonological, and hypernyms/hyponyms. To create the cognitive multiplex network, all of the layers were treated as undirected. Data for all layers except for free associations was obtained from WordData repository presented by WolframResearch, Champaign, IL, US, and available through Mathematica 11.3 program. The WordData dataset is based on WordNet 3.0 (Miller, 1995). WordNet 3.0 is a dictionary that includes information about word-word similarities as computed from English dictionaries (Stella et al., 2018). Specifically, the multiplex includes the following four layers:

- **Free associations layer:** created using data of associations elicited by participants, from the Small World of Words project (De Deyne et al., 2019). Only links that were elicited more than 10 times were considered eligible, for the association layer to feature the same link density of other multiplex layers.
- **Synonym layer:** consists of word-word relations that represent meaning overlapping between the words, such as *hot* and *warm*.
- **Phonological layer:** consists of word-word relations that represent one phoneme difference between words, such as *cat* (kæt) and *bat* (bæt).
- **Hypernyms/Hyponyms layer**: consists of word-word relations that represent generalization and specification, such as *bird* and *eagle*.

2.2.3.2. Multiplex measures. After creating the multiplex, we computed the Largest Viable Cluster (LVC, see Baxter et al. 2016) which is the largest cluster of words that are connected across all layers. Similar to Stella and Kenett (2019), we compute several cognitive multiplex network measures and use multivariate statistical analyses to determine which variables significantly differ across the groups.

We used the fluency task responses of each participant to identify where the participant "walks" on the network, and compute multiple measures for each participant (Stella & Kenett, 2019; **Table 1**). The measures focus on aspects such as the interaction of each participant with the LVC, the entropy of paths participants used in their mental navigation, the amounts of responses each participant generated in general, within, and outside of the LVC (see **Table 1** and **Figure 2** for an illustration).

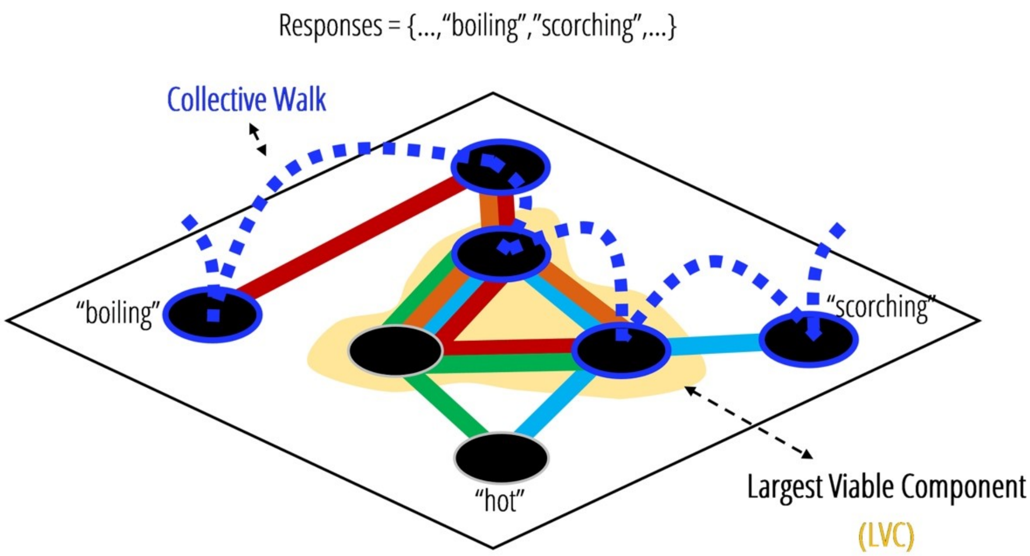
Table 1

List of multiplex network properties assessed from participants fluency list.

Name	Definition
Number of Responses	Number of responses in the list.
Coverage per Response	Average number of visited nodes in the multiplex shortest paths from one response to the next one.
Fraction of responses in LVC	Fraction of words in the list being part of the LVC.
Fraction of LVC Accesses	In the collective walk collating all shortest paths between response$_{i}$ and response_i+1 for all responses, check how many nodes in the LVC were visited over the total number of visited nodes.
Entropy of LVC Accesses	Always in the collective walk, put a 0 if a visited node is outside of the LVC, put 1 otherwise. On this binary list B, compute the Shannon entropy.
Entropy of LVC Coverage	Entropy of the collective walk w_{iN}, including nodes not in *l* but in the multiplex lexical network and being inside or outside the LVC.
Maximum Permanence in LVC	Maximum number of visited nodes in the collective walk w_{iN} being consecutively in the LVC.
Median Permanence in LVC	Median number of nodes in all the visits to the LVC during the collective walk w_{iN}.
Accesses to LVC from *hot*	Average number of visited nodes in the LVC in the multiplex shortest path between responses and *hot*.
Max Out	On the binary list B, check the length of the largest block containing all consecutive 0s.
Median Out	On the binary list B, compute the median of the lengths of all blocks containing consecutive 0s.
Distance from *hot* per response	For every response$_{i}$, measure the shortest path length between response$_{i}$ and the target category, e.g., *hot*. Sum the lengths and divide them by the number of responses.
Start in the LVC	Flag for the first response being in the LVC.
Fraction of typos	Percentage of incorrect spelling responses.
Norm 1	A normalized version of the maximum permanence in the LVC. Maximum number of visited nodes in the collective walk w_{iN} being consecutively in the LVC, divided by the number of responses.
Norm 2	This is a normalized version of the average permanence in the LVC. Median number of nodes in all the visits to the LVC during the collective walk w_{iN}, divided by the number of responses.

Figure 2.

Toy example of network features for a list of synonyms of hot, focusing on two consecutive responses.



Note – The LVC on a toy example for the multiplex lexical network is highlighted in yellow. A collective walk is the collection of shortest paths connecting any two consecutive responses. In here, the collective walk is highlighted as a dashed blue line. 3 nodes besides the responses are visited (hence, coverage is 3) but only 2 nodes are within the LVC. The shortest path starts from outside the LVC and thus corresponds to a binary sequence of (0,0,1,1,0), where 0 (1) indicates a node outside (within) the LVC, and on which entropy and permanence measures are computed (For more details, see **Table 2**).

2.2.4. Machine Learning Analysis

2.2.4.1. Data Splitting. The data for the machine learning model is split into two parts – training data and test data. The data splitting method we used was leave-one-out cross-validation (LOOCV; Alpaydin, 2020), in which all data points except one are used to train the model, and then the trained model is tested on the remaining data point. In this method, the split reoccurs multiple times so that each data point (i.e., the vector of measures relative to individual participants) is used as the test set exactly once. We use LOOCV to account for the contained sample size of this study while maximizing size and variability in the test set, thus reducing overfitting and producing more robust machine learning models (Alpaydin, 2020).

2.2.4.2. Models. To learn how to classify the data, and how to predict a score, there is a need for a learning model, that is based on a certain method or rule. In our case, we used a binary logistic regression approach for the classification model (Alpaydin, 2020). This method weighs the features' importance in predicting the score for each subject, and returns a binary prediction (low- or high- Openness). We used linear regression, which is similar to the binary logistic regression, but returns a continuous score and not a binary one. We chose linear regression for the prediction, as it allows us to identify the significance of each feature in predicting the Openness, and to easily compare goodness-of-fit between models, using Pearson's R. Importantly, to find the best regression models, both in prediction and in classification, we followed a stepwise regression model. For classification, we used a forward stepwise regression, where features are added to the regression model when significantly improving the model's area under the curve (AUC). For our prediction analysis, we used a backwards stepwise regression, where features are removed from the regression model when its removal significantly improves the model's *p*-value. Importantly, albeit using backwards stepwise regression as the default, the reason we chose to use forward stepwise regression for classification was that backwards regression converged with a minimal dataset of only 36 participants. This meant losing approximately 60% of the sample due to missing values in some of the chosen features. Therefore—to capitalize on the entire dataset and avoid overfitting based on a small sample—we chose to use forward regression instead, which converged leaving more participants in the analysis.

2.2.4. Statistical Analysis

The first step of the analysis was to calculate each of the multiplex measures for each participant, based on their list of *hot* synonyms. Next, we created two groups, consisting of participants with the lowest and highest 30% of Openness scores. Then, a comparison between group scores for each network measure was calculated using a Mann-Whitney U statistical test. In parallel, we conducted a Pearson's R test to measure the correlation between each of the multiplex measures and Openness scores. Finally, two computational analyses were conducted. First, all multiplex measures were used as the basis for a prediction model of Openness. Using the LOOCV approach and a backwards stepwise linear regression, a model was fitted to the data. Second, all multiplex measures were again used, this time as the basis for a classification model of Openness. A model was created using the LOOCV approach and a forward stepwise logistic regression.

2.2.5. Openness Facets Analysis

Using the model trained on Openness with the same training set, features, splitting method, and machine learning model, we attempted to predict the facets that the Openness to Experience trait includes. In the NEO-PI-3 questionnaire (McCrae et al., 2005) Openness score is built from Aesthetics, Actions, Ideas, Values, Fantasy and Feelings. In this analysis we tested the Openness trained model on these six facets. Such an analysis allows us to examine more specific relations between the mental lexicon and the different parts of Openness to Experience.

2.2.6. Model Specificity Analysis

After creating the model based on the Openness scores and finding the features that yield the best result, as a further test of model specificity for predicting Openness, we attempted to predict the other four Big-5 personality traits (Neuroticism, Extraversion, Agreeableness, and Conscientiousness) using the model trained on predicting Openness with the same training set, features, splitting method, and machine learning model. The only difference in this analysis was the test set, which included the additional four Big-5 personality traits. Such an analysis allows us to examine the specificity of our model and how uniquely it relates to Openness.

3. Results

3.1. Low- vs. High- Openness to Experience Group Analysis

We first computed for each participant their cognitive multiplex network properties representing their behavioral performance in the *hot* synonyms task. We then examined the difference between the low- and high- Openness groups across these cognitive multiplex network properties, via a non-parametric Mann-Whitney test.

This analysis revealed significant differences across the two groups in multiple cognitive multiplex network properties, most of which related to the LVC (**Table 2** and **Figure 3**). These properties included the Number of Responses, Coverage per Response, Fraction of Responses in LVC, Entropy of LVC Accesses, Maximum Permanence in LVC, Median Permanence in LVC, Maximum Out, Median Out, Distance from *hot* per Response, and Accesses to LVC from *hot*. Similar to higher creative individuals (Stella & Kenett, 2019), high Openness individuals tended to generate more synonyms to the word *hot*, synonyms that were less related to the LVC.

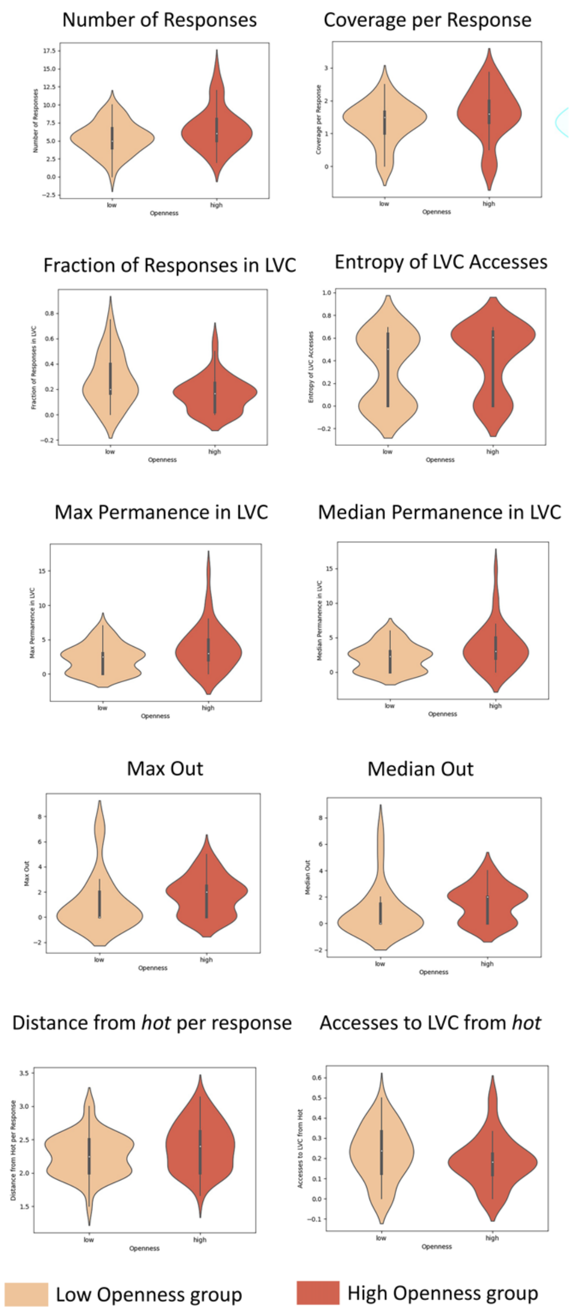
Table 2

Cognitive multiplex network variables that significantly differ between the low- and high- Openness to Experience groups.

Variable	N_{low}	N_{high}	Mann-Whitney U	*p*-value
Number of Responses	46	46	795.50	.019
Coverage per Response	45	45	790.50	.036
Fraction of responses in LVC	45	46	702.50	.004
Entropy of LVC Accesses	45	46	797.50	.027
Maximum Permanence in LVC	46	46	770.00	.011
Median Permanence in LVC	46	46	791.00	.017
Maximum Out	46	46	212.50	.020
Median Out	22	25	180.00	.013
Distance from h*ot* per Response	45	45	806.00	.045
Accesses to LVC from *hot*	45	45	788.00	.035

Figure 3.

Violin plots of difference in cognitive multiplex network properties across the low- and high- Openness to Experience groups.



Note - X-axis – the low- and high- Openness groups. Y-axis – the various multiplex network parameters.

3.2. Individual Differences Analysis

Next, we examined the relation between the various cognitive multiplex network properties and individual differences in Openness, via Pearson's correlation analysis. This analysis revealed several significant correlations between the various cognitive multiplex network properties and Openness (**Table 3** and **Figure 4**). These properties include: Number of responses, Coverage of response, Fraction of responses in LVC, Maximum permanence in LVC, Median permanence in LVC, and Distance from *hot* per response. These significant properties fully correspond with the low- and high- Openness group analysis. These results highlight the general significance of these properties in relation to Openness.

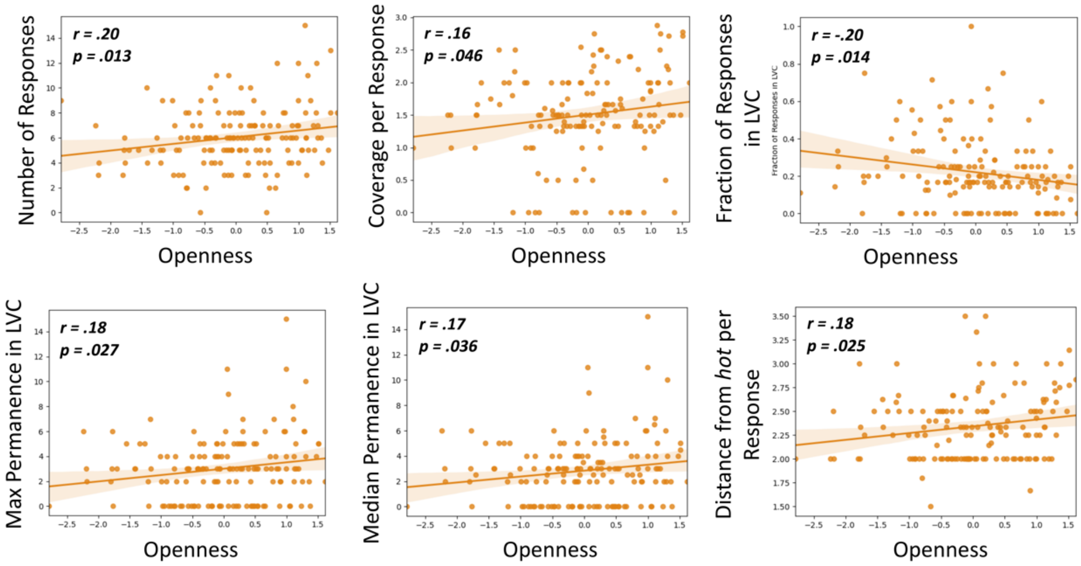
Table 3

Cognitive multiplex network variables that were significantly related to individual differences in Openness to Experience.

Variable	N	***r***	*p*-value
Number of Responses	154	.20	.013
Coverage per Response	148	.16	.046
Fraction of responses in LVC	152	-.20	.014
Maximum Permanence in LVC	154	.18	.027
Median Permanence in LVC	154	.17	.036
Distance from h*ot* per Response	148	.18	.025

Figure 4.

Scatter plots of the multiplex network properties across the sample.



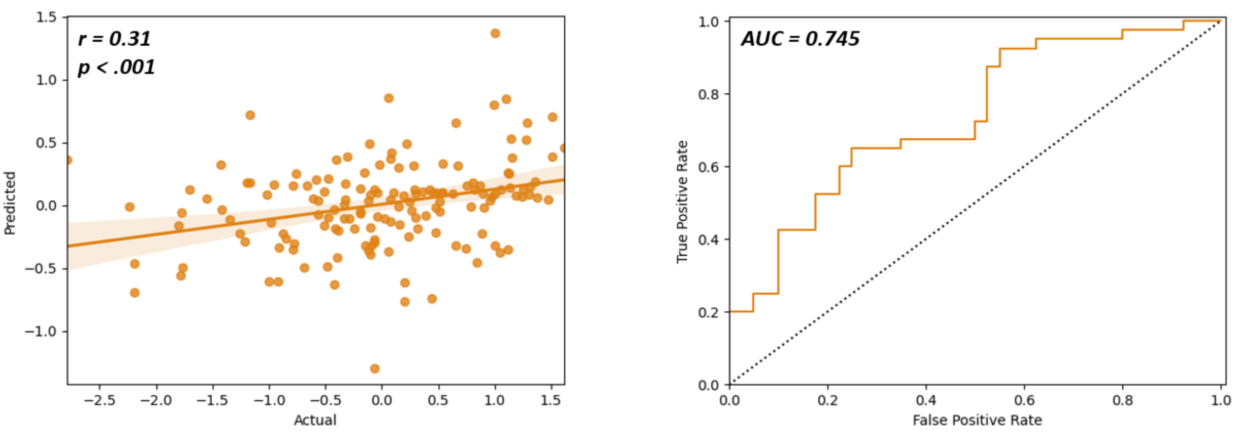
Note - Top row from left to right: Number of Responses; Coverage per Response; Fraction of Responses in LVC; Bottom row from left to right: Maximum Permanence in LVC; Median Permanence in LVC; Distance from *hot* per Response. X-axis – Openness to Experience scores, Y-axis - the various multiplex network parameters. Lighter orange background indicates confidence intervals.

3.3. Machine Learning Analysis

As a next step, we examine the performance of our machine learning model to predict and classify Openness (**Figure 5**). The prediction model, using a leave one out cross-validation splitting technique and a backwards stepwise linear regression, yielded a significant correlation between the predicted and actual score, $r(\mathbf{151}) = .31$, MSE = 0.74, $p < .001$, with the following features: Fraction of Responses in LVC, Entropy of LVC Accesses, Max Permanence in LVC, and Norm1. The classification model, using a leave one out cross-validation and a logistic regression classifier, yielded an AUC of .745 (N = 80), with the following features: Fraction of Responses in LVC, Entropy of LVC Accesses, Fraction of LVC Accesses, Median Permanence in LVC, Fraction of Incorrect Spellings, Accesses to LVC from Hot, and Distance from *hot* per Response.

Figure 5.

Machine learning model results.



Prediction

Classification

Note - Prediction: Correlation between predicted and actual Openness scores. Classification: Receiver operating characteristics (ROC) curve for the classification of low- and high- Openness scores.

3.4. Openness Facets Analysis

Next, we examine whether the success of our model predicting Openness is driven by any specific Openness facet that comprises the general Openness trait based on the NEO PI-3 Inventory (McCrae et al., 2005). We do so by testing the success of this trained model in capturing the different Openness facets – Actions, Aesthetics, Fantasy, Feelings, Ideas, and Values. Testing the prediction model on the facets of

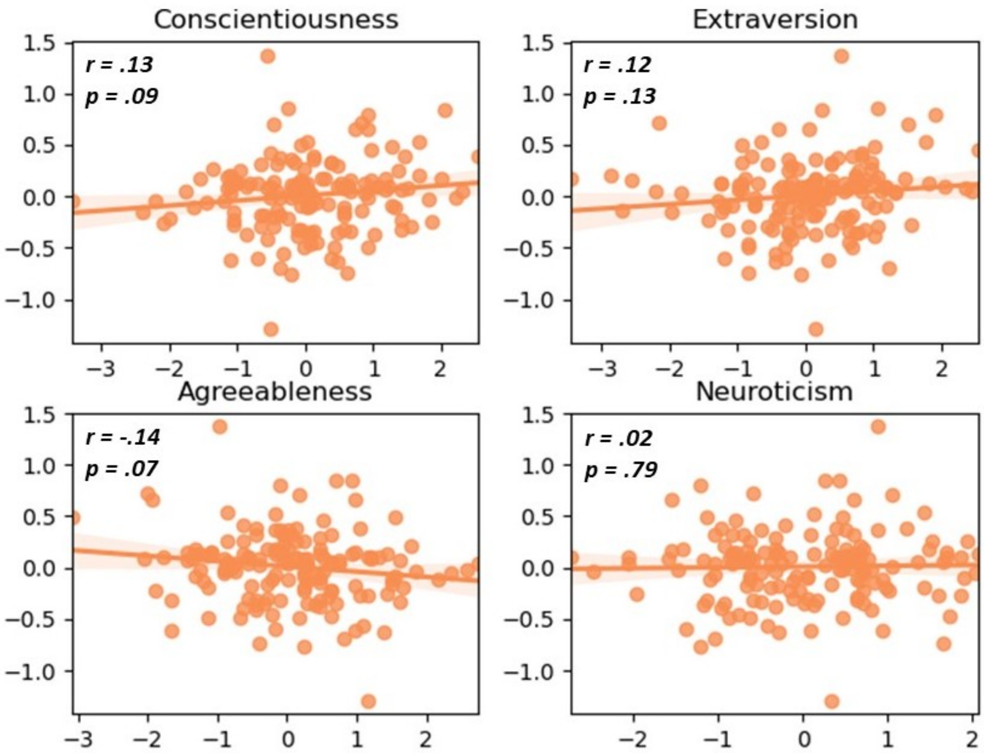
Openness to Experience score yielded the following results: Aesthetics, $r = .32$, $p < .001$, MSE = 0.9133; Actions, $r = .10$, $p = .243$, MSE = 1.07; Ideas, $r = .14$, $p = .110$, MSE = 1.05;Values, $r = .11$, $p = .184$, MSE = 1.06; Fantasy, $r = .01$, $p = .929$, MSE = 1.15; Feelings, $r = .24$, $p = .004$, MSE = 0.9495.

3.5. Specificity Analysis

Finally, we examined the specificity of our model for predicting Openness, and not any other personality trait. We do so by testing the success of this trained model in capturing the additional four personality traits from the Big-5 model. As expected, our prediction model—using leave one out cross-validation splitting and a backwards stepwise linear regression, and trained on the Openness data—yielded nonsignificant correlations between predicted scores and actual scores with the additional four personality traits (**Figure 6**): Conscientiousness, $r(151) = .13$, MSE = 1.01, $p = .09$; Extraversion, $r(151) = .12$, MSE = 1.05, $p = 0.13$; Agreeableness, $r(151) = -.14$, MSE = 1.22, $p = .07$; and Neuroticism, $r(151) = .02$, MSE = 1.07, $p = .79$. These results indicate that the model trained on Openness scores is unable to provide meaningful predictions for the additional four personality traits, providing evidence of its specificity in predicting Openness.

Figure 6.

Results of prediction of the Big 5 personality traits other than Openness, using the model trained to predict Openness.



***Note* -** X-axes – Predicted scores. Y-axes – Actual scores.

4. Discussion

In the current study, we examine how mental navigation through memory (Benigni et al., 2021)—operationalized via a semantic fluency task—predicts the personality trait of Openness to Experience. To this aim, we represent the mental lexicon as a cognitive multiplex network that consists of linguistic and conceptual information. We then demonstrate how this multiplex representation relates to differences between low- and high- Openness groups, relates to individual differences in Openness, can be used to construct a machine learning model that accurately predicts Openness, based on a model with high specificity.

Our work is based on recent applications of computational methods to study structure and processes in cognitive systems such as language and memory (Baronchelli et al., 2013; Borge-Holthoefer & Arenas, 2010; Günther et al., 2019; Hills & Kenett, 2022; Mandera et al., 2017; Siew et al., 2019). These advances have led to empirical investigation of the structure of memory as graphs, or networks, and the processes operating over them, such as mental navigation (Benigni et al., 2021; Todd & Hills, 2020). However, these studies largely treat different linguistic levels—such as semantics and phonology—separately, and only few studies have examined multidimensional cognitive systems, which represent more than one type of information. These studies deal with analyzing a cognitive multiplex network, which comprises of different layers, or networks, of information. Cognitive multiplex network research has demonstrated how such an approach can be uniquely used to study issues related to language, learning, development, creativity, and clinical research (Castro, 2022; Castro & Stella, 2019; Levy et al., 2021; Stella, 2019; Stella et al., 2017; Stella et al., 2018; Stella & Kenett, 2019).

Several cognitive multiplex network studies have highlighted the role of a core in the multiplex network that cuts across all of the layers: the largest viable cluster (Stella et al., 2018; Stella & Kenett, 2019). This core, the LVC, is composed of highly general, frequent, and conceptually concrete words which are considered to facilitate language comprehension and processing. Importantly, the LVC emerges from the multiplexity of the mental lexicon and cannot be identified in single-layer modelling approaches. Stella and Kenett (2019) have shown that higher creative individuals retrieve fewer words from the LVC and spend less time searching within it, in line with the idea that higher creative individuals search farther and more broadly through their memory (Kenett, forthcoming; Kenett & Faust, 2019).

Similar to Stella and Kenett (2019), we find several differences in our cognitive multiplex networks measures between the low- and high- Openness groups. Specifically, we found that the high Openness group generated more synonyms to *hot*, and had a smaller fraction of their responses inside the LVC, in addition to generating synonyms that were farther on the network from *hot* than the low Openness group's responses. These findings overlap with Stella and Kenett (2019) who showed that high creative individuals generate more responses than low creative individuals, and a smaller fraction of their responses are inside the LVC. The strong relation between Openness and creativity, and the consistency across our study and that of Stella and Kenett (2019), highlights and generalizes the use of cognitive multiplex networks to study complex cognitive behavior.

Moving beyond group effects, we examined the relation of the cognitive multiplex network parameters—based on performance in the *hot* synonym task—and individual differences in Openness. Our analysis revealed multiple significant relations that further highlight the role of the LVC in complex behavior. Similar to the findings of Stella and Kenett (2019), number of responses, fraction of responses in the LVC, coverage of responses, and entropy of responses were significantly related to Openness. Thus, our current study replicates the findings of Stella and Kenett (2019), further highlighting the role of the LVC in complex behavior. In addition, our study generalizes these findings by a different fluency task (synonyms to *hot* in our study vs. animal fluency in the Stella and Kenett [2019] study) and a different predicted complex behavior (Openness to Experience in our study vs. creativity in the Stella and Kenett [2019] study). Moreover, our findings expand our previous work (Stella & Kenett, 2019), by moving from between-group comparisons to demonstrating how this analysis can capture individual differences.

Notably, our supervised machine learning model results show a small to medium correlation between predicted and actual Openness scores of $r = .31$, and a .75 AUC score for classifying whether a participant is low or high in Openness. This classification ability is very close to the benchmark for clinical use, which is .80 (Jones & Athanasiou, 2005). Thus, our results highlight the ability to quantitatively predict Openness simply based on how participants retrieve synonyms to common words. Examining the specificity of our model, we find that it is specific to Openness as it is unsuccessful in predicting the other Big-5 personality traits. This finding strengthens the validity of our model and further highlights the close relation between the personality trait of Openness to Experience and cognition (Zillig et al., 2002).

To delve deeper into the possible theoretical implications of the results of our model, we conducted an exploratory post-hoc Openness facets analysis. Specifically, we examined the success of our model in predicting the different Openness facets from the NEO PPI-3 inventory (McCrae et al., 2005): *Action*, *Aesthetics*, *Fantasy*, *Feelings*, *Ideas*, and *Values*. We found that our model significantly predicted only two of these facets, namely *Aesthetics* and *Feelings*. A previous study by McCrae (1993) found that the *Aesthetics* facet of Openness was significantly—albeit weakly—correlated with different subscales of an IQ test of vocabulary, similarities, and object assembly. Moutafi et al. found that the *Actions* and *Ideas* facets were significantly related to measures of fluid intelligence (Moutafi et al., 2006). Finally, our results correspond to the findings of Christensen et al., who found that higher Openness individuals—characterized based on *Aesthetic*, *Fantasy*, and *Openness to Emotions* facets—had a richer semantic memory network structure compared to low Openness individuals (Christensen, Kenett, et al., 2018). Finally, Altaras-Dimitrijević (2012) has shown that gifted people can be characterized by a model that includes seven personality facets, including *Aesthetics* and *Fantasy*.

As such, some evidence suggests that different Openness facets relate to different aspects of complex cognitive capacities. Overall, *Aesthetics* and *Feelings* are similar to each other, grounded more in perceptual experiences of sensory (Aesthetics) and affective (Feelings) states linked to aesthetic experiences (McCrae, 1993; McCrae & Costa, 1997). *Fantasy*, while closely related to *Aesthetics* and *Feelings*, tends to be about daydreaming and lack of attentional focus. While *Fantasy* has been related to cognition (Altaras-Dimitrijević, 2012; Christensen, Kenett, et al., 2018), our model was not successful in predicting it. This may be due to the semantic fluency task being a focused goal-directed task that requires attentional focus (Ovando-Tellez, Benedek, et al., 2022), which is thus likely less related to the *Fantasy* facet.

Thus, much more nuanced research is needed to elucidate the relations of specific personality facets to cognition (Seeboth & Mõttus, 2018; Sindermann et al., 2021; Soutter & Mõttus, 2021). Furthermore, Openness has been related to several different, overlapping, facets, based on the questionnaire used for its assessment. Thus, future research is needed based on our findings examining different Openness facets (Christensen, Cotter, et al., 2018).

Our results provide further support that the mental lexicon can be modelled using multiplex networks, and that the characteristics of such a model are linked to complex cognitive traits, such as language, development, and creativity, in typical and clinical populations (Levy et al., 2021; Stella, 2019, 2020; Stella et al., 2017; Stella & Kenett, 2019). Furthermore, the link outlined here between the LVC and Openness indicates that multiplexity is an important feature of mental representations of linguistic knowledge. Since the LVC emerges from the multiplex interplay of semantic and phonological associations, our work implies that cognitive multiplex networks represent a natural and convenient framework for exploring cognitive traits, through quantitative and reproducible measurements, free from the constraints of subjective evaluations (e.g., self-report scales).

More generally, our findings provide further support for the link of personality traits (Openness to Experience) and cognitive systems (the mental lexicon). Traditionally, cognition and personality have been investigated separately, but a large body of work has linked creativity with Openness, a personality trait that has been referred to as the “creativity trait” (Christensen, Cotter, et al., 2018; Christensen, Kenett, et al., 2018; Kaufman et al., 2016; Oleynick et al., 2017). Christensen, Kenett et al. (2018) have recently shown a relation between semantic memory structure and Openness. The authors show how people high on Openness had a more flexible, richly connected semantic memory network (Christensen, Kenett, et al., 2018). Our study provides further evidence supporting the relation between Openness and the mental lexicon, pushing these two domains closer together. Thus, increased theoretical focus on the role of the mental lexicon in personality—and especially in Openness to Experience—is needed. Recent studies have begun theoretically studying personality as a complex dynamic system using network science methodology (Beck & Jackson, 2021). Based on our current study, future cognitive multiplex network research should expand to incorporate a “personality” layer to more directly study how cognition and personality impact each other and interact together to realize complex human behavior. Given the strong coupling between Openness and creativity, future research should conduct a similar cognitive multiplex network approach on a large sample of participants that includes assessment for both Openness and creativity, as well as consider various semantic fluency categories to identify the most valid and reliable way to build a more accurate prediction and classification model of high-level cognition.

It is important to note that the dataset for this research was smaller than optimal for machine learning, which might lead to biased results (Vabalas et al., 2019). Moreover, the prediction model's MSE was quite large, which means that although the relation between predicted and actual Openness scores is noticeable, prediction itself is only modestly accurate. To address these issues, we intend to replicate and extend our research in the future, using a considerably larger dataset (this should also assist in minimizing the MSE).

We based our definition on the idea that synonyms can have similar meanings, but they may not have the exact same meaning. It's worth noting that the classification of words as synonyms can be somewhat subjective and can vary depending on the context. In addition, words can have multiple meanings: some may be synonyms in one context but not in another. Importantly, there are also different types of synonyms, including exact synonyms, perfect synonyms, near-synonyms or quasi-synonyms, and relative synonyms. In our study, we considered near-synonyms or quasi-synonyms. Future studies should replicate and extend our current findings by applying more rigorously defined fluency tasks, such as a recent method suggested by Ovando-Tellez et al. (2022). Furthermore, it is important to note that most cognitive network research analyzing semantic fluency responses solely focuses on the animal category. This is due to this category being well-structured, taxonomic, and universal (Ardila et al., 2006; Christensen & Kenett, 2021a; Zemla & Austerweil, 2018). As these semantic fluency categories becomes less well-defined—such as synonyms for *hot*—the application of computational methods such as those applied here become more challenging. Further methodological research is needed to examine the reliability and generality of our approach on additional semantic fluency categories.

In summary, the results of this research show that it is possible to predict and classify personality scores using multiplex networks and a very short task, even when data size is limited. Therefore, our work adds another support to the assumption that these tools can be used to predict complex cognitive traits. To further investigate this direction of research, and to increase certainty regarding this method, it is important to replicate this study and its methods with a larger dataset, and aim at predicting more complex cognitive traits. Nevertheless, our study further demonstrates the ability to predict complex behavior from simple, behavioral tasks such as semantic fluency. These findings push personality and cognition closer together, and provide initial evidence for the ability to develop automatic, objective scoring of Openness. Such a quantitative direction has largely advanced creativity research over the past decade (Beaty & Johnson, 2021; Beaty et al., 2018; Dumas et al., 2021; Ovando-Tellez, Kenett, et al., 2022). Given that Openness and Creativity are so closely related, such quantitative methods should be further applied to advance Openness-and personality more generally—research.

Acknowledgements

We thank Dr. Alexander Christensen for his assistance in data analysis.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Random walks on semantic networks can resemble optimal foraging. *Psychological Review*, *122*(3), 558-569. https://doi.org/10.1037/a0038693

Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.

Altaras-Dimitrijević, A. (2012). A faceted eye on intellectual giftedness: Examining the personality of gifted students using FFM domains and facets. *Psihologija*, *45*(3), 231-256.

Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of semantic verbal fluency (ANIMALS). *International Journal of Psychology*, *41*(5), 324-332. https://doi.org/10.1080/00207590500345542

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.

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in cognitive sciences*, *17*(7), 348-360. https://doi.org/10.1016/j.tics.2013.04.010

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757-780. https://doi.org/10.3758/s13428-020-01453-w

Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapil, T. R., Kane, M. J., & Silvia, P. J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences*, *115*(5), 1087-1092. https://doi.org/10.1073/pnas.1713532115

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition*, *42*(7), 1-12. https://doi.org/10.3758/s13421-014-0428-8

Beck, E. D., & Jackson, J. J. (2021). Network approaches to representing and understanding personality dynamics. In D. Wood, S. J. Read, P. D. Harms, & A. Slaughter (Eds.), *Measuring and Modeling Persons and Situations* (pp. 465-497). Academic Press.

Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., & Neubauer, A. C. (2017). How semantic memory structure and intelligence contribute to creative thought: a network science approach. *Thinking & Reasoning*, *23*(2), 158-183. https://doi.org/10.1080/13546783.2016.1278034

Benigni, B., Dallabona, M., Bravi, E., Merler, S., & De Domenico, M. (2021). Navigating concepts in the human mind unravels the latent geometry of Its semantic space. *Complexity*, *2021*, 6398407. https://doi.org/10.1155/2021/6398407

Borge-Holthoefer, J., & Arenas, A. (2010). Semantic networks: Structure and dynamics. *Entropy*, *12*(5), 1264-1302. https://doi.org/10.3390/e12051264

Borodkin, K., Kenett, Y. N., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition*, *156*, 60-70. https://doi.org/10.1016/j.cognition.2016.07.014

Castro, N. (2022). Methodological Considerations for Incorporating Clinical Data Into a Network Model of Retrieval Failures [https://doi.org/10.1111/tops.12531]. *Topics in Cognitive Science*, *14*(1), 111-126. https://doi.org/https://doi.org/10.1111/tops.12531

Castro, N., & Siew, C. S. Q. (2020). Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proceedings of the Royal Society A*, *476*(2238), 20190825.

Castro, N., & Stella, M. (2019). The multiplex structure of the mental lexicon influences picture naming in people with aphasia. *Journal of Complex Networks*, *7*(6), 913-931.

Christensen, A. P., Cotter, K. N., & Silvia, P. J. (2018). Reopening openness to experience: A network analysis of four openness to experience inventories. *Journal of Personality Assessment*, 1-15.

Christensen, A. P., & Kenett, Y. N. (2021a). Semantic network analysis (SemNA): A tutorial on preprocessing, estimating, and analyzing semantic networks. *Psychological Methods*.

Christensen, A. P., & Kenett, Y. N. (2021b). Semantic network analysis (SemNA): A tutorial on preprocessing, estimating, and analyzing semantic networks. *Psychological Methods*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/met0000463

Christensen, A. P., Kenett, Y. N., Cotter, K. N., Beaty, R. E., & Silvia, P. J. (2018). Remotely close associations: Openness to experience and semantic memory structure. *European Journal of Personality*, *32*(4), 480-492. https://doi.org/doi:10.1002/per.2157

Dai, W., Brisimi, T. S., Adams, W. G., Mela, T., Saligrama, V., & Paschalidis, I. C. (2015). Prediction of hospitalization due to heart diseases by supervised learning methods. *International journal of medical informatics*, *84*(3), 189-197.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987-1006.

DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality*, *46*(1), 63-78.

Dumas, D., Organisciak, P., & Doherty, M. (2021). Measuring originality with human raters and text-mining models: A psychometric comparison of methods. *Psychology of Aesthetics, Creativity and the Arts*, *15*(4), 645-663. https://doi.org/https://doi.org/10.1037/aca0000319

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, *2*(4), 290-309.

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006-1033.

He, L., Kenett, Y. N., Zhuang, K., Liu, C., Zeng, R., Yan, T., Huo, T., & Qiu, J. (2021). The relation between semantic memory structure, associative abilities, and verbal and figural creativity. *Thinking & Reasoning*, *27*(2), 268-293. https://doi.org/10.1080/13546783.2020.1819415

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431-440. https://doi.org/10.1037/a0027373

Hills, T. T., & Kenett, Y. N. (2022). Is the mind a network? Maps, vehicles, and skyhooks in cognitive network science. *Topics in Cognitive Science*, *14*(1), 189-208.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, *19*(1), 46-54. https://doi.org/http://dx.doi.org/10.1016/j.tics.2014.10.004

Johnson, J. A. (1994). Clarification of factor five with the help of the AB5C model. *European Journal of Personality*, *8*(4), 311-334.

Jones, C. M., & Athanasiou, T. (2005). Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. *The Annals of thoracic surgery*, *79*(1), 16-20.

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment*, *82*(3), 291-305.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321-340.

Kaufman, S. B., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., & DeYoung, C. G. (2016). Openness to experience and intellect differentially predict creative achievement in the arts and sciences. *Journal of Personality*, *84*(2), 248-258.

Kenett, Y. N. (forthcoming). Creatively searching through semantic memory structure: A short integrative revie. In L. J. Ball & F. Vallée-Tourangea (Eds.), *Routledge International Handbook of Creative Cognitio*.

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in human neuroscience*, *8*(407), 1-16. https://doi.org/10.3389/fnhum.2014.00407

Kenett, Y. N., Beaty, R. E., Silvia, P. J., Anaki, D., & Faust, M. (2016). Structure and flexibility: Investigating the relation between the structure of the mental lexicon, fluid intelligence, and creative achievement. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(4), 377-388. https://doi.org/10.1037/aca0000056

Kenett, Y. N., & Faust, M. (2019). A semantic network cartography of the creative mind. *Trends in Cognitive Sciences*, *23*(4), 271-274. https://doi.org/10.1016/j.tics.2019.01.007

Kenett, Y. N., Levy, O., Kenett, D. Y., Stanley, H. E., Faust, M., & Havlin, S. (2018). Flexibility of thought in high creative individuals represented by percolation analysis. *Proceedings of the National Academy of Sciences*, *115*(5), 867-872. https://doi.org/10.1073/pnas.1717362115

Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletun & Review*, *28*(1), 40-80. https://doi.org/10.3758/s13423-020-01792-x

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate behavioral research*, *39*(2), 329-358.

Levy, O., Kenett, Y. N., Oxenberg, O., Castro, N., De Deyne, S., Vitevitch, M. S., & Havlin, S. (2021). Unveiling the nature of interaction between semantics and phonology in lexical access based on multilayer networks. *Scientific Reports*, *11*(1), 14479. https://doi.org/10.1038/s41598-021-93925-y

Lin, Y., Zhou, K., & Li, J. (2018). Prediction of slope stability using four supervised learning methods. *IEEE Access*, *6*, 31169-31179.

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57-78. https://doi.org/http://dx.doi.org/10.1016/j.jml.2016.04.001

McCrae, R. R. (1993). Openness to experience as a basic dimension of personality. *Imagination, Cognition and Personality*, *13*(1), 39-55.

McCrae, R. R., & Costa, P. T. (1997). Conceptions and correlates of openness to experience. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 825-847). Academic Press.

McCrae, R. R., Costa, P. T., & Martin, T. A. (2005). The NEO–PI–3: A more readable revised NEO personality inventory. *Journal of Personality Assessment*, *84*(3), 261-270.

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, *38*(11), 39-41.

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences*, *16*(1), 31-42.

Oleynick, V. C., DeYoung, C. G., Hyde, E., Kaufman, S. B., Beaty, R. E., & Silvia, P. J. (2017). Openness/Intellect: The core of the creative personality. In G. J. Fiest, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *Cambridge handbook of creativity and personality research* (pp. 9-27). Cambridge University Press.

Ovando-Tellez, M., Benedek, M., Kenett, Y. N., Hills, T., Bouanane, S., Bernard, M., Belo, J., Bieth, T., & Volle, E. (2022). An investigation of the cognitive and neural correlates of semantic memory search related to creative ability. *Communications Biology*, *5*(1), 1-16.

Ovando-Tellez, M., Kenett, Y. N., Benedek, M., Bernard, M., Belo, J., Beranger, B., Bieth, T., & Volle, E. (2022). Brain connectivity–based prediction of real-life creativity is mediated by semantic memory structure. *Science advances*, *8*(5), eabl4294.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36.

Seeboth, A., & Mõttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *European Journal of Personality*, *32*(3), 186-201.

Siew, C. S. Q., Wulff, D. U., Beckage, N. M., & Kenett, Y. N. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity*, *2019*, 2108423. https://doi.org/https://doi.org/10.1155/2019/2108423

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence*, *41*(5), 328-340. https://doi.org/10.1016/j.intell.2013.05.004

Silvia, P. J., & Sanders, C. E. (2010). Why are smart people curious? Fluid intelligence, openness to experience, and interest. *Learning and Individual Differences*, *20*(3), 242-245.

Sindermann, C., Mõttus, R., Rozgonjuk, D., & Montag, C. (2021). Predicting current voting intentions by Big Five personality domains, facets, and nuances–a random forest analysis approach in a German sample. *Personality Science*, *2*, 1-21.

Soutter, A. R. B., & Mõttus, R. (2021). Big five facets' associations with pro-environmental attitudes and behaviors. *Journal of Personality*, *89*(2), 203-215.

Stella, M. (2019). Modelling early word acquisition through multiplex lexical networks and machine learning. *Big Data and Cognitive Computing*, *3*(1), 10.

Stella, M. (2020). Multiplex networks quantify robustness of the mental lexicon to catastrophic concept failures, aphasic degradation and ageing. *Physica A: Statistical Mechanics and its Applications*, *554*, 124382. https://doi.org/https://doi.org/10.1016/j.physa.2020.124382

Stella, M., Beckage, N. M., & Brede, M. (2017). Multiplex lexical networks reveal patterns in early word acquisition in children [Article]. *Scientific reports*, *7*, 46730. https://doi.org/10.1038/srep46730

Stella, M., Beckage, N. M., Brede, M., & De Domenico, M. (2018). Multiplex model of mental lexicon reveals explosive learning in humans. *Scientific reports*, *8*(1), 2259. https://doi.org/10.1038/s41598-018-20730-5

Stella, M., & Kenett, Y. N. (2019). Viability in multiplex lexical networks and machine learning characterizes human creativity. *Big Data and Cognitive Computing*, *3*(3), 45.

Todd, P. M., & Hills, T. T. (2020). Foraging in mind. *Current Directions in Psychological Science*, *29*(3), 309-315.

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLoS ONE*, *14*(11), e0224365.

von Stumm, S. (2018). Better open than intellectual: The benefits of investment personality traits for learning. *Personality and Social Psychology Bulletin*, *44*(4), 562-573.

Zemla, J. C., & Austerweil, J. L. (2018). Estimating semantic networks of groups and individuals from fluency data.

Zillig, L. M. P., Hemenover, S. H., & Dienstbier, R. A. (2002). What do we assess when we assess a Big 5 trait? A content analysis of the affective, behavioral, and cognitive processes represented in Big 5 personality inventories. *Personality and Social Psychology Bulletin*, *28*(6), 847-858.