Received: 3 April 2023 | Revised: 11 October 2023 | Accepted: 13 December 2023

DOI: 10.1111/jopy.12915

ORIGINAL ARTICLE

WILEY

Surfing the OCEAN: The machine learning psycholexical approach 2.0 to detect personality traits in texts

Federico Giannini[1] | **Marco Marelli**[2] | **Fabio Stella**[1] | **Dario Monzani**[3] | **Luca Pancani**[2]

[1]Department of Informatics, Systems and Communication, University of Milan-Bicocca, Milan, Italy

[2]Department of Psychology, University of Milan-Bicocca, Milan, Italy

[3]Department of Psychology, Educational Science and Human Movement, University of Palermo, Palermo, Italy

Correspondence
Luca Pancani, Department of Psychology, University of Milan-Bicocca, Milan, Italy.
Email: luca.pancani@unimib.it

Abstract

Objective: We aimed to develop a machine learning model to infer OCEAN traits from text.

Background: The psycholexical approach allows retrieving information about personality traits from human language. However, it has rarely been applied because of methodological and practical issues that current computational advancements could overcome.

Method: Classical taxonomies and a large Yelp corpus were leveraged to learn an embedding for each personality trait. These embeddings were used to train a feedforward neural network for predicting trait values. Their generalization performances have been evaluated through two external validation studies involving experts ($N=11$) and laypeople ($N=100$) in a discrimination task about the best markers of each trait and polarity.

Results: Intrinsic validation of the model yielded excellent results, with R^2 values greater than 0.78. The validation studies showed a high proportion of matches between participants' choices and model predictions, confirming its efficacy in identifying new terms related to the OCEAN traits. The best performance was observed for agreeableness and extraversion, especially for their positive polarities. The model was less efficient in identifying the negative polarity of openness and conscientiousness.

Conclusions: This innovative methodology can be considered a "psycholexical approach 2.0," contributing to research in personality and its practical applications in many fields.

KEYWORDS

Big Five, natural language processing, psycholexical approach, word embedding

1 | THE PSYCHOLEXICAL APPROACH TO PERSONALITY

The psycholexical approach is a methodology that allows the identification of psychological constructs from the human language. The basic assumption of this approach was brilliantly stated by Goldberg (1981): "Those individual differences that are of most significance in the daily transactions of persons with each other will eventually become encoded into their language. The

Dario Monzani and Luca Pancani contributed equally to this work.

more important is such a difference, the more people will notice it and wish to talk of it, with the results that eventually they will invent a word for it" (pp. 141–142). Indeed, language can be seen as a repository of concepts, ideas, and values that help people in structuring reality and making sense of their everyday experiences (Berger & Luckmann, 1966). The traditional psycholexical approach has focused on identifying the terms that could convey information about personality characteristics. This approach has been mainly applied to develop self-assessment scales to evaluate people's personality characteristics (e.g., Caprara et al., 1993; Cattell, 1943a, 1943b; Costa & McCrae, 1992). This approach eventually led to the development of methodologies to infer the personality of individuals by analyzing texts produced by them (e.g., Moreno et al., 2021).

Recent developments of the psycholexical approach were applied to infer personality characteristics in spontaneous texts describing people and nonhuman objects (Gosling & John, 1999; Tanasescu et al., 2013). Although these early attempts showed how human personality can be meaningfully applied to nonhuman targets, they were still limited to analyzing a set of specific words (i.e., markers) known as conveying personality information.

The present work fits into this last perspective by aiming to develop a machine learning model to infer personality traits of nonhuman targets (in our case, venues) by analyzing how they are spontaneously described in natural language. Specifically, we propose a psycholexical approach 2.0 that overcomes prior attempts by (1) getting personality-relevant information from all the words (vs. adjectives, as done in the past) by applying recent developments in the Natural Language Processing field and (2) using enriched embeddings that allow enhancing personality-related information encoded in such distributed representations, to optimize their applicability for research in personality psychology. Indeed, such an innovative approach might widely improve how we measure personality by both introducing a method that can be easily applied to spontaneously produced texts and overcoming the limitations of self-report questionnaires and other automatic methodologies, similar to our approach but based on few known terms.

2 | OVERVIEW OF THE PSYCHOLEXICAL APPROACH

In 1884, the famous polymath scientist Francis Galton was presumably the first who realized that language could convey information about people's personality. In his pioneering work "Measurement of character," Galton scanned the dictionary and identified 1000 words expressing people's character. Although this work represents only a first and unsystematic attempt, Galton's methodology laid the foundation of the psycholexical approach to studying personality. Indeed, during the last century, many efforts have been made to pinpoint representative terms of personality characteristics. A more organized theoretical and empirical contribution to the psycholexical approach can be attributed to Klages (1926) and Baumgarten (1933), who applied this method to the German language. A few years later, Allport and Odbert (1936) built a catalog of almost 18,000 personality-relevant words, identifying terms that can "distinguish the behavior of one human being from that of another" (Allport & Odbert, 1936, p. 24). The terms were classified into four categories: (1) neutral terms designating possible personal traits (e.g., active, introverted, conscientious), (2) terms primarily descriptive of temporary moods or activities (e.g., afraid, blue, liberal), (3) terms conveying social and character judgments of personal conduct (e.g., effective, irresistible, worthy), and (4) metaphorical and doubtful terms, which consisted of a miscellaneous category of words describing human beings (e.g., athletic, cannibal, and drinking).

Starting from this list, Cattell (1943a, 1943b) adopted an innovative methodology to identify the major dimensions of personality. Specifically, the author focused only on Allport and Odbert's first category, as it included terms that most closely resemble the definition of personality as an enduring and stable pattern of characteristics. Through a set of subsequent semantic and empirical reductions, Cattell (1957) developed a 16-factor personality model. Despite Cattell's contribution having been widely debated, it represented a crucial stimulus for other personality researchers. Among them, Tupes and Christal (1961) conducted correlational analyses on other-reported ratings deriving from eight different samples, identifying "five relatively strong and recurrent factors" (p. 14) of personality: surgency, agreeableness, dependability, emotional stability, and culture. Beyond sharing many similarities with some of Cattell's factors, these five dimensions can be considered the precursors of the Big Five personality traits. A further crucial contribution can be attributed to Norman (1967). Based on previous research and collecting an enormous set of ratings, Norman (1967) developed an exhaustive taxonomy (about 2800 terms) related to personality and proposed the first version of the contemporary Big Five model. Specifically, he identified the best markers for each of the poles of the five factors. Goldberg (1980, 1981, 1982) continued and refined Norman's work by developing a taxonomy of 1710 terms to be used in empirical research as items to measure the Big Five traits.

Along with these major contributions, the history of the psycholexical approach is full of attempts to find the best taxonomy to account for personality traits. Accordingly, this approach was applied to several languages (e.g., Caprara & Perugini, 1994; De Raad, 1992; Isaka, 1990; Ostendorf, 1990), developing different taxonomies that, beyond marginal cultural differences, have generally led to the identification of Big Five traits of personality: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (aka OCEAN). The strong consistency among the results of several studies represents significant evidence for the validity of the Big Five model and confirms the basic assumption of the psycholexical approach: The investigation of personality can be performed through the exploration of language. Most of the taxonomies developed so far consisted mainly of adjectives, as lexical terms that encode and describe qualities of persons. Thus, several scales aimed at measuring the Big Five consisted of adjective checklists (García et al., 2004; Goldberg, 1990, 1992; Saucier, 1994a; Saucier & Goldberg, 1996) or short statements that include several adjectives (e.g., the BFQ by Caprara et al., 1993; the NEO-PI-R by Costa & McCrae, 1992; the BFI by John et al., 1991).

In the last decades, the traditional psycholexical approach has been mostly abandoned. Probably, once the transcultural validity of the Big Five model was demonstrated and the best markers of each factor were identified, the scientific interest in this field has been exhausted. However, the psycholexical approach could go beyond these goals. For instance, detecting the Big Five markers in spontaneous descriptions of people and nonhuman objects could allow inferring these targets' personality. Indeed, terms used to describe people's personality can also be used to describe other entities, such as animals or venues. For instance, a review (Gosling & John, 1999) showed a large overlap in the use of the Big Five traits to describe humans and other animal species. An empirical attempt to describe the personality of inanimate objects was conducted by Tanasescu et al. (2013), who explored reviews of venues counting the Big Five markers (taken from Saucier & Goldberg, 1996) to obtain a score for each venue in each trait automatically. The authors matched these scores with ratings from participants, obtaining encouraging correspondence between automatic and human evaluations. Although they received little attention, these pioneering works laid the foundations of a new method for extracting information about personality from texts regarding nonhuman entities. We argue that current developments in computer science and linguistics may be exploited to improve this approach, breaking new grounds in the study of personality.

3 | VECTOR-SPACE MODELS OF LINGUISTIC DATA

Nowadays, methods from computer science, and in particular the Natural Language Processing field, offer a unique opportunity for a resurgence of the psycholexical approach and a new quantitative, automatic, data-driven investigation of the relationship between personality and language. Indeed, in the past few years, substantial advancements were brought to different psychology domains by applying vector-space models trained on large samples of language usage. Such models, often labeled as distributional semantic models, have been particularly impactful in cognitive science, with extremely influential proposals like Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), Hyperspace Analogue to Language (HAL Burgess, 1998), and the "bound encoding of the aggregate language environment" model (BEAGLE; Jones & Mewhort, 2007). These systems are funded on the distributional hypothesis, stating that word meanings can be approximated (or even learned) by an analysis of the contexts in which a word appears (Harris, 1954; Lenci, 2018). Using computational systems able to capture such contextual distributions, this approach ends up representing word meaning as vectors in a high-dimensional space, or *word embeddings*: The more two words appear in similar contexts, the more their corresponding vectors will "live" close to each other in the defined space, the more their corresponding meaning will be perceived as related by human participants.

Vector-space models are particularly appealing from a cognitive perspective for both empirical and theoretical reasons (Günther et al., 2019). First, independent predictions obtained from these models are shown to align well with behavioral and neural data from a number of different psychological tasks, ranging from semantic priming (Jones et al., 2006), to explicit intuitions (Landauer & Dumais, 1997), to sentence comprehension (Frank & Willems, 2017). Second, these systems have their foundations in psychologically plausible learning mechanisms (Mandera et al., 2017), such as those described by the Rescorla–Wagner equations (Rescorla & Wagner, 1972), and are believed to well-capture the passage from episodic to semantic memory via dimensionality reduction (Landauer & Dumais, 1997). These premises have granted the widespread application of such an approach in cognitive science (see Günther et al., 2019, for a review and a thorough discussion of their validity as cognitive models).

Importantly, the impact of vector-space models goes beyond the obvious domains of investigation such as the psychology of language and semantic memory. Several studies have shown how predictions from vector-space models can help shading lights on open issues in domains such as grounded cognition (Louwerse, 2007), numerical processing (Rinaldi & Marelli, 2020), episodic memory (Gatti et al., 2022), conceptual combination (Marelli et al., 2017), emotion representations (Rotaru & Vigliocco, 2020), social stereotypes (Caliskan et al., 2017), and human judgments (Bhatia et al., 2019). Personality psychology is also a straightforward domain of application. Consider the observations of Goldberg (1981) and Berger and Luckmann (1966) reported at the beginning of the paper: If language acts as a repository for concepts, ideas, and, more generally, life experiences (including interactions with other people), there must be a way to induce such representations by moving from language statistics, in a manner that is analogous to what vector-space models do. Indeed, similar arguments have been advanced, with respect to such systems, by Louwerse: In the context of his symbol-interdependence hypothesis, he observed how we speak of things in the world and our experience with them; as a result, language ends up encoding grounded information, that can be extracted by the application of proper computational methods to linguistic data (Louwerse, 2011).

Applying vector-space modeling to personality psychology also offers an important methodological advancement, allowing researchers to rely on semantic representations rather than purely lexical ones. In fact, most of the classical studies in the domain have focused on lists of adjectives (as presented above), limiting research to a rather small portion of the whole lexicon of a language. Word embeddings avoid this substantial limitation associated with the grammatical class of the items: In principle, personality-related information can be induced for any word for which an embedding is available, on the ground of its distributional similarity with "pivot" terms for which the personality pattern is known (e.g., the very adjectives considered in classical studies).

Evidence in this respect has been provided, showing that it is possible to induce personality information from textual data. For example, a systematic literature review (Ahmad et al., 2020) identified 30 scientific papers between 2007 and 2019 that used a variety of methods (i.e., deep learning, supervised and unsupervised machine learning) to detect personality characteristics from texts. A recent meta-analysis by Moreno et al., including 23 different studies, indicates effect sizes ranging from $r=0.26$ to $r=0.30$, depending on the personality trait under consideration. Furthermore, a work by Cutler and Condon (2022) exploits the recent advances in Natural Language Processing technologies to represent 435 personality descriptors and then compare these vectorial representations to a principal component analysis carried out on survey-based ratings.

Here, we propose a new approach in which, prior to the actual language-to-personality mapping, personality-related information is highlighted in the textual data through an enrichment process. We thus try to build a specific vector space for each personality trait and use it to infer word' loadings on that trait.

4 | THE PRESENT RESEARCH

The rationale of the present research is that the OCEAN model allows to describe personality traits of objects according to human experience. In this paper, we developed a machine learning model to infer personality traits of human experience objects, like venues, by exploiting how they are mentioned, described, and commented in natural language text: the psycholexical approach 2.0. Specifically, we developed an effective embedding for each personality trait, that is, an embedding in which words are "coherent" with the personality trait. Moreover, we evaluated the performance of machine learning models in identifying the markers of personality traits in two ways. In the first step, we applied a standard model validation procedure, evaluating the models' accuracies in predicting how single words (i.e., words in isolation) captured different personality traits. Data and source code are available at https://github.com/.../personality_predictioncoding. In the second step, we ran two different studies to test whether the terms identified by the machine learning model as markers of each trait (and polarity) were correctly recognized by both experts in personality psychology (Study 1) and laypeople (Study 2). The data that support the findings of both studies are available from the corresponding author. This research was not preregistered. All studies were conducted in compliance with the Declaration of Helsinki ethical standards.

5 | DEVELOPMENT OF THE MACHINE LEARNING MODELS

5.1 | Resources

5.1.1 | GloVe embedding

A word embedding is a representation of words' meaning, typically obtained by combining language models and learning techniques. Word embedding aims to map words or phrases from a given vocabulary to vectors of real numbers with a fixed size. In particular, it involves a mathematical embedding from a high-dimensional space to a small-dimension continuous vector space. In such vector space, the distance between two words depends on their semantic similarity.

GloVe is the embedding produced by a global log-bilinear regression model that combines the advantages of global matrix factorization and local context window methods. Indeed, it can capture the corpus statistics and perform well on the analogy task.

The produced vector space contains a meaningful substructure, and the model outperforms related models on similarity, analogy, and named entity recognition tasks (Pennington, 2014).

Many versions of GloVe pre-trained vectors are made publicly available. In this paper, we use the version trained on 6GB tokens from the corpus extracted from Wikipedia and English Gigaword 5th Edition. The associated vocabulary consists of 400 K terms, each one represented as a 100-dimensional vector.

5.1.2 | Scoring dataset

The personality scoring dataset was created through a scoping review aimed at identifying scientific papers that report relevant data on adjectives measuring the Big Five traits. A web search was performed on Google Scholar combining the following keywords: Big Five, adjectives, marker, loadings, and exploratory factor analysis. We selected papers according to five criteria: (1) the paper and the adjectives had to be in English; (2) the complete list of adjectives used in the study(s) had to be reported; (3) participants had to evaluate one single adjective at a time (i.e., bipolar scales were excluded); (4) participants had to evaluate their own personality (i.e., studies using peer-assessment were excluded); and (5) loadings on each of the Big Five trait had to be reported for each adjective.

The scoping review led to select seven papers (see Table 1), two of which reported data on two different samples, for a total of 4059 cases. Adjectives reported in each paper were then combined in a single list along with their loadings on the five traits. If an adjective was reported by more than one study, a weighted average was computed for each loading based on the studies' sample size to consider the relative magnitude of each loading. Accordingly, we applied Equation (1), where λ represents the loading, N represents the study sample size, S is the total number of considered studies, and i is the number of studies in which the adjective was found.

$$\frac{\sum_{i=1}^{S} \lambda_i * N_i}{\sum_{i=1}^{S} N_i} \tag{1}$$

In other words, loadings were multiplied by the number of cases of each study in which they were found; then, the resulting values were summed across studies and divided by the number of total cases. For instance, the adjective "adaptable" was found in Saucier and Goldberg (1996; $N=899$) and Somer and Golberg (1999; $N=182$), with the highest loadings on agreeableness ($\lambda=0.27$ and $\lambda=0.45$, respectively). The weighted average for agreeableness was computed as:

TABLE 1 List of papers selected in the scoping review, sample size of the studies considered, and number of total and trait-specific adjectives.

		Number of adjectives					
	Sample size	**Total**	**O**	**C**	**E**	**A**	**N**
Goldberg (1992)[a],[b]—Study 4 (self-sample)	320	100	20	20	20	20	20
Saucier (1994b)[a]	636	40	8	8	8	8	8
Benet and Waller (1995)							
American sample	569	43	11	7	8	9	8
Spanish sample	435	43	11	7	8	9	8
Saucier and Goldberg (1996)[a]	899	435	65	86	102	127	55
Somer and Goldberg (1999)[a]—Study 1	182	273	33	65	86	55	34
Perugini et al. (2000)[a]	337	50	10	10	10	10	10
Ledesma et al. (2011)							
Validation sample	372	67	10	12	11	16	18
Replication sample	309	67	10	12	11	16	18

Note: O = openness, C = conscientiousness, E = extraversion, A = agreeableness, N = neuroticism.

[a]The study originally measured the inverse of neuroticism, namely emotional stability.

[b]Extraversion was originally referred to as surgency, whereas openness as intellect.

$$\frac{(0.27 \times 899) + (0.45 \times 182)}{899 + 182} = 0.30$$

The same procedure was applied for each of the adjectives on each of the five loadings. The resulting scoring dataset included 616 adjectives and constituted the list of known terms used in the following modeling steps. Among the adjectives, 26.46% displayed the highest loading on agreeableness, 23.70% on extraversion, 19.97% on conscientiousness, 14.94% for neuroticism, and 14.94% for openness. Concerning the polarity of each trait, the percentage of adjectives showing the highest loading on positive polarity was 48.47% for agreeableness, 54.47% for conscientiousness, 54.79% for extraversion, 56.52% for neuroticism, and 59.78% for openness. Supporting information Appendix C reports a detailed analysis of the scoring dataset and illustrates the score distribution of individual traits.

5.1.3 | Yelp reviews corpus

We use a reviews corpus made publicly available by the Yelp Dataset Challenge (https://www.yelp.com/dataset). It consists of more than 5 million reviews of about 200 K venues located in more than 10 different metropolitan areas, from which we extracted a subset containing 300 K reviews to reduce training time and memory consumption.

5.2 | Procedure

5.2.1 | Enriching word embeddings using personality texts

The goal of this initial step was to process GloVe word embeddings in order to develop five new word embeddings, each one specifically tuned for one of the Big Five traits. The rationale for this approach was that pairs of terms, which are close in a trait-specific personality word embedding, should have a similar influence on the corresponding personality trait. Therefore, we aimed at increasing the geometric coherence between each of the specific word embedding and the associated personality trait score function, in order to better define this latter using the very terms' positions in the word embedding space. The rationale here was that the personality score was well approximated via a locally smooth function. Therefore, we built vector representations, which encapsulated semantic information for each personality trait, in such a way that the position of a word, in the trait-specific personality word embedding, maximally informed about how the word contributes to score the given personality trait. Indeed, such a localized representation of words could help to improve prediction accuracy of personality traits.

We implemented these tunings by training, for each specific personality trait, a Convolutional Neural Network (CNN) on the Yelp reviews corpus. It is worthwhile to mention that venue reviews are a good case study to serve the purpose. Indeed, venues are objects of human experience and capture strong human components; thus, it is possible to assign them personality traits (Graham & Gosling, 2011; Tanasescu et al., 2013). Specifically, as clearly stated by Tanasescu et al. (2013, p. 76), people generally use the same words to describe places and human personality “due to the personality of the individuals that frequent the venue, or because of characteristics of the venues.” Therefore, we assumed that the OCEAN model could be used to describe objects of human experience. Each trait-specific personality embedding was obtained starting from the common GloVe word embedding. Then, the embedding was trained to predict the average personality scores of the reviews for each personality trait. The average personality score of a venue review and, for a given personality trait, was computed by averaging score values across all known terms, mentioned in the given venue review, which were associated with the given personality trait. Therefore, venue reviews not containing any known term were discarded.

In this way, word vector representations were modified to improve prediction performance for each specific personality trait. The training phase thus aimed to modify the weights in the embedding layer to reduce the model's loss function computed on each review's predicted score. This approach allowed detecting which words are most influential in order to predict score values for each personality trait. This step outputted five trait-specific word embeddings, each embedding consisting of the matrix of weights of the embedding layers of the corresponding CNN. The CNN architecture and the detailed training process are described in Supporting information Appendix A.

5.2.2 | Personality score learning

Once we had tuned the GloVe embedding to improve personality score predictions, we were able to learn five models, one for each personality trait. In particular, each Feedforward Neural Network (FNN) model was trained to predict the personality score based on the known terms, which were inputted to the model using their trait-specific word embedding. Once trained, the FNN can be used to estimate personality scores also for unknown terms. The architecture of these models is detailed in Supporting information Appendix B).

5.2.3 | Intrinsic validation of the model

Before validating the model against human intuitions, we tested our approach of having five separate trait-specific word embeddings. To do so, we applied the K-Nearest Neighbors (KNN) test described in Supporting information Appendix D. This test provides initial evidence that our specific enriched embeddings encapsulate more information about personality traits than the original GloVe embedding. To better prove this statement, we conduct a further test by training FNN models to predict the personality scores of a word given its embedding representations. The only terms whose personality scores were given were the known terms. Therefore, we could compute performance for predictions of known terms' scores only.

Our model validation procedure partitioned the set of known terms into 10 folds, to be used in a 10-fold cross-validation (10-fcv) procedure to estimate R^2 and MSE performance measures. Each training cycle consisted of 300 epochs, while performances were computed after each epoch. We considered the performances related to the best-performing epoch for each training cycle. To validate the efficacy of our method, we conducted experiments using three types of word embeddings. The first two are generic embeddings and represent all five personality traits together, taken as baselines. The former is the original GloVe embedding that did not go through the enrichment process described above; the latter is a uniquely enriched word embedding that represents all five personality traits together. Finally, the third type of embedding consists of enriching a specific word embedding for each trait. Additionally, we compared a single FNN model, trained to predict scores for all five personality traits at once, with five separate trait-specific FNN models. For each experiment, we conducted a 10-fold cross-validation, in which we computed the performance by averaging, for each trait, the measured performances on the validation sets of the 10 cross-validation rounds. We standardized the personality scores before starting each 10-fcv round and computed the performances on standardized scores.

Results in Table 2 provide strong evidence that enriching the original word embedding enhances performance significantly. The FNN models trained with the enriched embedding demonstrated superior performance to those trained with the standard GloVe embedding. This enrichment process generated new embeddings that capture word similarities directly linked to specific personality traits. Moreover, using a specific model for each personality trait resulted in the best solution. When utilizing a generic word embedding (either GloVe or the uniquely enriched embedding), the five specific FNN models (each corresponding to a distinct personality trait) outperformed the singular FNN model that covers all traits. Furthermore, FNN models trained using the five specific enriched word embeddings performed better than those trained with the single enriched word embedding. The trait-specific enriching procedure tuned the vectorial representations to express similarities and semantics closely associated with each trait. These results reinforce the value of employing tailored word embeddings for enhanced performance.

Scores of personality traits E and N were the most difficult to predict, while the remaining personality traits performed well, with trait C obtaining the best performance. Considering the best configuration (trait-specific FNN models with trait-specific embeddings), MSE was always smaller than 0.21, while R^2 values were always greater than 0.78. Therefore, for each personality trait, a simple model effectively predicted personality score values using trait-specific word embedding to represent known terms. This evidence supported again the rationale that the geometric coherence achieved by tuned word embeddings helps predict personality scores. However, as already stressed, this test was about known terms only. Generalization to unknown terms was further evaluated through a coherence test, described in detail in Supporting information Appendix E.

TABLE 2 Intrinsic validation of the mode: ten-fold cross-validation test results.

	GloVe				Enriched unique embedding				Enriched specific embeddings	
	Specific models		Unique model		Specific models		Unique model		Specific models	
	MSE	R^2	MSE	R^2	MSE	R^2	MSE	R^2	MSE	R^2
O	0.654	0.36	0.695	0.321	0.19	0.798	0.243	0.743	0.172	0.817
C	0.451	0.521	0.502	0.479	0.125	0.875	0.165	0.835	0.128	0.874
E	0.636	0.36	0.7	0.29	0.25	0.728	0.293	0.686	0.203	0.79
A	0.541	0.451	0.598	0.392	0.22	0.771	0.248	0.747	0.168	0.829
N	0.564	0.409	0.631	0.35	0.223	0.769	0.29	0.699	0.206	0.787

6 | EXTERNAL VALIDATION

Two studies were conceived to validate the model performance in identifying each trait's negative and positive polarity markers. The design of both studies was identical, but Study 1 involved experts in personality psychology, whereas Study 2 was conducted on laypeople.

6.1 | Study 1

6.1.1 | Materials and procedure

Eleven researchers and professionals with strong expertise in personality psychology and the Big Five theory took part in the validation experiment. No remuneration was offered to them. We involved only people with a Ph.D. in psychology and at least 5 years of expertise in the personality field. Experts received an individual invitation to the validation study consisting of 100 trials in a two-alternative forced-choice task.

Trials were divided in 10 blocks that were randomly presented to the experts. Each block focused on one of the two polarities (positive or negative) of the five personality traits. In each trial, experts were presented with a specific polarity of the trait (e.g., high extraversion) and a set of its best markers reported in the literature (e.g., sociable, talkative, and energetic) to increase consistency in traits' definitions among the raters. Along with this information, the experts were asked to select which of the two presented words best describes the polarity of that trait.

For each pair, one word (i.e., the marker) was retrieved from the most representative terms identified by the trait-specific machinelearning model. Specifically, the 10 unknown terms associated with the highest (lowest) marker indices for a specific trait were considered the most representative of its positive (negative) polarity after an a-priori exclusion of the following categories:

1. Proper names of people and places.
2. Terms with less than 10 occurrences in the reviews used to tune the embedding.
3. Terms whose average score of the five nearest known terms in the embedding was lower than 0.15.

Conversely, the other word (i.e., the distractor) was randomly extracted from the whole list of unknown terms (excluding the most representative ones). In each block, the word pairs were randomly presented to the experts.

For example, in the block focusing on the positive polarity of extraversion, the experts were asked to select which of the following two stimuli best describes it: superstar (the marker) vs. polity (the distractor).

6.1.2 | Results

A set of one-tailed binomial tests was conducted to evaluate whether the experts' probability of agreement (i.e., choosing the term predicted by the model as representative of the trait and polarity reported by the trial) was significantly greater than what was expected by chance (50:50). The probability of correctness was evaluated considering different targets: (1) the whole test (i.e., 11 experts judging 100 term pairs, for a total of $N = 1100$), (2) specific polarity (i.e., 11×50, $N = 550$), (3) specific trait (i.e., 11×20, $N = 220$), and trait per polarity (i.e., 11×10, $N = 110$). The results are reported in Table 3; considering we ran 18 tests, p values were corrected by the false discovery rate method, computing the Benjamini–Hochberg adjusted p value. The high probability of agreement observed for most of the targets and the significance of their associated tests indicated a high proportion of matches between experts' choices and model predictions, confirming the model's good performance in identifying new terms related to the Big Five traits. However, there was a low agreement for unknown terms identified as representative of the negative polarity of openness, conscientiousness, and neuroticism, namely the experts' probability of agreement was lower than what expected by chance.

6.2 | Study 2

6.2.1 | Materials and procedure

One hundred participants were enrolled through the prolific online recruitment platform (www.prolific.co), and they were reimbursed with £3.50. The sample consisted of English native speakers (nationality: 97 American, one Irish, one Canadian, and one Italian) with a mean age of 33.60 ($SD = 10.61$), and it was balanced for gender (48 females and 52 males). Study design, that is, the two-alternative forced-choice task, and item lists were identical to Study 1. However, before starting each block, participants were presented with the trait and polarity (as well as the set of its best markers) to be evaluated in subsequent trials, and were asked to familiarize with the specific personality trait before starting the task. Moreover, we added a further trial per block as attention checks, increasing the total trials from 100 to 110 (i.e., 11 per block). One word of each attention check was randomly extracted from the whole list of unknown terms, whereas the other (the correct one) was taken from the set of best markers reported above the trial.

TABLE 3 Study 1 results of the one-tailed binomial probability tests.

Target	*N*	*P* correctness	95% CI	B-H *p* value
Whole test	1100	0.685	0.662, 0.709	<0.001
Polarity				
Positive	550	0.756	0.724, 0.786	<0.001
Negative	550	0.615	0.579, 0.649	<0.001
Trait				
Openness	220	0.573	0.515, 0.627	0.022
Conscientiousness	220	0.668	0.612, 0.721	<0.001
Extraversion	220	0.750	0.697, 0.798	<0.001
Agreeableness	220	0.800	0.750, 0.843	<0.001
Neuroticism	220	0.636	0.580, 0.690	<0.001
Trait × Polarity				
Openness (+)	110	0.691	0.611, 0.763	<0.001
Openness (−)	110	0.455	0.374, 0.537	0.780
Conscientiousness (+)	110	0.818	0.747, 0.876	<0.001
Conscientiousness (−)	110	0.518	0.436, 0.600	0.410
Extraversion (+)	110	0.800	0.727, 0.861	<0.001
Extraversion (−)	110	0.700	0.620, 0.772	<0.001
Agreeableness (+)	110	0.773	0.697, 0.837	<0.001
Agreeableness (−)	110	0.827	0.757, 0.884	<0.001
Neuroticism (+)	110	0.700	0.620, 0.772	<0.001
Neuroticism (−)	110	0.573	0.490, 0.653	0.424

Note: B-H *p* value = Benjamini–Hochberg adjusted *p* value.

6.2.2 | Results

A set of one-tailed binomial tests was conducted to evaluate whether the participants' probability of agreement (i.e., choosing the term predicted by the model as representative of the trait and polarity reported by the trial) was significantly greater than what was expected by chance (50:50). As in Study 1, the probability of agreement was evaluated considering different targets: (1) the whole test (i.e., 100 participants judging 100 term pairs, for a total of $N = 10{,}000$), (2) polarity (i.e., 100×50, $N = 5000$), 3) trait (i.e., 100×20, $N = 2000$), and trait per polarity (i.e., 100×10, $N = 1000$). As a preliminary step, we evaluated the probability of correctness of the attention checks (i.e., 100×10, $N = 1000$). The results are reported in Table 4; considering we ran 18 tests, *p* values were corrected by the false discovery rate method, computing the Benjamini–Hochberg adjusted *p*-value. The attention checks' probability of correctness was extremely high, indicating high compliance of participants. Considering the different targets of the analysis (i.e., traits, polarities, and traits × polarities), the laypeople's behavior was similar to that of experts. Specifically, although the probabilities of agreement observed in Study 2 were generally lower than in Study 1, the results confirmed the model's good performance in identifying new terms related to the Big Five traits. Indeed, the probability of agreement observed was significantly higher than what was expected by chance for most of the targets, indicating a good proportion of matches between participants' choices and model predictions. As in Study 1, the negative polarity of openness and conscientiousness showed that the probability of agreement was not significantly higher than what was expected by chance. However, in contrast with Study 1, this did not occur for low neuroticism.

7 | GENERAL DISCUSSION

The present paper proposes the "psycholexical approach 2.0" an innovative methodology to retrieve information about personality traits from written texts, combining the old-fashioned psycholexical approach and the modern Natural Language Processing methodology, specifically vector-space modeling. Specifically, starting from the literature addressing to what extent adjectives can measure the OCEAN personality traits, we developed five machine learning models, mapping how the words reported in 300 K reviews are associated with each polarity of each

TABLE 4 Study 2 results of the one-tailed binomial probability tests.

Target	*N*	*P* correctness	95% CI	B-H *p* value
Attention check	1000	0.939	0.925, 0.951	<0.001
Whole test	10,000	0.646	0.638, 0.654	<0.001
Polarity				
Positive	5000	0.712	0.701, 0.723	<0.001
Negative	5000	0.580	0.569, 0.592	<0.001
Trait				
Openness	2000	0.584	0.565, 0.602	<0.001
Conscientiousness	2000	0.600	0.581, 0.618	<0.001
Extraversion	2000	0.707	0.689, 0.723	<0.001
Agreeableness	2000	0.741	0.724, 0.757	<0.001
Neuroticism	2000	0.601	0.583, 0.619	<0.001
Trait × Polarity				
Openness (+)	1000	0.703	0.678, 0.727	<0.001
Openness (−)	1000	0.464	0.438, 0.490	0.987
Conscientiousness (+)	1000	0.692	0.667, 0.716	<0.001
Conscientiousness (−)	1000	0.507	0.481, 0.533	0.361
Extraversion (+)	1000	0.795	0.773, 0.816	<0.001
Extraversion (−)	1000	0.618	0.592, 0.643	<0.001
Agreeableness (+)	1000	0.729	0.705, 0.752	<0.001
Agreeableness (−)	1000	0.752	0.729, 0.774	<0.001
Neuroticism (+)	1000	0.642	0.616, 0.667	<0.001
Neuroticism (−)	1000	0.560	0.534, 0.586	<0.001

Note: B-H p value = Benjamini–Hochberg adjusted p value.

trait. We tuned GloVe word embeddings to be specific to each personality trait, in such a way that close terms in trait-specific word embeddings had a similar influence on the corresponding personality trait. This goal was reached by training, for each personality trait, a machine learning model on the reviews corpus. Then, starting from adjective personality loadings in literature, we used trait-specific word embeddings to train five new machine learning models aimed at predicting loadings on personality traits given a certain word. The model validation yielded favorable results, indicating that about 80% of the variance of the five personality traits was explained by the proposed algorithm.

Our approach brings a new perspective in the literature aimed at predicting personality from the language. For example, Cutler and Condon (2022) used contextual word embeddings, which can account for contextual factors by representing the same word differently depending on the context. However, in our use case, we aimed to have a unique representation for each word to predict its loading on a specific personality trait. In this way, we impose a specific context for each word: Its vectorial representation must be helpful to predict its loading for a personality trait. Cutler and Condon (2022) submitted a query for each word to the model to get a unique representation. Suppose one wants to represent different words having different parts of speech (adjectives, verbs, adverbs, and nouns); a specific query should be built for each part of the speech. Our methodology is more robust and scalable; it builds word representations regardless of their part of speech. Furthermore, Cutler and Condon (2022) only represent 435 known personality descriptors and compare the pairwise correlations between words in the new vectorial space to the ones in the principal component analysis carried out on survey-based ratings. In the present paper, we proposed three further steps:

1. We started from a generic word embedding and then built a specific word embedding for each personality trait.
2. Using the traits-specific word embeddings, we trained models to infer the personality loadings of all the language words.
3. By selecting the words with higher predicted loadings, we extended the set of marker terms for each personality trait.

The external validation showed that experts in personality psychology and laypeople agreed on the correctness of the best markers of most traits and polarities (as identified by the machine learning model). Each trait, irrespective of the polarity, was associated with a probability of agreement significantly higher than what was expected by chance; this was observed in both samples, indicating the reliability of our proposed model in capturing dimensions of personality. Generally speaking, the results were comparable across the two samples, even though the experts' agreement with the markers identified by the model seemed slightly higher than those observed for laypeople. Concerning the traits, experts and laypeople showed the highest probability of agreement when considering both the global traits and the two polarities of extraversion and agreeableness. The high performance on these two traits could be ascribed to the high number of adjectives included in the initial scoring dataset. Indeed, considering the entire scoring dataset, 26.46% and 23.70% of adjectives showed primary loading on agreeableness and extraversion, respectively. However, the predominance in terms of markers of these two traits over the others might be related to their very nature. Agreeableness and extraversion are intrinsically related to social and interpersonal dimensions (Graziano & Tobin, 2002; Lucas & Diener, 2001; Schaefer et al., 2012), and this could have increased their relevance in language (and the number of terms used to describe these characteristics). For instance, extraversion and agreeableness were found to be key in developing social capital and, more specifically, predicted people's social–emotional resources (Tulin et al., 2018). Furthermore, it is highly plausible that the terms related to extraversion and agreeableness were the most represented in the corpus used in our study (i.e., the Yelp dataset), since it consisted of reviews on public venues in which the social dimension is crucial. For the remaining traits, data showed a different pattern according to the polarity to be evaluated. Indeed, participants were more likely to choose the markers identified by the model when facing trials about the positive polarity.

Our work differs from the ones described by Ahmad et al. (2020) as it focuses on the personality of the item described in a text and not on the author's one. Furthermore, we applied similar approaches but with different goals. We did not propose a methodology for classifying an item according to a personality trait; we developed, instead, an algorithm to predict the personality loadings of all the language words. These predictions were used to expand the set of marker terms for each personality trait. In future research, using these new sets of terms, it will be possible to build unsupervised models (such as constrained topic models) that are able to classify a venue given its reviews. We also combined deep learning techniques with approaches that did not require manual text labeling. The generic word embedding was, in fact, specialized using a self-supervised strategy that assigned to a review the average of its known terms' loadings. This word embedding enrichment represents the crux of our proposal, highlighting the personality-related information within words' representations. Through our experimental campaign, we substantiated the effectiveness of this approach in enhancing GloVe embeddings and facilitating the prediction of word personality scores. We found that constructing distinct enriched word embeddings for each personality trait outperforms the alternatives of creating a single enriched word embedding for all traits and relying solely on GloVe embeddings.

The present study has some limitations that pertain to both the specific materials used and the approach. First, the papers used to develop the scoring dataset are relatively dated since they were published 30 to 10 years ago. The lack of more recent articles reflects the loss of interest in the psycholexical approach in the last decades. At the same time, the consideration of older articles could be problematic because the use of some adjectives could change over time (as well as their link with personality traits), and some of them could not be used anymore. However, these papers reported methodologically strong research and provided all the details we needed to develop the scoring dataset (e.g., factor loadings). Moreover, due to the lack of more recent articles adding new terms, we could only use 616 adjectives, of which only 571 were represented in the GloVe embedding space. Having a larger number of known terms would have allowed us to better map the personality function in the GloVe embedding space and thus build more coherent trait-specific word embeddings. In addition, the final FNN models would have more training data to rely one, allowing for a higher generalizability of its predictions.

8 | CONCLUSION

The psycholexical approach 2.0 can be an effective methodology to retrieve psychologically relevant information regarding inanimate objects from large text corpora. Unlike other approaches aimed at inferring the personality of the authors of texts, our methodology focuses on extracting personality characteristics of items (i.e., venues) *described* by texts. Such a methodology has been successfully applied to detect OCEAN personality traits of venues, but its practical implications are potentially infinite. For instance, as Ahmad et al. (2020) highlighted, the OCEAN model represents only one of the theoretical frameworks developed to classify personality traits. We argue that our approach might be effectively adopted to detect personality traits belonging to different research traditions, such as Cattell's 16 personality factor model (Cattell et al., 1970) or Eysenck's Giant Three model (Eysenck, 1994), or even other psychological constructs not related to personality. Similarly, we investigated the OCEAN personality of venues. Still, we argue that the same approach can also be applied to people and other inanimate objects, such as retail items, movies, cars, and games.

Beyond identifying psychological constructs, the psycholexical approach 2.0 might be integrated into current recommendation systems to predict venues, products, and items that a specific user might be interested in. Specifically, new items can be recommended based on their similarities with the OCEAN profiles of users' previously preferred/purchased items. Moreover, recommendations could be even more fine-tuned if users' personalities are measured and matched with products with a similar OCEAN personality profile. In conclusion, we argue that the psycholexical approach 2.0 can substantially impact research and practice related to the identification of personality features from textual data.

AUTHOR CONTRIBUTIONS

Conceptualization: FG, MM, FS, DM, LP; Data curation: FG; Formal analysis: FG, DM, LP; Investigation: DM, LP; Methodology: FG, MM, FS, DM, LP; Supervision: MM, FS, DM, LP; Writing - original draft preparation: FG, MM, DM, LP; Writing - review & editing: FG, MM, FS, DM, LP.

ACKNOWLEDGMENTS

We would like to thank researchers and professionals for their helpful contribution in participating in Study 1.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors report no conflicts of interest in this work.

DATA AVAILABILITY STATEMENT

Data and source code for the development of the machine learning model are available at https://github.com/federicogiannini13/personality_prediction; the scoring dataset, the list of stimuli used in Study 1 and Study 2, and the data that support their findings are available at https://osf.io/xtp3n/.

ETHICAL STATEMENT

All studies were conducted in compliance with the Declaration of Helsinki's standards.

PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES

N/A.

PREREGISTRATION

This research and its analysis were not preregistered; therefore, it should be considered exploratory.

ORCID

Federico Giannini https://orcid.org/0000-0002-4210-6271
Marco Marelli https://orcid.org/0000-0001-5831-5441
Fabio Stella https://orcid.org/0000-0002-1394-0507
Dario Monzani https://orcid.org/0000-0001-9544-2933
Luca Pancani https://orcid.org/0000-0002-7834-3826

REFERENCES

Ahmad, H., Asghar, M. Z., Khan, A. S., & Habib, A. (2020). A systematic literature review of personality trait classification from textual content. *Open Computer Science*, *10*(1), 175–193.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*, i–171.

Baumgarten, F. (1933). 'Die Charaktereigenschaften'. [The character traits]. In *Beitraegezur Charakter- und Persoenlichkeitsforschung*. A. Francke.

Benet, V., & Waller, N. G. (1995). The big seven factor model of personality description: Evidence for its cross-cultural generality in a Spanish sample. *Journal of Personality and Social Psychology*, *69*(4), 701–718.

Berger, P., & Luckmann, T. (1966). *The social construction of reality*. Penguin Book.

Bhatia, S., Richie, R., & Zou, W. (2019). Distributed semantic representations for modeling human judgment. *Current Opinion in Behavioral Sciences*, *29*, 31–36.

Burgess, C. (1998). From simple associations to the building blocks of language: Modeling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, & Computers*, *30*(2), 188–198.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "Big Five questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences*, *15*(3), 281–288.

Caprara, G. V., & Perugini, M. (1994). Personality described by adjectives: Generalizability of the Big Five to the Italian lexical context. *European Journal of Personality*, *8*, 357–369.

Cattell, R. B. (1943a). The description of personality. I. Foundations of trait measurement. *Psychological Review*, *50*, 559–594.

Cattell, R. B. (1943b). The description of personality. II. Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, *38*, 476–507.

Cattell, R. B. (1957). *Personality and motivation structure and measurement*. World Book.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the sixteen personality factor questionnaire (16 PF): In clinical,*

educational, industrial, and research psychology, for use with all forms of the test. Institute for Personality and Ability Testing.
Costa, P. T., & McCrae, R. R. (1992). *NEO personality inventory-revised (NEO PI-R)*. Psychological Assessment Resources.
Cutler, A., & Condon, D. M. (2022). Deep Lexical Hypothesis: Identifying personality structure in natural language. *arXiv preprint arXiv*:2203.02092.
De Raad, B. (1992). The replicability of the Big Five personality dimensions in three word-classes of the Dutch language. *European Journal of Personality*, *6*, 15–29.
Eysenck, H. J. (1994). The Big Five or Giant three: Criteria for a paradigm. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 37–51). Erlbaum.
Frank, S. L., & Willems, R. M. (2017). Word predictability and semantic similarity show distinct patterns of brain activity during language comprehension. *Language, Cognition and Neuroscience*, *32*(9), 1192–1203.
Galton, S. F. (1884). Measurement of character. *Fortnightly Review*, *42*, 179–185.
García, O., Aluja, A., & García, L. (2004). Psychometric properties of Goldberg's 50 personality markers for the Big Five model. *European Journal of Psychological Assessment*, *20*, 310–319.
Gatti, D., Rinaldi, L., Marelli, M., Mazzoni, G., & Vecchi, T. (2022). Decomposing the semantic processes underpinning veridical and false memories. *Journal of Experimental Psychology: General*, *151*(2), 363–389.
Goldberg, L. R. (1980). *Some ruminations about the structure of individual differences: Developing a common lexicon for the major characteristics of human personality*. In Annual Meeting of the Western Psychological Association, Honolulu, Hawaii.
Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. *2*, pp. 141–165). Sage.
Goldberg, L. R. (1982). From ace to zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. *1*, pp. 203–234). Erlbaum.
Goldberg, L. R. (1990). An alternative "description of personality": The Big Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.
Goldberg, L. R. (1992). The development of markers for the Big Five factor structure. *Psychological Assessment*, *4*, 26–42.
Gosling, S. D., & John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science*, *8*(3), 69–75.
Graham, L., & Gosling, S. (2011). Can the ambiance of a place be determined by the user profiles of the people who visit it? In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. *5*, pp. 145–152). AAAI Press.
Graziano, W. G., & Tobin, R. M. (2002). Agreeableness: Dimension of personality or social desirability artifact? *Journal of Personality*, *70*(5), 695–728.
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.
Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2–3), 146–162.
Isaka, H. (1990). Factor analysis of trait-terms in everyday Japanese language. *Personality and Individual Differences*, *11*, 115–124.
John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *Big Five inventory*. University of California.
Jones, M. N., Kintsch, W., & Mewhort, D. J. (2006). High-dimensional semantic space accounts of priming. *Journal of Memory and Language*, *55*(4), 534–552.
Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, *114*, 1–37.
Klages, L. (1926). *The science of character*. Allen & Unwin.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240.
Ledesma, R. D., Sánchez, R., & Díaz-Lázaro, C. M. (2011). Adjective checklist to assess the Big Five personality factors in the Argentine population. *Journal of Personality Assessment*, *93*(1), 46–55.
Lenci, A. (2018). Distributional models of word meaning. *Annual Review of Linguistics*, *4*, 151–171.
Louwerse, M. M. (2007). Symbolic or embodied representations: A case for symbol interdependency. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 107–120). Lawrence Erlbaum Associates Publishers.
Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, *3*(2), 273–302.
Lucas, R. E., & Diener, E. (2001). Understanding extraverts' enjoyment of social situations: The importance of pleasantness. *Journal of Personality and Social Psychology*, *81*(2), 343–356.
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.
Marelli, M., Gagné, C. L., & Spalding, T. L. (2017). Compounding as abstract operation in semantic space: Investigating relational effects through a large-scale, data-driven computational model. *Cognition*, *166*, 207–224.
Moreno, J. D., Martinez-Huertas, J. A., Olmos, R., Jorge-Botana, G., & Botella, J. (2021). Can personality traits be measured analyzing written language? A meta-analytic study on computational methods. *Personality and Individual Differences*, *177*, 110818.
Norman, T. (1967). *2,800 personality trait descriptors: Normative operating characteristics for a university population*. University of Michigan.
Ostendorf, F. (1990). *Sprache und persönlichkeitsstruktur: Zur validität des fünf-faktoren-modells der persönlichkeit*. Roderer.
Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). Association for Computational Linguistics.
Perugini, M., Gallucci, M., & Livi, S. (2000). Looking for a simple Big Five factorial structure in the domain of adjectives. *European Journal of Psychological Assessment*, *16*(2), 87–97.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokosy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Rinaldi, L., & Marelli, M. (2020). The use of number words in natural language obeys Weber's law. *Journal of Experimental Psychology: General*, *149*(7), 1215–1230.
Rotaru, A. S., & Vigliocco, G. (2020). Constructing semantic models from words, images, and emojis. *Cognitive Science*, *44*(4), e12830.
Saucier, G. (1994a). Trapnell versus the lexical factor: More ado about nothing? *European Journal of Personality*, *8*, 291–298.
Saucier, G. (1994b). Mini-markers: A brief version of Goldeberg's unipolar Big-Five markers. *Journal of Personality Assessment*, *63*(3), 506–516.
Saucier, G., & Goldberg, L. R. (1996). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, *10*, 61–77.
Schaefer, M., Heinze, H. J., & Rotte, M. (2012). Touch and personality: Extraversion predicts somatosensory brain response. *NeuroImage*, *62*(1), 432–438.
Somer, O., & Goldberg, L. R. (1999). The structure of Turkish trait-descriptive adjectives. *Journal of Personality and Social Psychology*, *76*(3), 431–450.
Tanasescu, V., Jones, C. B., Colombo, G., Chorley, M. J., Allen, S. M., & Whitaker, R. M. (2013). The personality of venues: Places and the five-factors ('Big Five') model of personality. In *2013 Fourth International Conference on Computing for Geospatial Research and Application* (pp. 76–81). IEEE.
Tulin, M., Lancee, B., & Volker, B. (2018). Personality and social capital. *Social Psychology Quarterly*, *81*(4), 295–318.
Tupes, E. C., & Christal, R. C. (1961). *Recurrent personality factors based on trait ratings* (Technical Report). USAF.

SUPPORTING INFORMATION

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

How to cite this article: Giannini, F., Marelli, M., Stella, F., Monzani, D., & Pancani, L. (2024). Surfing the OCEAN: The machine learning psycholexical approach 2.0 to detect personality traits in texts. *Journal of Personality*, *92*, 1602–1615. https://doi.org/10.1111/jopy.12915