Please, Tell Me About Yourself: Automatic Personality Assessment Using Short Self-Presentations

Ligia Maria Batrinca
CIMec, Univ. of Trento and FBK-irst
Corso Bettini, 38068 Rovereto, Italy
batrinca@fbk.eu

Nadia Mana
FBK-irst
via Sommarive 18, 38123 Povo, Italy
mana@fbk.eu

Bruno Lepri
FBK-irst and MIT Media Lab
via Sommarive 18, 38123 Povo, Italy
lepri@fbk.eu

Fabio Pianesi
FBK-irst
via Sommarive 18, 38123 Povo, Italy
pianesi@fbk.eu

Nicu Sebe
DISI, University of Trento
via Sommarive 14, 38123 Povo, Italy
sebe@disi.unitn.it

ABSTRACT
Personality plays an important role in the way people manage the images they convey in self-presentations and employment interviews, trying to affect the other's first impressions and increase effectiveness. This paper addresses the automatically detection of the Big Five personality traits from short (30-120 seconds) self-presentations, by investigating the effectiveness of 29 simple acoustic and visual non-verbal features. Our results show that Conscientiousness and Emotional Stability/Neuroticism are the best recognizable traits. The lower accuracy levels for Extraversion and Agreeableness are explained through the interaction between situational characteristics and the differential activation of the behavioral dispositions underlying those traits.

Categories and Subject Descriptors
H.1.2 [**User/Machine Systems**]: Human Information Processing;
I.5.4 [**Pattern Recognition Applications**]: Computer Vision, Signal Processing;
J.4 [**Computer Applications**]: Social and Behavioral Sciences – *Psychology, Sociology, Economics.*

General Terms
Algorithms, Measurement, Experimentation, Human Factors.

Keywords
Personality trait detection, Big Five, Self-Presentation.

1. INTRODUCTION
Social psychology research showed that personality plays an important role in the way people manage the images they convey in self-presentations and employment interviews, trying to affect the audience first impressions and increase effectiveness [21].

At the same time, other studies provided evidence that interviewers perform personality inferences during personnel selection interviews: for instance, Huffcut et al. [17] found that personality traits and social skills were the most frequently measured constructs.

ICMI'11, November 14–18, 2011, Alicante, Spain.

Moreover, the type and degree of interview structure seem to moderate this correspondence. Roth et al. [36] found that the extent to which interviews inadvertently measure an applicant's personality seems to depend on the extent to which interpersonal skills are allowed to play a role throughout the interview process.

Nonverbal cues play a major role in this complex process ([15], [34]). DeGroot and Gooty [11] examined the mutual influences between personality attributions, performance ratings and interviewees' non-verbal behaviors. Using a structured behavioral interview setting, they found that raters can make personality attributions even using only one channel of information (e.g., acoustic, visual, etc.) and that these attributions mediate the relationships between the interviewee's nonverbal cues and performance ratings. More in detail, Conscientiousness attributions explain the relationship between visual cues and interview ratings. Extraversion attributions mediate the relationship between vocal cues and interview ratings, and Neuroticism attributions had a suppressing effect for both visual and vocal cues. Regarding Extraversion, the results show that the interviewers infer this trait only from the speaker's voice characteristics.

Our goal is to build machines able to automatically detect personality traits from simple acoustic and visual non-verbal signals, extracted from short (30-120 seconds) self-presentations. Following much psychosocial work on first impressions formation, we exploit the concept of *thin slices* [1] to refer to the short amount of expressive behaviors that we, humans, rely on to produce impressively precise judgments on an individual's or group's properties, such as personality, teaching capabilities, negotiation outcomes, etc.

To our knowledge, this is the first time that the self-presentation scenario is addressed. Given the mediating role of personality in the assessment of interviewees' performance, systems capable of understanding personality traits during short self-presentations would be very helpful to interviewers in their personnel and student selection job. At the same time, they could be the core of coaching systems that support people in their preparation for job interviews or, more generally, that help people manage the first impressions they convey, helping them in defining and achieving personal goals. The rest of the paper is organized as follows: the next section reports some previous works addressing automatic recognition of personality traits. Section 3 describes the corpus used in this work. Section 4.1 and 4.2 present the acoustic and visual features extracted from the collected data, while Section 4.2

and Section 4.4 present the analysis and selection of these features. The automatic machine learning approaches used to automatically predict personality traits are described in Section 5 and the achieved results are discussed in Section 5.1. Finally, Section 6 draws the conclusions.

2. RELATED WORK

Previous work on automatic recognition of personality is relatively recent. Pioneering work addressing this task was carried out by Argamon et al. [2] in 2005. They used word categories based on systemic functional grammar (SMG) and relative frequency of function words to train support vector machines (SVMs) for the recognition of two (Extraversion and Emotional Stability) of the five Big Five traits, measured by means of self-reports. Similarly, in 2006 Oberlander and Nowson [29] trained Naïve Bayes classifiers and SVMs, using n-gram features, to isolate four (Agreeableness, Conscientiousness, Extraversion and Neuroticism) of the Big Five traits on a corpus of personal weblogs. Also Mairesse et al. ([24], [25]) worked on recognition of the Big Five personality traits, while systematically investigating the usefulness of different sets of acoustic and textual features, as suggested by psycholinguistic and psychosocial literature. They separately trained their automatic recognition algorithms on personality data coming from self-reports and observed data. The results showed that: (a) Extraversion is the easiest personality trait to model from spoken language; (b) prosodic features play a major role, and (c) the automatic recognition was closer to observed personality than to self-reports.

Olguin et al. [30] collected various behavioral measures, extracted from daily activities of 67 professional nurses in a hospital. The data were collected by means of a socio-meter badge, a wearable device integrating a number of sensors (accelerometer, infrared sensor and microphone) measuring aspects such as physical and speech activity, level of proximity to relevant objects (people, but also beds, etc.), number of face-to-face interactions with others, and social networks parameters. Although the data collection was not targeting prediction of personality traits, a lot of information about personality came out from the correlation analysis conducted by the authors.

More recently, Pianesi et al. [33] and Lepri et al. [22] showed the feasibility of automatically recognizing the Extraversion and Locus of Control personality traits, using simple non-verbal features. This approach was based on the assumption that: a) personality shows up in the course of social interaction and b) thin slices of social behavior are enough to classify personality traits. The first assumption was realized by exploiting classes of acoustic features encoding specific aspects of social interaction (*Activity*, *Emphasis*, *Mimicry*, and *Influence*) and three visual features (head, body, and hands fidgeting). For the second, they demonstrated personality assessment based on inferences from 1-minute-long behavioral sequences.

Kalimeri et al. [19] studied the classification performance of two models incorporating theoretically motivated hypotheses about personality trait relationships and their behavioral manifestations. In the first model, the classification task exploits the relationship between people's personality traits and their behavior to infer traits from observed behaviors. The second model includes the context, meant as an additional causal factor beside personality traits.

Zen et al. [39] used proxemic features (e.g., number of intimate, personal, and social relationships; minimum distance between two subjects, etc.), extracted from video tracking and head pose estimation, to investigate the automatic recognition of two personality traits (Extraversion and Neuroticism). Mohammadi et al. [28] showed how prosodic features can be used to predict the personality assessments of human experts on a collection of 640 speech samples.

Recently, de Oliveira et al. [31] showed that variables derived from the users' mobile phone call behavior as captured by call detail records and social network analysis of the call graph can be used to automatically infer the users' personality traits defined by the Big Five. On the same line, Chittaranjan et al. [10] analyzed the relationship between smartphone usage and self-assessed personality. Their study is based on a large-scale dataset of 8 months of real usage of smartphones by 83 people and personality surveys that are suitable for large mobile or online studies Applications usage, call and SMS logs contained several meaningful relationships to the Big-Five personality framework.

3. SELF-PRESENTATIONS CORPUS

3.1 Participants

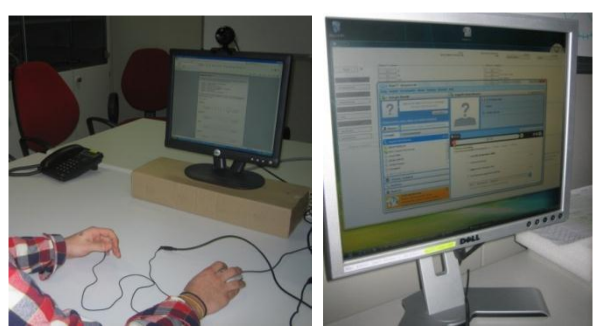
The 89 participants of our study were recruited among employees of a research centre (9 subjects from the administration and 14 subjects from the various research areas), university students (43 subjects) and other external people (23 subjects).

The distribution of the participants was quite balanced in terms of gender (46 male and 43 female) and age (47 young people, i.e. under 25, and 42 adults, i.e. over 25).

Due to the large participation of students, the average age was quite low (29 years) but in any case falling within the adult range (over 25).

3.2 Technical Setup and Recording Procedure

The subjects, involved in individual sessions, were invited to sit in front of a monitor with a webcam on its top (see Figure 1-a). They used the computer mouse just to fill in an online questionnaire (see Section 3.4) and to accept the Skype call from the experimenter at the beginning of the experimental session. A clip-on microphone, worn close to the shirt or sweater collar, was used to communicate with the experimenter.



(a) **(b)**

Figure 1. A view of the experimental setting on the subject's side (a) and on the experimenter side (b)

At the beginning of the session, the experimenter provided each subject with the necessary details. Then she asked the subject to sign the informed consent form and to fill in an online personality questionnaire (see section 3.4.1). After checking that the subjects properly wore the clip-on microphone and securing that the webcam was properly placed for frontal and central shots, the experimenter left for the recording room. Afterwards, she called the subject via Skype, informing him/her that he/she could start the self-introduction session.

3.3 Task Description

The subjects were asked to introduce themselves in front of a camera. Possible topics (e.g. talking about their job, last read book, last holiday, preferred food, preferred sport, etc.) were suggested. However, they were left absolutely free to choose any of them or any other topics they liked.

In order to obtain a good quality frontal image of the subject to be later used for initialization purposes, the participant was initially asked to look into the camera for a few seconds without moving.

The length of resulting self-presentations ranged from 30 up to 120 seconds.

3.4 Personality

We used the Big Five model, asking to the subjects to fill in the Italian version of the questionnaire [18], presented online.

3.4.1 Big Five Questionnaire

The Big Five questionnaire owes its name to the five traits that it takes as constitutive of people's personality:

1. Extraversion vs. Introversion
(sociable, assertive, playful vs. aloof, reserved, shy)
2. Agreeableness vs. Disagreeable
(friendly, cooperative vs. antagonistic, faultfinding)
3. Conscientiousness vs. Un-conscientiousness
(self-disciplined, organized vs. inefficient, careless)
4. Emotional stability vs. Neuroticism
(calm, unemotional vs. insecure, anxious)
5. Creativity or Openness to experience
(intellectual, insightful vs. shallow, unimaginative)

In the standard questionnaire, validated on the Italian language [32], each trait is investigated through ten items, each of them assessed by means of 1 to 7 points scale.

3.4.2 Questionnaire Scores

The Big Five scores were calculated summing single raw scores, properly inverted, per each personality trait. The results (average and standard deviation values) are reproduced in Table 1.

Table 1. Averages and standard deviations of Big Five scores

	Women		Men	
	<25	≥25	<25	≥25
Agree	49.05 (6.383)	50.68 (5.867)	47.65 (7.985)	51.45 (7.564)
Consc	41.76 (9.643)	46.36 (8.693)	39.45 (9.944)	48.25 (9.124)
EmSt	36.57 (7.047)	36.59 (5.518)	40.50 (7.458)	44.65 (6.968)
Creat	45.43 (6.896)	46.55 (7.482)	48.69 (6.279)	47.00 (7.861)
Extra	47.00 (7.791)	45.50 (8.245)	42.04 (9.075)	42.55 (8.370)

In order to investigate the dependence of traits' scores on gender and age, the latter was split in two classes: people younger than 25 and people 25 or more years old. The analysis was conducted by means of a series of ANOVA with dependent variables, the factorial scores for Extraversion, Agreeableness, Conscientiousness, Emotional Stability and Creativity, and Age and Gender as predictors. The results are reported in Table 2.

As evident from the table, no effects were found for Agreeableness and Creativity. Conscientiousness exhibited an Age effect ($F=11.387$, $p<.05$) and Emotional Stability and Extraversion a Gender effect ($F=17.054$ and $F= 4.859$ respectively, $p<.05$). According to our data, therefore, people tend to become more conscientious with age (47.26 vs. 40.43) and independently of their gender; on the other hand, men are more emotionally stable than women (42.30 vs. 36.58), whereas women are more extravert (46.23 vs. 42.26).

Table 2. Dependence of personality traits' scores on subjects' gender and age. F-statistics values for main and interaction effects from ANOVA analyses. *, p<.05; **, p<.01; ***p<.001; n.s., not significant.

	Gender	Age	Gender*Age
Agree	n.s.	n.s.	n.s.
Consc	n.s	11.387***	n.s.
EmSt	17.054***	n.s.	n.s.
Creat	n.s.	n.s.	n.s.
Extra	4.859*	n.s.	n.s.

4. NON-VERBAL CUES AND FEATURES

4.1 Acoustic and Visual Cues

Many acoustic and visual cues have been shown to be related to personality traits. In particular, the importance of acoustic features, such as pitch and intensity, for personality has often been pointed out ([12], [37]). Furnham [14] discussed how extraverted people are characterized by a particular speaking style: they talk more, louder, faster and have fewer hesitations. Others pointed out that frequency of hand movements/gestures was positively correlated with Extraversion [9]. Riggio [35] suggested that extraverts have a higher head movement frequency than introverts and they change their posture more often. Moreover, they maintain eye-contact for longer [26], have higher speaking and gestural fluency. Laughing was associated with high scores in Openness to Experience/Creativity [27]. Many visual cues have been correlated to Agreeableness ([13], [8]). Conscientiousness was correlated with gaze, speaking fluency, speech rate and hand movements ([13], [8], [3]).

Drawing on this and related literature, we derived a set of audio and visual cues, extending their usage also to traits for which less evidence was available, e.g., Creativity. Pitch and acoustic intensity were automatically extracted using Praat [6]. The pitch algorithm is based on the autocorrelation method. Setting the default values for the time step and pitch ceiling (600 Hz) parameters, but lowering the pitch floor from 70 Hz to 50 Hz, the algorithm, uses a time step of 0.015 seconds and a window length of 0.06 seconds. It computes 67 pitch values per second. Intensity is calculated taking into consideration the minimum periodicity frequency in the signal [7]. Visual cues (see in Table 3) were hand-annotated by means of the ANVIL annotation tool [20].

Table 3. Visual cues, manually annotated

Cue	Label	Notes
Eye-Gaze	Up down right left closedEyeLid desktopCtc camCtc	a) upward directed -gaze; b) downward directed-gaze; c) rightward directed gaze; d) leftward directed gaze; e) lids closed for longer than 2 sec; f) -gaze directed toward the desktop; g) -gaze directed toward the webcam.
Frowning	Yes, no	
Hand Movement	MovFace MovAir MovBody Stillface StillAir StillBody	a) hand(s) on the face; b) hand(s) moving in the air; c) hand(s) on parts of the body, except on the face.
Head orientation	Left right down up front rightSide leftSide rightIncl leftIncl	a) head oriented left; b) head oriented right; c) head oriented down; d) head oriented up; e) head oriented frontal; f) head tilted right; g) head tilted left; h) head half oriented right; i) head half oriented left
Mouth Fidgeting	Smile tongueLips biteLips tightLips retractLips	a) smiles; b) subject passing her tongue over her lips; c) subject biting her lips; d) subject pressing her lips; e) subject moving her lips by lowering both mouth corners.
Posture	Back straight forward	a) subject leaning back; b) straight posture; c) subject leaning forward.

4.2 Extracted Features

From the acoustic and visual cues of the previous section we extracted the 29 features of Table 4, to be used in our classification experiments: 17 visual features, 9 acoustic features and 3 features concerning the total speech time, the average duration of voiced segments and the length of the self-presentation.

4.3 Feature Analysis

The (linear) relationships between our features and the Big Five personality traits were analyzed by means of number of backward linear regression analyses, one per each trait, with all our features as predictors.

Table 5 reports for each of the final models the retained predictors and the portion of dependent variable variance explained. As can be seen, the linear models account only for a small portion of variance.

Table 6 reports the partial correlations between the retained predictors and the various traits. Agreeable people tend to gain more often the straight posture, to have a lower maximum pitch and produce longer presentations. More conscientious people smile more frequently and use a longer portion of their presentation for speaking than less conscientious ones; but they move their head around less, tend to exhibit a lower average pitch and lower minimal vocal energy, and produce shorter (on average) voiced segments. More creative people lean towards the camera more often than less creative ones but gesticulates less. Higher emotional stability is associated with a greater number of rather short leaning forward events, lower amount of gesticulation and lower vocal intensity.

Table 4. Feature list

	Ref	Label	Description
Visual Features	1	avDurationBack	Average duration of leaning back episode
	2	avDurationCam	Average duration of looking into the webcam episodes
	3	avDurationDown	Average duration of looking down episodes
	4	avDurationFrow	Average duration of frownings
	5	avDurationFwd	Average duration of leaning forward episodes
	6	avDurationStr	Average duration of straight posture episodes
	7	freqBack	Rate of leaning backward postures
	8	freqCam	Rate of looking into the webcam events
	9	freqDown	Rate of looking down events
	10	freqForward	Rate of leaning forward postures
	11	freqFrowning	Rate of frownings
	12	freqHandMoving	Rate of hand-movement events.
	13	freqHandStill	Rate of hand-still events
	14	freqMouth	Rate of lip moving or biting events
	15	freqSmile	Rate of smiles
	16	freqStraight	Rate of straight postures
	17	NrHeadOrient	Rate of head orientation change
Acoustic Features	18	pitch_mean	Mean of Pitch
	19	pitch_min	Minimum of Pitch
	20	pitch_max	Maximum of Pitch
	21	pitch_med	Median of Pitch
	22	pitch_sdev	SD of Pitch
	23	i_mean	Mean of Intensity
	24	i_min	Minimum of Intensity
	25	i_max	Maximum of Intensity
	26	i_sdev	SD of Intensity
Additional Features	27	avTimeVoiced	Av. duration of voiced segments
	28	TimeVoiced	Portion of self-presentation taken by speech
	29	videoLength	Total length of self-presentation

Finally, extraverts produce longer voiced segments, but smile and frown less frequently and maintain for shorter time the straight posture in front of the camera.

Table 5. Retained predictors and portion of dependent variable variance explained

	Retained predictors	R^2
Agree	16, 20, 29	.127
Consc	15, 17, 18, 23, 27, 28	.188
Creat	10, 12	.107
EmSt	5, 10, 12, 13, 23	.148
Extra	6, 11, 15, 27	.172

Table 6. Partial correlations between the retained predictors and the various traits; see Table 4 for features' reference numbers.

Feat	16	20	29
Agree	.303	-.216	.217

Feat	15	17	18	23	27	28
Consc	.203	.208	-.337	-.207	-.183	.274

Feat	10	12
Creat	.321	-.195

Feat	5	10	12	13	23
EmSt	-.280	.290	-.258	.287	-.195

Feat	6	11	15	27
Extra	-.259	-.184	-.265	.240

4.4 Feature Ranking

A known problem in classification tasks is to find strategies to reduce the dimensionality of the feature space in order to avoid over-fitting. In our experiment, we applied the Weka implementation of the Support Vector Machine Recursive Feature Elimination (SVM-RFE) algorithm (called Support Vector Machine attribute evaluation method in Weka) in order to evaluate the importance of a feature. This algorithm was introduced by Guyon et al. [16] in a cancer classification problem with the goal of finding a subset of features which maximize the performance of the classifier. Roughly speaking, a linear SVM is a hyper-plane that separates two classes of examples (positive and negative) maximizing the separation margin. The SVM creates a weight vector, where a weight is assigned to each feature. The weight vector is used to determine the least important feature, defined as the one with the *smallest* weight. At each iteration, the least important feature is removed and the procedure is repeated on the reduced feature set. This method is used with a Ranker search method and the features are ranked according to the square of the weights assigned to them. Hence, the first feature is the most relevant for the classification task at hand and the last feature of the least relevant one.

For reasons of space, we do not report here the feature rankings produced; see the next section for more on this topic.

5. AUTOMATIC PREDICTION OF PERSONALITY TRAITS

For the sake of our classification experiments, all personality traits' scores were quantized (Low/High) along their median values (Agreeableness = 50, Consciousness = 44, Creativity = 46, Emotional Stability = 39, and Extraversion = 45).

Three machine learning algorithms, namely Naïve Bayes, SVM with linear kernel and SVM with RBF one, were used in 5 binary classification tasks, one per personality trait. The bound-constrained SVM classification algorithm was used for the two SVM classifiers. The cost parameter C and the RBF kernel parameter γ were estimated trough an inner leave-one- out cross validation on the training set of the first fold using the first 88 subjects for the parameter estimation. The best performing values of the parameters were then kept fixed for the outer-cross validation.

Naïve Bayes is a simple probabilistic classifier that applies the Bayes theorem and assumes that the presence/absence of a particular feature of a given class (e.g. a personality state) is unrelated to the presence/absence of any other feature. The main advantage of using Naïve Bayes is that it only requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification.

For each classifier and for each trait, we executed 29 classification runs, each exploiting a subset of the features aggregated according to the ranking provided by the SVM-RFE algorithm. We started from the single feature experiment using the first ranked feature, then executed the 2-feature experiment with the first two ranked features, and so on.

The leave-one-out cross-validation strategy was employed. Hence, 89 models for each personality trait were trained on 88-subject subsets, evaluating them against the remaining ones and finally averaging the results.

5.1 Results and Discussion

Table 7 reports, for each classifier and for each trait, the feature combination producing the highest accuracy value. All the accuracy values reported are statistically significant, according to binomial tests that compared the observed accuracy to that of the baseline classifier that exploits the observed frequencies of the two classes. The significance was set at $p \leq 0.01$ and the resulting threshold values for accuracy where 0.618 for Agreeableness, Conscientiousness and Emotional Stability, and 0.629 for Creativity and Extraversion. The highest accuracy values, greater or equal to 70%, were obtained on Conscientiousness (SVM-RBF), Emotional Stability (SVM-RBF and SVM-Lin) and Extraversion (SVM-RBF and Bayes). To further characterize the predictive power of our features and the behavior of the classifiers, we investigated how they worked on the two classes each trait was split into. We started from the confusion matrix of the conditions in Table 7 and compared the hits for the High and Low classes with those expected from the baseline classifier. The comparison was conducted by means of Pearson residuals, standardized scores - they are $N(0, 1)$ - that measure the difference between observed and expected outcomes. On hits, the absolute value of a Pearson residual measures how much the classifier performs better (positive sign) or worse (negative sign) than the baseline in terms of recall. For errors, the reverse is true. Here we focus on hits.

Table 7. Accuracy (Acc), recall on the Low class (Rec1) and on the High class (Rec2) for the best conditions per each personality trait. The features used are also reported; see Table 4 for the reference number.

		SVM-RBF	SVM-Lin	Bayes
Agree	Acc, Rec1, Rec2	65.16, 50, 80	65.16, 50.00, 80.00	64.04, 34.09, 93.33
	Features	8, 22, 6, 15, 29, 10	8, 22, 6, 15, 29, 10	8, 22, 6, 15, 29, 10, 16
Consc	Acc, Rec1, Rec2	73.03, 73.91, 72.09	68.53, 76.09, 60.47	
	Features	19, 22, 27,6, 14, 24, 20, 17, 25	19, 22, 27,6, 14, 24, 20, 17, 25, 12, 9, 16	
Creat	Acc, Rec1, Rec2	64.04, 72.00, 53.85	66.29, 76.00, 53.85,	64.04, 68.00, 58.97
	Features	16, 8, 15, 26, 12, 7, 6, 4, 24, 20, 14, 13, 2, 5, 1, 17, 9, 11, 21, 3, 22	16, 8, 15, 26	16, 8, 15, 26, 12, 7, 6, 4, 24, 20, 14, 13, 2, 5, 1, 17, 9, 11
EmSt	Acc, Rec1, Rec2	76.40, 80.44, 72.09	75.28, 82.61, 67.44	64.04, 45.65, 83.72
	Features	25, 22, 20, 13, 12, 10, 16, 9, 2, 19, 8	25, 22, 20, 13, 12, 10, 16, 9, 2, 19	25, 22, 20, 13, 12, 10, 16, 9, 2
Extra	Acc, Rec1, Rec2	70.78, 76.60, 64.29		69.66, 74.47, 64.29
	Features	24		24

Finally, we took advantage of the $N(0, 1)$ distribution of Pearson residual and fixed a threshold of ±3sd for the statistical significance of the difference between observed and expected hits. Table 8 reports the results.

Table 8. Pearson residual for the condition of Table 4. Only residuals greater or equal in absolute value than 3 are reported.

	SVM-RBF		SVM-Lin		Bayes	
	Low	**High**	**Low**	**High**	**Low**	**High**
Agree		4.25		3.95		5.74
Consc	3.02	3.12	3.31			
Creat						
EmSt	3.90	3.12	4.20			4.65
Extra	3.27				2.97	

SVM-RBF was the only classifier yielding balanced performances on at least two traits: Conscientiousness and Emotional Stability. In all the other cases, good performances arose only from either the Low or the High class. For instance, with Agreeableness good performances are limited to the High class whereas with Extraversion they are limited to the Low class. For Creativity, no classifier yielded significant performances on any class.

We think it is by no chance that Conscientiousness and Emotional Stability are the traits that yielded the best results, both in terms of accuracy values and of the level of balance between the recall values for the Low and High classes. Conscientiousness, in fact, relates to the individual's capacities for behavioral and cognitive control. Conscientious individuals are described as responsible, attentive, careful, persistent, orderly, and planful; those low on this trait are irresponsible, unreliable, careless, and distractible. High conscientiousness has been connected to positive engagement within task-related behavior [4]. Apparently, the request of introducing themselves in front of a monitor, with camera and microphone on, activated our subjects' Conscientiousness dispositions, doing so for both those high and those low in this trait. Those dispositions, in turn, affected some of the considered behaviors, including: the dynamics of pitch (minimal and maximal pitch, and pitch range as captured by pitch_sdev) and of voice energy (minimal and maximal intensity) on the acoustic side; posture (average duration of episodes of sitting straight in front of the monitor) and head movements and lip-related events on the visual one.

Emotional stability and its counterpart, neuroticism, concern the susceptibility to negative emotions, which we one might think are elicited by the task of introducing oneself in front of a computer screen while being audio-video recorded. Emotional Stability/Neuroticism includes both anxious and irritable distress. The former is inner-focused and includes dispositions to anxiety, sadness and insecurity. Irritable distress, in turn, involves outer-directed hostility, anger, frustration and irritation. We can easily figure out that the self-introduction task activates one of those two sets of dispositions, depending of the person's internal constitution. The behavioral signs that proved effective concerned pitch dynamics (maximum and minimum pitch, and pitch standard deviation), maximum voice intensity, and several visual features: dynamic of hand movements, posture dynamics (straight and forward position), camera fixation and camera aversion, hand fidgeting.

According to this line of explanation, both the request of executing a task and the nature of the task (introducing oneself) activated dispositions connected to Conscientiousness and Emotional Stability. We think that a similar rationale can be given to account for the performances with other traits.

A good accuracy level was obtained with Extraversion too, but this result was basically due to the performance with introverts. One might suggest that the unbalance might be (at least partially) due to the fact that the situation/task was not appropriate for fully activating Extraversion-related dispositions, especially with extraverts. It has been argued, in fact, that the core of Extraversion lies in the tendency to behave in a way so as to engage, attract and enjoy social attention, *i.e.*, extraverts invest time and energy in activities that attract the attention of others ([5],[23]). One possible consequence of this view is that Extraversion-related disposition are activated to a lesser extent in situation like the one we are considering here where there are no others whom to attract attention from, and that this should affect especially extraverts' behavior. Notice, finally, that only one feature (minimal vocal intensity) is used in the best case.

Agreeableness does not reach the same levels of accuracy as Conscientiousness and Emotional Stability. However, the recall for the High class, both measured in absolute term (see Table 7) and through Pearson residuals (see Table 8) is very high: in absolute terms, it reaches 80% with SVM-RBF and SVM-Lin and 93% with Naïve Bayes. This trait includes a number of dispositions that foster congenial social behavior: generosity, consideration, cooperation, willingness to accommodate others wishes, etc. Agreeable people do not aim to attract social attention like extraverts, but to please others. We believe that, again, the key to understanding the performance of our features and classifiers with this trait is in the nature of the situation: it activates a pleasing attitude that somehow masks non-agreeable dispositions (being aggressive, rude, manipulative, etc.).

Finally, the dispositions linked to Creativity seem to be uniformly activated to much a lesser extent by the task of introducing oneself.

6. CONCLUSION

The aim of this paper was to contribute to advance the state of the art in the automatic analysis of people personality. In particular, we investigated the feasibility of detecting the Big Five in short videos of self-presentation. In doing so, we adopted a thin-slice perspective and we investigated the contribution of a large set of different acoustic and visual non-verbal cues.

The main findings of our paper are the following:

a) Conscientiousness and Emotional Stability are the easiest traits to automatically detect during self-presentation. The reason could be that the first trait is connected to engagement within task-related behavior and that the second one is connected to the emotional reactions (e.g. distress) it elicits.

b) Our task does not seem to activate the full range of dispositions of Agreeableness and Extraversion. For the latter, the reasons can be that introducing oneself in front of a computer screen does not provide enough social audience to let the social attention dispositions of extraverts fully activate. As to Agreeableness, we have invoked the masking effect of the necessity of pleasing the experimenter, implicit in the nature of the situation.

On the practical side, our results are a first important step towards automatic systems assisting either interviewers or interviewees in improving their performance in job interviews. On a more theoretical side, they emphasize the influence of the situation for a full unfolding of the behavioral dispositions tied to personality traits.

Of course, more work is needed to fully explore the automatic analysis of personality traits in self-introduction, e.g., by considering even larger sets of non-verbal feature, as well as verbal ones (e.g., lexical choice, presence of emotion-related words; topic dynamics, etc.); using larger samples and/or exploiting regression or ordinal techniques. Finally, we also consider the possibility of extending the work to interviewer/interviewee interactions by collecting new data to model this scenario too, where the system can work with the mediation of an interviewer.

7. ACKNOWLEDGMENTS

The data collection and research presented in this paper have been supported by the EU NETCARITY Project - Integrated Project of the Sixth Framework Programme (IST-2006-045508). Bruno Lepri's research is funded by PERSI project inside the Marie Curie COFUND – 7th Framework.

8. REFERENCES

[1] Ambady, N. and R. Rosenthal. 1992. 'Thin slices' of expressive behaviors as predictors of interpersonal consequences. A meta analysis. Psychological Bulletin. 111, 156-274.

[2] Argamon, S., Dhawle, S., Koppel, M., and Pennbaker, J. 2005. Lexical predictors of personality type. In Proceedings of Interface and the Classification Society of North America.

[3] Aronovitch, C. D. 1976. The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psyhcology*, 99, 207-220.

[4] Ashton MC, Lee K. 2001. A theoretical basis for the major dimensions of personality. *European Journal of Personality*. 15:327–53.

[5] Ashton, M. C., Lee, K., and Paunonen, S. V. 2002. What is the Central Feature of Extraversion? Social Attention Versus Reward Sensitivity. *Journal of Personality and Social Psychology*, 83, 245-252.

[6] Boersma, P., Weenik, D. – Praat: doing phonetics by computer. Computer Program, Version 5.2.05. www.praat.org

[7] Boersma, P., Weenik, D. – Praat: doing phonetics by computer. User Manual, Version 5.2.05. www.praat.org

[8] Borkenau, P. and Liebler, A. 1992. Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645-657.

[9] Brebner, J. 1985. Personality theory and movement. In B. D. Kirkcaldy (Ed.), Individual differences in movement (pp. 27-41). Lancaster, England: MTP Press Limited.

[10] Chittaranjan, C., Blom, J., and Gatica Perez, D. 2011. Who's Who with Big Five: Analyzing and Classifying Personality Traits with Smartphones. In Proceedings of IEEE International Symposium on Wearable Computers (ISWC), San Francisco.

[11] DeGroot, T., and Gooty, J. 2002. Can Nonverbal Cues be Used to Make Meaningful Personality Attributions in Employment Interviews. In Journal of Business Psychology, 24, pp. 179-192.

[12] Dewaele, J. M., Furnham, A. 2000. Personality and speech production: a pilot study of second language learners. In *Personality and Individual Difference*.

[13] Funder, D. C. and Sneed, C. D. 1993. Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, 64, 470-490.

[14] Furnham, A. 1990. Language and personality. In K. R. Scherer and H. Giles and W. Robinson (Eds). *Handbook of Language and Social Psychology,* Winley.

[15] Gifford, R., Ng, C., & Wilkinson, M. 1985. Nonverbal cues in the employment interview: Links between applicant qualities and interviewer judgments. The Journal of Applied Psychology, 70.

[16] Guyon I, Weston J., and Barnhill S. 2002. Gene Selection for Cancer Classification using Support Vector Machines. In Machine Learning, volume: 46, Issue: 1, Publisher: Springer, Pages: 389-422.

[17] Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. 2001. Identification and meta-analytic assessment of psychological constructs measured in employment interviews. Journal of Applied Psychology, 86, 897−913.

[18] John, O. P., Srivastava, S. 1999. The Big five trait taxonomy: History, measurement and theoretical perspectives. In Pervian, L. A. & John, O. P., (Eds.) Handbook of personality theory and research. Guilford Press. New York.

[19] Kalimeri, K., Lepri, B., and Pianesi, F. 2010. Causal Modelling of Personality Traits: Extraversion and Locus of Control. Proceedings of the Second International Workshop on Social Signal Processing (SSPW 2010)

[20] Kipp, M. 2001. Anvil – A generic Annotation Tool for Multimodal Dialogue. In Proceedings of the 7th European Conference on Speech Communication and Technology, Eurospeech, pp. 1367-1370.

[21] Leary, M.R. and Allen, A.B. (in press). Personality and Persona: Personality Processes in Self-Presentation. In Journal of Personality.

[22] Lepri, B., Mana, N., Cappelletti, A., Pianesi, F., Zancanaro, M. 2009. Modeling Personality of Participants during Group Interaction. In Proceedings UMAP – User Modeling, Adaptation and Personalization. Trento, Italy.

[23] Lepri, B., Subramanian, R., Kalimeri, K., Staiano, J., Pianesi, F., and Sebe, N. 2010. Employing Social Gaze and Speaking Activity for Automatic Determination of the Extraversion Trait. In *Proceedings of the ACM International Conference on Multimodal Interfaces, ICMI 2010*. Beijing, China, 2010.

[24] Mairesse, F., and Walker, M. 2006. Automatic recognition of personality in conversation. In Proceedings of HLT-NAACL.

[25] Mairesse, F., and Walker, M. 2006. Words mark the nerds: Computational models of personality recognition through language. In Proceedings of the 28th Annual Conference of the Cognitive Science Society, pp. 543–548.

[26] Mehrabian, A. 1969. Significance of posture and position in the communication of attitude and status relationships. In *Psychological Bulletin* 71(5), 359-372.

[27] Mergl, R., Vogel, M., Prässl, A., Graf, B., Karner, M., Mavrogiorgou, P., Hegerl, U.,, Juckel, G. 2006. Facial Expressions and Personality: *A Kinematical Investigation during an Emotion Induction Experiment*. Neuropsychobiology, 54, 114-119.

[28] Mohammadi, G., Vinciarelli, A., and Mortillaro, M. 2010. The Voice of Personality: Mapping Nonverbal Vocal Behavior into Trait Attributions Proceedings of the Second International Workshop on Social Signal Processing (SSPW 2010).

[29] Oberlander, J. and Nowson, S. 1992. Whose thumb is it anyway? Classifying author personality from weblog text. In Proceedings of the Annual Meeting of the ACL. Association for Computational Linguistics, 627-634.

[30] Olguin, D., Gloor, P.A., and Pentland, A. 2009. Wearable Sensors for Pervasive Healthcare Management. Proceedings of 3rd International Conference on Pervasive Computing Technologies for Healthcare, London, UK.

[31] Oliveira, R.; Karatzoglou, A.; Concejero, P.; Armenta, A.; Oliver, N. 2011. Towards a psychographic user model from mobile phone usage. *Proc. of 29th ACM conference on human factors in computing systems, Works-in-Progress session (ACM CHI)*. Vancouver, Canada.

[32] Perugini, M. and Di Blas, L. (2002). Analyzing Personality related Adjectives from an Eticemic Perspective: The Big Five marker Scales (BFMS) and the Italian AB5C Taxonomy. In B. De Raad and M. Perugini (eds.) Big Five Assessment, Hogrefe und Huber Publishers, Göttingen, 281-304.

[33] Pianesi, F., Mana, N., Cappelletti, A., Lepri, B., and Zancanaro 2008. Multimodal Recognition of Personality Traits in Social Interactions. In Proceedings of International Conference on Multimodal Interfaces, Chania, Crete, Greece.

[34] Raza, S. M., & Carpenter, B. N. 1987. A model of living decisions in real employment interviews. The Journal of Applied Psychology, 72, 596–603.

[35] Riggio, R., Friedman, H. 1986. Impression formation: The role of expressive behavior. In *Journal of Personality and Social Psychology* 50(2), 421-427.

[36] Roth, P. L., Van Iddekinge, C. H., Huffcutt, A. I., Eidson, C. E., Jr., & Schmit, M. J. 2005. Personality saturation in structured interviews. International Journal of Selection and Assessment, 13, 261−273.

[37] Scherer, K. R. 1979. Personality markers in speech. In K. R. Scherer and H. Giles (Eds)., *Social Markers in Speech.* Cambridge University Press.

[38] van Dam, K. 2003. Trait perception in the employment interview: A five-factor model perspective. International Journal of Selection and Assessment, 11, 43−55

[39] Zen G., Lepri B., Ricci E., Lanz O. 2010. Space Speaks - Towards Socially and Personality Aware Visual Surveillance. Multimodal Pervasive Video Analysis Workshop - MPVA 2010 - workshop in conjunction with ACM MM 2010, Firenze, Italy, October 29th, 2010.