Handwriting and Drawing Features for Detecting Personality Traits: An Analysis on Big Five Sub-dimensions*

Anna Esposito[1], Terry Amorese[1], Michele Buonanno[1], Marialucia Cuciniello[1], Antonietta M. Esposito[2], Marcos Faundez-Zanuy[3], Laurence Likforman-Sulem[4], Maria Teresa Riviello[1], Carmine Spagnuolo[5], Alda Troncone[1], and Gennaro Cordasco[1]

[1]Department of Psichology, University of Campania "L. Vanvitelli", Viale Ellittico, 31 - 81100 Caserta, and IIASS, Italy
Email: anna.esposito@unicampania.it, terry.amorese@unicampania.it michele.buonanno@studenti.unicampania.it, marialucia.cuciniello@unicampania.it, mariateresa.riviello@unicampania.it, alda.troncone@unicampania.it, gennaro.cordasco@unicampania.it

National Institute of Geophysics and Volcanology (INGV/OV), Via Diocleziano, 328, 80125 Napoli, Italy; Email: antonietta.esposito@ingv.it

[3]Escola Superior Politecnica, TecnoCampus Mataro-Maresme Mataro, Carrer d'Ernest Lluch, 32, 08302 Mataró, Barcelona, Spain
Email: faundez@tecnocampus.cat

[4]Télécom ParisTech, Université Paris-Saclay, 19 Place Marguerite Perey, 91120 Palaiseau, France; Email: likforman@telecom-paristech.fr

[5]Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132 - 84084 Fisciano (SA), Italy; Email: cspagnuolo@unisa.it

Abstract: Handwriting and Drawing are functional tasks involving physical and cognitive processes. Recently they have been investigated for detecting cognitive and motor disorders. In this work, handwriting/drawing features are investigated for identifying connections with personality traits. For this purpose, an experiment comprising seven handwriting/drawing tasks has been administrated to 78 young adults (mean age=24.6 ± 2.4 years) equally balanced by gender. Handwriting and Drawing activities - both on and close to the paper – had been recorded online through a digitizing tablet able to measure handwriting and drawing features such as pressure, speed, dimension, and inclination of

* A preliminary version of this paper was presented at the 10th IEEE International Conference on Cognitive Infocommunication (CogInfoCom 2019), [4]

each pen-stroke on the paper. Participants were asked to fill the Big Five Personality Questionnaire (BFQ) and according to the scores obtained for each of the 5 dimensions and 10 Big Five sub-dimensions, were partitioned into three categories: low, typical, and high. To evaluate whether the recorded handwriting/drawing features are connected with personality traits ANOVA repeated measures have been performed with gender and group category (low, typical, and high) as between and the listed handwriting/drawing features as within factors. The analyses show significant differences among low, typical and, high BFQ scores for the main Big Five dimensions and the ten Big Five sub-dimensions, indicating that personality traits can be revealed by a quantitative analysis of the proposed handwriting/drawing features.

Keywords: Personality traits; Big Five; Handwriting and drawing tasks; Graphology

1 Introduction

The act of writing and drawing by hand is the result of a complex interaction of physical and mental processes involving several cognitive, kinesthetic, and perceptual-motor skills. Indeed, handwritten texts convey considerable information on how some areas of the human brain are working [24]. Neurology, for example, utilizes writing and drawing tasks as a noninvasive method for diagnosing and monitoring of disorders such as Alzheimer's and Parkinson's disease, and developmental impairments among others [10, 12, 25].

Handwriting and drawing are a physical way for human of expressing themselves, and therefore creative processes involving conscious and unconscious factors among self's motivations, emotions, mood, and temperament [22]. Thus, handwriting and drawing analyses may allow identifying both neurological psychological individual's characteristics.

Personality may be among these characteristics [13]. Personality is related to a pattern of relatively unique and individual's traits giving consistency and individuality to a person's behavior [5]. It includes thoughts, emotions, and attitudes, both innate and learned, influencing an individual's "interactions with, and adaptations to, the intrapsychic, physical, and social environments" [16]. Indeed, that topic is increasingly attracting researchers, as detecting individual's personality is ca way to predict individual behaviors and preferences [9] across different application's contexts from health care [8] to marketing [23].

Personal individual traits are commonly measured by validated questionnaires made up of items aimed to envisage respondents' attitudes and conducts.

The most used questionnaire to identify individual's personality traits is the "Big Five Questionnaire" (BFQ) [3, 19]. Such a questionnaire is based on the empirically "Five Factors Model" [6], representing the personality in terms of five dimensions, each consisting of two sub-dimensions, (see details in Section 2.2).

The Big Five Model has been exploited for different purposes, form preventive medicine [11] to recruitment [1].

The oldest and widespread method to identify and understand personality through handwriting analysis is graphology [14]. Graphology provides some hints for assessing individuals' personality by evaluating strokes, patterns and pressure applied while writing. Unfortunately, graphology is based on subjective non-standardized evaluations carried out by graphologists requiring be very costly and time-consuming analyses. On the other hand, modern drawing devices like graphics tablets and digital pens or web whiteboards can gather data to allow quantitative (objective) handwriting and drawing analyses. Gathered data consist of timing information associated with pen position (on a 2D space), pen inclination (two angles with respect to the paper) and pen state (drawing a stroke on the paper or in-air between two strokes). In this paper, we define a stroke as a continuous line drawn using the same pen state (on paper or in-air). Some devices enable to measure also the pressure applied on the paper. Such data has been effectively exploited for impairment detection and personality traits assessment [17, 18, 21] as well as for the detection of mood and emotional states [2].

The present study investigates the link between handwriting/drawing features and personality traits as measured by the "Big Five" questionnaire. A preliminary analysis of these data has been reported in [4], where it has been shown that the Big Five main dimensions describing personality can be significantly detected through the following set of computable handwriting features:

a) the time and number of strokes required to perform a handwriting/drawing task;
b) the pressure applied on the paper;
c) the space occupied by the strokes;
d) the inclination of handwriting strokes.

The results reported in the present study differs from those reported and discussed in [4] since it provides a deeper investigation assessing the capability of the above-mentioned features to characterize the 10 Big Five sub-dimensions.

2 Handwriting Analysis for the Detection of Personal Traits

2.1 Handwriting/Drawing Digital Analysis

Thanks to the recent development of new technological tools (digital scanners, touch screens, digital pens, and graphics tablets), handwriting analysis can be performed on a computerized platform. Computerized online handwriting/drawing analysis provides two main advantages: a) data collection of several invisible features, such as pen pressure, pen tilt, in air movements; b) data collection of timings, which can be exploited to dynamically analyze the entire handwriting/drawing process instead of a single static set of on a paper strokes representing the final handwriting/drawing task. These online measures had revealed themselves very useful to detect (through appropriately defined handwriting/drawing tasks) diseases, like dementia and /or brain stroke risk factors [15, 20, 22].

In this work, handwriting/drawing analysis was performed using the INTUOS WACOM Series 4 digital tablet and the Intuos Inkpen writing device. The tablet captures digitally signals describing the Inkpen movements performed by the subjects while writing/drawing an A4 plain paper placed on the tablet screen.

During each handwriting/drawing activity, the following data are acquired continuously at intervals of approximately 8 milliseconds (125 Hz):

1) Time in milliseconds;
2) Pen tip x-position;
3) Pen tip y-position;
4) Pen status (*in air*, coded as 0 or *on paper*, coded as 1);
5) Azimuth angle of the pen with respect to the plane of the writing surface;
6) Altitude angle of the pen with respect to the plane of the writing surface;
7) Pressure applied on the writing surface (only when the status is *on paper*).

Each acquisition is stored as one line in a text file; a comma separates the seven acquired values.

2.2 Personal Traits: The Big Five and Sub-Dimensions

Big Five questionnaire (BFQ) has identified five basic dimensions for character description and evaluation. Each of them considers two sub-dimensions. These dimensions do not represent specific theoretical views, but are derived from the analysis of natural language terms that people use to describe themselves and

others. Each sub-dimension is evaluated using 12 items, of which 6 are positive and 6 are negative, in order to control the desirability bias, i.e. "the tendency of research subjects to give socially desirable responses instead of choosing responses that are reflective of their true feelings" [7].

BFQ comprises also a Lie Scale (L) designed to evaluate subjects' propensity to provide false information. The scale L is composed of 12 items and involves ideal behaviors in society. Overall, BFQ comprises 132 items, each of which is answered through 5 point a Likert scale (1 = "very false for me" to 5 = "very true for me").

The Big Five main dimensions and the two sub-dimensions associated to each of them are described as follows:

1) *Openness to Experience*: refers to the ability and tendency of people to explore and create new experiences, showing curiosity, imagination, creativity, appreciation of art and out-of-the-box ideas. The two sub-dimensions are *Intellect* and *Openness*. Intellect reflects the tendency to participate in abstract and intellectual information, while Openness reflects the tendency to participate in aesthetic and sensory information (perception and imagination).

On this scale, people score according to a dichotomy: inventive/curious (higher score) and consistent/cautious (lower score).

2) *Conscientiousness*: refers to a person who is careful, thoughtful, responsible, organized, hardworking, accomplished, and persevering. The secondary dimensions of consciousness are *Industriousness* and *Orderliness*. *Industriousness* refers to the ability to engage in continuous, goal-oriented endeavors, which is related to productivity and professional ethics, while Orderliness refers to the tendency of arrangement, organization, and systemization, which is related to qualities such as cleanliness and diligence.

On this scale, people score according to a dichotomy: efficient/organized (higher scores) vs. easy-going/careless (lower scores)

3) *Extroversion*: refers to people who express positive emotions easily, are sociable, gregarious, decisive, talkative, and active. The two sub-dimensions are *Enthusiasm* and *Assertiveness*. Enthusiasm means friendliness, sociability, and a tendency to experience positive influences, while Assertiveness reflects a tendency toward agency, driving force, and social dominance

On this scale, people score according to a dichotomy: outgoing/energetic (higher scores) vs. solitary/reserved (lower scores).

4) *Agreeableness*: refers to a person who tends to be compassionate, trusting and helpful, flexible and tolerant. The sub-dimensions of agreeableness are *Compassion* and *Politeness*. Compassion refers to the tendency to care for others emotionally, while Politeness refers to the tendency to show good manners, observe social norms, and avoid aggression.

On this scale, people score according to a dichotomy: friendly/compassionate (higher scores) vs. alleging/detached (lower scores).

5) *Neuroticism*: Refers to people who lack stability and control of their emotions and are easily emotional, anxious, depressed, angry, worried, and insecure. The sub-dimensions of neuroticism are *Withdrawal* and *Volatility*. Withdrawal refers to the susceptibility to negative inward influences (inhibition), while Volatility refers to emotional instability, difficulty in controlling emotional impulses, and sensitivity to negative outward influences (disinhibition).

On this scale, people score according to a dichotomy: sensitive/nervous (higher scores) vs. secure/confident (lower scores).

3 The Design of the Experiment

To investigate the link among handwriting/drawing features and personality traits, in terms of Big Five and sub-dimensions, an experimental protocol has been devised. The protocol consisted in administering both the Big Five questionnaire and a series of handwriting and drawing tasks, through the above-mentioned digital tablet. Participants were properly informed about the protocol and signed informed consent before being enrolled in the experiment.

3.1 Subjects

A total of 78 participants (40 men and 38 women, mean age: 24.6, SD: 2.4) were recruited among the students of the University of Campania "Luigi Vanvitelli" in Caserta (southern Italy). They were first asked to fill in the Big Five questionnaire and then to perform the seven handwriting/drawing tasks described in Figure 1.

For each of the Big Five main dimensions and sub-dimensions, participants were divided into the group categories, “low”, “typical” and “high”, based on the BFQ scores obtained. In particular, subjects with BFQ scores less than 46 were classified as "low", those with a BFQ scores between 46 and 54 were classified as "typical" and those with BFQ scores greater than 54 were classified as "high".

Tables 1 and 2 show the distribution of participants according to the gender (FEMALE: F and MALE: M) and the three categories (LOW, TYPICAL and HIGH) defined above, for each Big Five (Table 1) and each sub-dimension (Table 2).

Table 1
Big Five main dimension and associated subjects' distributions with respect to their gender and group category (high, typical and low)

	High		Typical		Low	
	F	M	F	M	F	M
Extraversion	17	16	13	17	8	7
Agreeableness	12	18	8	14	18	8
Conscientiousness	16	16	5	14	7	10
Neuroticism	7	14	18	15	13	11
Openness to Experience	16	27	14	9	8	4

Table 2
Big Five sub-dimension and associated subject's distribution with respect to their gender and group category (high, typical and low)

		High		Typical		Low	
		F	M	F	M	F	M
Extraversion	Enthusiasm	16	17	14	12	8	11
	Assertiveness	17	19	15	12	6	9
Agreeableness	Compassion	16	16	15	15	7	9
	Politeness	11	16	11	14	16	10
Conscientiousness	Industriousness	20	18	15	13	3	9
	Orderliness	14	16	17	13	7	11
Neuroticism	Withdrawal	9	14	17	14	12	12
	Volatility	7	14	14	12	17	14
Openness to Experience	Intellect	20	27	12	7	6	6
	Openness	17	25	8	9	13	6

3.2 Tasks and Features

The handwriting/drawing tasks, performed by each subject were and acquired through the tablet were the following:

• Pentagons drawing;

• House drawing);

•*Four Italian words written in capital letters* (BIODEGRADABILE (biodegradable), FLIPSTRIM (flipstrim), SMINUZZAVANO (collapse), CHIUNQUE (anyone));

• *A series of loops with the left hand*;

• *A series of loops with the right hand*;

• *A clock drawing task* (with the hours and clock hands);

• The following *phonetically complete Italian sentence to be written in cursive letters*: I pazzi chiedono fiori viola, acqua da bere, tempo per sognare (crazy people are looking for purple flowers, drinking water, and time to dream).

Figure 1 illustrates the seven tasks as requested to each participant.

From the set of dynamic features recorded by the tablet and described in Section 2.1, it is possible to compute more features such as pen acceleration, velocity, instantaneous trajectory angle, instantaneous displacement, time characteristics, and ductus-based characteristics. In addition, in this article, we use the method defined in [2] to classify the strokes into three categories:

- **down strokes**, when the pen touches the surface. They are recorded with state 1 (since we are using an Inkpen, they appear on the paper and provide feedback to the subjects);
- **up strokes**, when the pen is close but does not touch the surface. They do not appear on the paper, but since the pen is close to the surface they are recorded with state 0 by the digital tablet;
- **idle strokes**, when the pen is far from the surface. The tablet does not recognize them, but they are still recognizable using time stamps.

In this paper, the following 17 handwriting/ drawing features are considered, grouped into five categories:

Pressure features, computed by measuring the pressure applied by the pen on the tablet surface during a specific task (only down strokes are considered):

1) P_{min}, minimum pressure recorded during a specific task;
2) P_{max}, maximum pressure recorded during a specific task;
3) P_{avg}, average pressure recorded during a specific task;
4) P_{sd}, standard deviation of pressures recorded during a specific task;

5) P_{10}, 10^{th} percentile. Assuming assume that the recorded pressures are sorted in ascending order, the 10^{th} percentile is the value that divides the data so 10% is below it;
6) P_{90}, 90^{th} percentile;

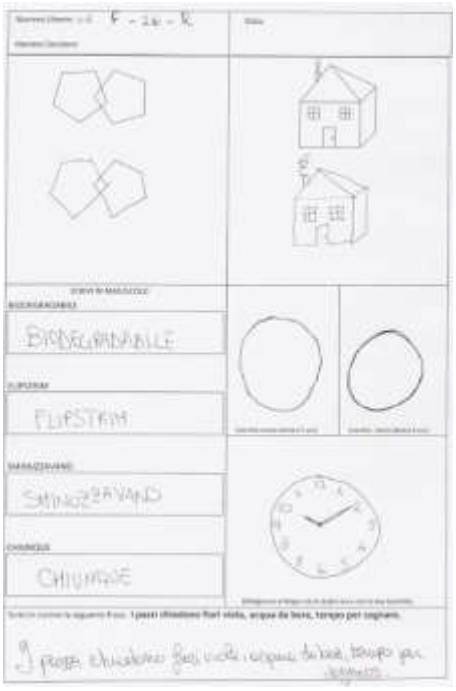

Figure 1
The seven handwriting/drawing tasks as they have been proposed to each participant

Ductus features, based on the number of strokes, in each pen status, performed during the task:

7) N_{up}, number of up strokes (the strokes that do not appear on the paper, but have been performed close to the surface);
8) N_{down}, number of down strokes (the strokes that appear on the paper);
9) N_{idle}, number of idle strokes (the number of times the pen is moved away from the paper/tablet, during a task).

Time features, based on the time spent, in each pen status, during the task:

10) T_{up}, time spent on up strokes;
11) T_{down}, time spend on down strokes;
12) T_{idle}, time spent in the idle status;
13) T_{total}, total time elapsed for the task ($T_{total}= T_{up} + T_{down} + T_{idle}$).

Space features, based on the area used by the strokes:

14) S_{bb}, sum of spaces used by on paper strokes. For each stroke, we compute the smallest axis aligned box containing the stroke and sum its area;
15) S_{avg}, average Euclidean distance between consecutive down strokes;
16) S_{total}, sum of the Euclidean distance between consecutive down strokes.

Inclination features, based on the inclination of strokes:

17) I_{avg}, average inclination of down strokes.

3.2 Results on the Main Big Five Dimension

Table 3 below provides a summary of the significant results observed for the main Big Five dimensions and discussed in detail in [4].

In this work, we provide a more detailed description and a deeper analysis of each Big Five sub-dimension. To this end, several repeated measures analysis of variance (ANOVAs) have been performed for each feature (divided into 5 categories), each task, and each Big Five sub-dimensions. It is worth mentioning that a single repeated measures analysis of variance was performed on the Inclination category, because it contains a single feature. The Bonferroni post-hoc test was performed with a significance equal to $\alpha = .05$.

Table 3
For each dimension, the table provides the task and the corresponding handwriting/drawing features for which it has been observed a significant difference among groups

Big Five Dimension	Task	Features
Extraversion	Loops with the left hand	T_{down}, T_{total}
	Loops with the right hand	T_{down}, T_{total}
Agreeableness	Clock drawing	N_{idle}
	House drawing	T_{total}, N_{up}, N_{down}, N_{idle}
	Writing in cursive letters	S_{bb}, N_{up}, N_{down}, N_{idle} s
	Pentagons drawing	N_{up}, N_{down}, N_{idle}
Conscientiousness	Loops with the right hand	T_{down}, T_{total}
	Pentagons drawing	N_{up}, N_{down}, N_{idle}
	House drawing	N_{up}, N_{down}, N_{idle}
	Clock drawing	N_{up}, N_{down}, N_{idle}
Openness to Experience	Writing in capital letters	P_{max}, P_{sd}, P_{10}, P_{90}, N_{up}, N_{down}, N_{idle}, S_{bb}, S_{avg}
	Loops with the left hand	T_{down}, T_{total}
	Loops with the right hand	T_{down}, T_{total}
	Writing in cursive letters	N_{up}, N_{down}, N_{idle}
Neuroticism	House drawing	S_{bb}, S_{avg}, S_{total},

4 Results

In the following, we provide the significant results obtained, grouped by features categories. A summary of these results is also provided in the following Table 4.

4.1 Pressure Features

- **Extraversion -> Assertiveness:**

a) Task 5: A series of loops with the right hand. There are significant differences between the three Assertiveness group categories [$F(2,72)<3.636$; $p=0.031$]. Bonferroni post-hoc tests showed a significant difference ($p=0.030$) between the score of participants in the *typical* group (M=559.830; SD=18.995) and the score of participants in the *low* group (M=473.367; SD=25.849). In particular, this difference is due to the scores of the following features:

  - P_{max}: The score of the *low* group (M=786.13; SD=161.43) is significantly different ($p=0.047$) with respect to the score of the *typical* group (M=887.74; SD=121.03);
  - P_{avg}: The score of the *low* group (M=645.08; SD=189.84) is significantly different ($p=0.026$) with respect to the score of the *typical* group (M=768.85; SD=141.13);
  - P_{90}: The score of the *low* group (M=732.47; SD=174.15) is significantly different ($p=0.033$) with respect to the score of the *typical* group (M=849.56; SD=135.68).

4.2 Ductus Features

- **Agreeableness -> Politeness:**

a) Task 6: A clock drawing task. There are significant differences between the three Politeness group categories [$F(2,72)<4.435$; $p=0.015$]. Bonferroni post-hoc tests showed a significant difference ($p=0.004$) between the score of participants in the *low* group (M=21.37; SD=1.63) and the score of participants in the *high* group (M=28.16; SD=1.59). In particular, this difference is due to the scores of the following features:

  - N_{up}: The score of the *high* group (M=28.99; SD=1.26) is significantly different ($p=0.043$) with respect to the score of the *low* group (M=24.41; SD=1.30) and significantly different ($p=0.014$) with respect to the score of the *typical* group (M=23.67; SD=1.30);
  - N_{down}: The score of the *high* group (M=29.63; SD=1.28) is significantly different ($p=0.043$) with respect to the score of the *low* group (M=25.01; SD=1.32) and significantly different ($p=0.014$) with respect to the score of the *typical* group (M=24.24; SD=1.32);

- N_{idle}: The score of the *low* group (M=14.68; SD=3.16) is significantly different (p=0.041) with respect to the score of the *high* group (M=25.84; SD=3.07) and significantly different (p=0.016) with respect to the score of the *typical* group (M=27.56; SD=3.16).

4.3 Time Features

- **Extraversion -> Enthusiasm:**

a) Task 5: A series of loops with the right hand. There are significant differences between the three Enthusiasm group categories [F(2,72)<3.831; p=0.026]. Bonferroni post-hoc tests showed a significant difference (p=0.023) between the score of participants in the *low* group (M=3655.22; SD=264.46) and the score of participants in the *high* group (M=4909.44; SD=273.20). In particular, this difference is due to the scores of the following features:

- T_{down}: The score of the *low* group (M=6970.42; SD=3614.54) is significantly different (p=0.006) with respect to the score of the *high* group (M=9773.61; SD=2638.94);
- T_{total}: The score of the *low* group (M=7314.53; SD=3811.99) is significantly different (p=0.022) with respect to the score of the *high* group (M=9793.00; SD=2653.03).

- **Agreeableness -> Compassion:**

a) Task 3: Four Italian words in capital letters. There are significant differences between the three Compassion group categories [F(2,72)<4.800; p=0.011]. Bonferroni post-hoc tests showed a significant difference (p=0.009) between the score of participants in the *low* group (M=19350.45; SD=1057.52) and the score of participants in the *typical* group (M=15329.87; SD=766.24). In particular, this difference is due to the scores of the following features:

- T_{up}: The score of the *low* group (M=16458.50; SD=6911.05) is significantly different (p=0.044) with respect to the score of the *typical* group (M=12625.13; SD=4025.33);
- T_{idle}: The score of the *low* group (M=5986.37; SD=6388.06) is significantly different (p=0.036) with respect to the score of the *typical* group (M=3406.73; SD=1734.23);
- T_{total}: The score of the *low* group (M=39417.57; SD=14216.22) is significantly different (p=0.009) with respect to the score of the *typical* group (M=28166.33; SD=4273.03).

- **Agreeableness -> Politeness:**

a) Task 4: A series of loops with the left hand. There are significant differences between the three Politeness group categories [$F(2,72)<4.392$; $p=0.016$]. Bonferroni post-hoc tests showed a significant difference ($p=0.022$) between the score of participants in the *low* group ($M=4113.86$; $SD=400.60$) and the score of participants in the *high* group ($M=5680.64$; $SD=389.23$). In particular, this difference is due to the scores of the following features:

   - T_{down}: The score of the *low* group ($M=7681.88$; $SD=3855.59$) is significantly different ($p=0.020$) with respect to the score of the *high* group ($M=10282.04$; $SD=3306.64$) and significantly different ($p=0.023$) with respect to the score of the *typical* group ($M=10513.24$; $SD=3901.02$);
   - T_{total}: The score of the *low* group ($M=8102.42$; $SD=4063.94$) is significantly different ($p=0.019$) with respect to the score of the *high* group ($M=11120.26$; $SD=3895.06$).

- Task 5: A series of loops with the left hand. There are significant differences between the three Politeness group categories [$F(2,72)<5.972$; $p=0.004$]. Bonferroni post-hoc tests showed a significant difference ($p=0.009$) between the score of participants in the *low* group ($M=3637.38$; $SD=302.70$) and the score of participants in the *typical* group ($M=5005.07$; $SD=302.55$). In particular, this difference is due to the scores of the following features:

   - T_{down}: The score of the *low* group ($M=7099.92$; $SD=3712.54$) is significantly different ($p=0.012$) with respect to the score of the *high* group ($M=9302.37$; $SD=2690.20$) and significantly different ($p=0.003$) with respect to the score of the *typical* group ($M=9876.94$; $SD=2651.90$);
   - T_{total}: The score of the *low* group ($M=7320.81$; $SD=3842.68$) is significantly different ($p=0.022$) with respect to the score of the *high* group ($M=9326.07$; $SD=2711.47$) and significantly different ($p=0.006$) with respect to the score of the *typical* group ($M=9907.84$; $SD=2625.55$).

- Task 6: A clock drawing task. There are significant differences between the three Politeness group categories [$F(2,72)<3,879$; $p=0.025$]. Bonferroni post-hoc tests showed a significant difference ($p=0.005$) between the score of participants in the *low* group ($M=13782.99$; $SD=1500.79$) and the score of participants in the *high* group ($M=19468.68$; $SD=1458.21$). In particular, this difference is due to the scores of the following features:

   - T_{up}: The score of the *low* group ($M=11498.73$; $SD=4468.06$) is significantly different ($p=0.026$) with respect to the score of the *high* group ($M=16926.52$; $SD=7929.66$);
   - T_{total}: The score of the *low* group ($M=26440.96$; $SD=11594.21$) is significantly different ($p=0.025$) to the score of the *high* group ($M=39754.11$; $SD=20125.58$).

- **Conscientiousness -> Orderliness:**

a) Task 3: Four Italian words in capital letters. There are significant differences between the three Orderliness group categories [$F(2,72)<4.117$; $p=0.020$]. Bonferroni post-hoc tests showed: i) a significant difference ($p=0.028$) between the score of participants in the *low* group ($M=19118.55$; $SD=1026.73$) and the score of participants in the *high* group ($M=15860.77$; $SD=777.14$); ii) a significant difference ($p=0.016$) between the score of participants in the *low* group ($M=19118.55$; $SD=1026.73$) and the score of participants in the *typical* group ($M=15696.32$; $SD=782.40$). In particular, this difference is due to the scores of the following feature:

  - T_{total}: The score of the *low* group ($M=38550.22$; $SD=12719.49$) is significantly different ($p=0.041$) with respect to the score of the *high* group ($M=31792.77$; $SD=5250.42$) and significantly different ($p=0.030$) with respect to the score of the *typical* group ($M=31267.60$; $SD=7741.35$).

- **Conscientiousness -> Industriousness:**

a) Task 3: Four Italian words in capital letters. There are significant differences between the three Industriousness group categories [$F(2,72)<4.117$; $p=0.020$]. Bonferroni post-hoc tests showed: i) a significant difference ($p=0.017$) between the score of participants in the *low* group ($M=21077.16$; $SD=1379.89$) and the score of participants in the *high* group ($M=15874.71$; $SD=672.47$); ii) a significant difference ($p=0.046$) between the score of participants in the *low* group ($M=21077.16$; $SD=1379.89$) and the score of participants in the *typical* group ($M=16287.81$; $SD=784.32$). In particular, this difference is due to the scores of the following features:

  - T_{idle}: The score of the *low* group ($M=6587.75$; $SD=7220.71$) is significantly different ($p=0.001$) with respect to the score of the *high* group ($M=3838.71$; $SD=2372.66$) and significantly different ($p=0.002$) with respect to the score of the *typical* group ($M=3873.04$; $SD=2163.73$);
  - T_{total}: The score of the *low* group ($M=39513.25$; $SD=12697.87$) is significantly different ($p=0.003$) with respect to the score of the *high* group ($M=31671.74$; $SD=6010.51$) and significantly different ($p=0.011$) with respect to the score of the *typical* group ($M=32429.64$; $SD=9190.34$).

b) Task 4: A series of loops with the left hand. There are significant differences between the three Industriousness group categories [$F(2,72)<7.475$; $p=0.001$]. Bonferroni post-hoc tests showed a significant difference ($p=0.006$) between the score of participants in the *low* group ($M=2737.72$; $SD=650.66$) and the score of participants in the *high* group ($M=5531.10$; $SD=317.09$). In particular, this difference is due to the scores of the following features:

- T_{down}: The score of the *low* group (M=6753.92; SD=4022.90) is significantly different (p=0.001) with respect to the score of the *high* group (M=10393.47; SD=3585.75) and significantly different (p=0.014) with respect to the score of the *typical* group (M=9434.86; SD=3724.66);
- T_{total}: The score of the *low* group (M=6905.50; SD=4168.29) is significantly different (p=0.001) with respect to the score of the *high* group (M=11080.37; SD=4061.33) and significantly different (p=0.016) with respect to the score of the *typical* group (M=9796.93; SD=3741.61).

c) Task 5: A series of loops with the right hand. There are significant differences between the three Industriousness group categories [$F(2,72)<11.041$; $p<<0.001$]. Bonferroni post-hoc tests showed: i) a significant difference (p=0.001) between the score of participants in the *low* group (M=2344.45; SD=471.93) and the score of participants in the *high* group (M=4805.04; SD=229.99); ii) a significant difference (p=0.013) between the score of participants in the *low* group (M=2344.45; SD=471.93) and the score of participants in the *typical* group (M=4447.34; SD=268.24). In particular, this difference is due to the scores of the following features:

- T_{down}: The score of the *low* group (M=5973.17; SD=3301.74) is significantly different ($p<<0.001$) with respect to the score of the *high* group (M=9655.47; SD=2954.48) and significantly different (p=0.001) with respect to the score of the *typical* group (M=8717.04; SD=3016.30);
- T_{total}: The score of the *low* group (M=6039.42; SD=3340.61) is significantly different ($p<<0.001$) with respect to the score of the *high* group (M=9672.32; SD=2966.07) and significantly different (p=0.001) with respect to the score of the *typical* group (M=8922.14; SD=3053.51).

- **Openness to Experience -> Intellect:**

a) Task 4: A series of loops with the left hand. There are significant differences between the three Intellect group categories [$F(2,72)<6.114$; p=0.004]. Bonferroni post-hoc tests showed: i) a significant difference (p=0.004) between the score of participants in the *low* group (M=3148.16; SD=572.72) and the score of participants in the *high* group (M=5335.02; SD=292.65); ii) a significant difference (p=0.010) between the score of participants in the *low* group (M=3148.16; SD=572.72) and the score of participants in the *typical* group (M=5322.08; SD=471.78). In particular, this difference is due to the scores of the following features:

- T_{down}: The score of the *low* group (M=5878.83; SD=3311.66) is significantly different (p=0.002) with respect to the score of the *high* group (M=9932.30; SD=3304.53) and significantly different (p=0.002) with respect to the score of the *typical* group (M=10674.26; SD=4307.50);

- T_{total}: The score of the *low* group (M=6296.33; SD=3675.26) is significantly different (p=0.003) with respect to the score of the *high* group (M=10608.87; SD=3774.48) and significantly different (p=0.014) with respect to the score of the *typical* group (M=10740.05; SD=4368.18).

b) Task 5: A series of loops with the right hand. There are significant differences between the three Intellect group categories [F(2,72)<8,627;p<<0.001]. Bonferroni post-hoc tests showed: i) a significant difference (p=0.001) between the score of participants in the *low* group (M=2846.81; SD=424.14) and the score of participants in the *high* group (M=4802.80; SD=216.73); ii) a significant difference (p=0.004) between the score of participants in the *low* group (M=2846.81; SD=424.14) and the score of participants in the *typical* group (M=4647.25; SD=349.39). In particular, this difference is due to the scores of the following features:

- T_{down}: The score of the *low* group (M=5627.92; SD=3181.35) is significantly different (p<<0.001) with respect to the score of the *high* group (M=9305.30; SD=2862.221) and significantly different (p=0.004) with respect to the score of the *typical* group (M=9356.79; SD=3211.24);
- T_{total}: The score of the *low* group (M=5694.17; SD=3229.410) is significantly different (p<<0.001) with respect to the score of the *high* group (M=9441.11; SD=2872.27) and significantly different (p=0.005) with respect to the score of the *typical* group (M=9356.79; SD=3211.24).

4.4 Space Features

- **Agreeableness -> Politeness:**

a) Task 5: A series of loops with the right hand. There are significant differences between the three Politeness group categories [F(2,72)<3.557; p=0.034]. Bonferroni post-hoc tests showed a significant difference (p=0.030) between the score of participants in the *low* group (M=3913.62; SD=657.32) and the score of participants in the *typical* group (M=6344.11; SD=656.99). In particular, this difference is due to the scores of the following feature:

- S_{bb}: The score of the *low* group (M=11832.27; SD=7476.26) is significantly different (p=0.027) to the score of the *typical* group (M=19218.80; SD=11845.08).

- **Conscientiousness -> Orderliness:**

a) Task 3: Four Italian words in capital letters. There are significant differences between the three Orderliness group categories [F(2,72)<6.183; p=0.003]. Bonferroni post-hoc tests showed a significant difference (p=0.003) between the score of participants in the *low* group (M=112500.147; SD=5463.13) and

the score of participants in the *high* group (M=88811.80; SD=4135.11). In particular, this difference is due to the scores of the following features:

- S_{bb}: The score of the *low* group (M=255313.83; SD=64520.85) is significantly different (p=0.003) to the score of the *high* group (M=201763.87; SD=56429.60);
- S_{total}: The score of the *low* group (M=79326.67; SD=19118.84) is significantly different (p=0.009) to the score of the *high* group (M=63650.60; SD=14186.71);
- S_{avg}: The score of the *high* group (M=1062.40; SD=216.63) is significantly different (p=0.018) with respect to the score of the *low* group (M=1262.17; SD=244.79) and significantly different (p=0.015) with respect to the score of the *typical* group (M=1244.37; SD=260.77).

Table 4

For each Big Five sub-dimension, the table provides the dimension, the task and the corresponding features for which it has been observed a significant difference among groups

Big Five Sub-dimension	**Big Five Dimension**	**Task**	**Features**
Enthusiasm	Extraversion	Loops with right hand	T_{down}, T_{total}
Assertiveness	Extraversion	Loops with right hand	P_{max}, P_{avg}, P_{90}
Compassion	Agreeableness	Writing in capital letters	T_{up}, T_{idle}, T_{total}
Politeness	Agreeableness	Clock drawing	N_{up}, N_{down}, N_{idle}, T_{up}, T_{total}
	Agreeableness	Loops with left hand	T_{down}, T_{total}
	Agreeableness	Loops with right hand	T_{down}, T_{total}, S_{bb}
Industriousness	Conscientiousness	Writing in capital letters	T_{idle}, T_{total}
	Conscientiousness	Loops with left hand	T_{down}, T_{total}
	Conscientiousness	Loops with right hand	T_{down}, T_{total}
Orderliness	Conscientiousness	Writing in capital letters	T_{total}, S_{bb}, S_{total}, S_{avg}
Volatility	Neuroticism	Copy of a house drawing	I_{avg}
	Neuroticism	Clock drawing	I_{avg}
Intellect	Openness to Experience	Loops with left hand	T_{down}, T_{total}
	Openness to Experience	Loops with right hand	T_{down}, T_{total}

4.5 Inclination Feature

Being the Inclination category composed of a single feature (I_{avg}), a single repeated measurement ANOVA was performed considering the results of all the tasks.

- **Neuroticism -> Volatility:**

There are significant differences between the three Volatility group categories [F(2,71)<5,047; p=0.009]. Bonferroni post-hoc tests showed a significant difference (p=0.003) between the score of participants in the *low* group (M=1.75; SD=0.009) and the score of participants in the *typical* group (M=2.20; SD=0.10). In particular, this difference is due to the scores in the following tasks:

- Task 2: A copy of a house drawing. The score of the *low* group (M=3.48; SD=1.75) is significantly different (p=0.020) to the score of the *typical* group (M=4.95; SD=2.61);
- Task 6: A clock drawing task. The score of the *low* group (M=1.12; SD=0.56) is significantly different (p=0.011) to the score of the *high* group (M=1.92; SD=1.45).

Conclusion

The results of the proposed statistical analyses show that online handwriting and drawing features collected through appropriately defined tasks can be used to distinguish among different personality dimensions and sub-dimensions of the Big Five Questionnaire. With more details, the analysis carried out in this paper and in [4] shows that the considered features discriminate among 4 out of the 5 Big Five dimensions and 8 out of the 10 Big-Five sub-dimensions. Table 4synthesizes the features for which significant differences are observed allowing categorizing participants with different personalities. A comparison among tasks shows that although all the envisaged tasks have proven to be useful, the simpler ones (such as loops or writing in capital letters) are more informative about subject's personalities. Finally, comparing the different types of features, it can be observed that timing ductus and pressure features are the more promising and may deserve deeper investigation.

Acknowledgement

The research leading to these results has been funded by the EU Horizon 2020 Research and Innovation Program under funding agreements N. 769872 (EMPATHIC) and N. 823907 (MENHIR). Partial support comes also from the Italian project SIROBOTICS, funded by MIUR, PNR 2015-2020, DD 1735 13/07/2017 and the Italian project ANDROIDS, Università della Campania "Luigi Vanvitelli" programma V:ALERE 2019, DR 906, 04/10/2019, n. 157264.

References

[1] J. Camps, J. Stouten, and M. Euwema: The Relation Between Supervisors' Big Five Personality Traits and Employees' Experiences of Abusive Supervision. In Frontiers in Psychology, 7, pp. 1-11, 2016

[2] G. Cordasco, F. Scibelli, M. Faundez-Zanuy, L. Likforman-Sulem, and A. Esposito: Handwriting and Drawing Features for Detecting Negative Mood. In WIRN 2017: Quantifying and Processing Biomedical and Behavioral Signals, Smart Innovation, Systems and Technologies, Vol. 103, pp. 73-86, 2017

[3] C. G. De Young, L. C. Quilty, and J. B. Peterson: Between facets and domains: 10 aspects of the Big Five. In Journal of Personality and Social Psychology, 93 (5), pp. 880-896, 2007

[4] A. Esposito, T. Amorese, M. Buonanno, M. Cuciniello A. M. Esposito, M. Faundez-Zanuy, L. Likforman-Sulem, M. T. Riviello, A. Troncone and G. Cordasco: Handwriting and Drawing Features for Detecting Personality Traits. In IEEE 10th International Conference on Cognitive Infocommunications, 2019

[5] J. Feist and G. J. Feist: Theories of personality, Boston: McGraw-Hill, 2009

[6] L. R. Goldberg: Personality structure: emergence of the five factors model. In Annual Review of Psychology, 41, pp. 417-440, 1990

[7] P. Grimm: Social desirability bias. Wiley international encyclopedia of marketing. https://doi.org/10.1002/9781444316568.wiem02057, 2010

[8] A. Hajek, J. O. Bock, and H. H. Knig: The role of personality in health care use: Results of a population-based longitudinal study in Germany. In PloS one, 12(7), 2017

[9] D. Hassabis, R. N. Spreng, A. A. Rusu, C. A. Robbins, R. A. Mar, D. L. Schacter: Imagine all the people: how the brain creates and uses personality models to predict behavior. Cerebral cortex, 24(8), (2014)

[10] D. Impedovo, G. Pirlo: Dynamic Handwriting Analysis for the Assessment of Neurodegenerative Diseases: A Pattern Recognition Perspective. In IEEE Reviews in Biomedical Engineering, 12, pp. 209-220, 2019

[11] S. Israel, T. E. Moffitt, D. W. Belsky, R. J. Hancox, R. Poulton, B. Roberts, W. Murray Thomson, A. Caspi: Translating personality psychology to help personalize preventive medicine for young adult patients. In J. Pers. Soc. Psychol., 106, pp. 484-498, 2014

[12] K. H. James, R. J. Jao, V. Berninger: The development of multileveled writing systems of the brain: Brain lessons for writing instruction. In MacArthur, C. A., Graham, S., Fitzgerald, J. (Eds.), Handbook of writing research 2nd ed., pp. 116-129, 2016

[13] P. M. Joshi, A. Agarwal, A. Dhavale, R. Suryavanshi, S. Kodolikar: Handwriting Analysis for Detection of Personality Traits using Machine Learning Approach. In International Journal of Computer Applications 130, pp. 40-45, 2015

[14] S. Kedar, V. Nair and S. Kulkarni: Personality Identification through Handwriting Analysis: A Review. In International Journal of Advanced Research in Computer Science and Software Engineering, 5(1), 2015

[15] H. Kim: The clockme system: Computer-assisted screening tool for dementia. Ph.D. dissertation, College Computing, Georgia Inst. Technol., 2013

[16] R. R. Larsen and D. M. Buss: Personality Psychology: Domains of Knowledge About Human Nature, 3rd edition, McGraw-Hill Inc. 2008

[17] T. Li-Ping Tang: Detecting honest people's lies in handwriting. In Journal of Business Ethics, 106(4), pp-389-400, 2012

[18] L. Likforman-Sulem, A. Esposito, M. Faundez-Zanuy, S. Clmenon, and G. Cordasco: Emothaw: A novel database for emotional state recognition from handwriting and drawing. In IEEE Transactions on Human-Machine Systems, 2016

[19] R. R. McCrae and O. P. John: An Introduction to the Five-Factor Model and Its Applications. In Journal of Personality, 60(2), pp. 175-215, 1992

[20] C. O'Reilly and R. Plamondon: Design of a neuromuscular disorders diagnostic system using human movement analysis. In 11th International Conference on Information Science, Signal Processing and their Applications, ISSPA 2012, pp. 787-792, 2012

[21] R. Plamondon and S. N. Srihari: On-line and off-line handwriting recognition: A comprehensive survey. In IEEE Trans. Pattern Anal. Mach. Intell, 22(1), pp. 63-84, 2000

[22] R. Plamondon, C. O'Reilly, and C. Ouellet-Plamondo: Strokes against stroke - strokes for strides. In Pattern Recognition, 47(3) pp. 929-944, 2014

[23] J. Rawat and B. Mann: Role of consumer personality and involvement in understanding customer experience. In Innovative Marketing, 12(3), pp. 19-33, 2016

[24] M. Schiegg, D. Thorpe: Historical Analyses of Disordered Handwriting: Perspectives on Early 20th-Century Material from a German Psychiatric Hospital. Written Communication, 34(1), pp. 30-53, 2017

[25] L. Trojano, G. Gainotti: Drawing Disorders in Alzheimer's Disease and Other Forms of Dementia. Journal of Alzheimer's disease, JAD, 53(1), pp. 31-52, 2016