Excessive Self-Presentation on Facebook
One Year of Analysis of Online Posting

Andrea Guazzini*, Cristina Cecchini†, Elisa Guidi‡, Monica Milani‡, Patrizia Meringolo*†

*†Department of Education and Psychology, University of Florence, Via di San Salvi, 12, Building 26, 50135 Florence, Italy
‡Department of Information Engineering, University of Florence, Via di San Marta, 3, 50139 Florence, Italy
†Department of Informatics and Systems, University of Florence, Via di San Marta, 3, 50139 Florence, Italy
‡Centre for the Study of Complex Dynamics (CSDC), University of Florence, Via Sansone, 1, 50019 Sesto Fiorentino, Florence, Italy

*andrea.guazzini@unifi.it; †cristina.cecchini@unifi.it; ‡elisa.guidi@unifi.it; ‡monica.milani@istitutogk.it; †patrizia.meringolo@unifi.it

Abstract—Facebook and social networks in general have exploded in popularity in the last several years, becoming a social institution for teenagers, who use it for self-presentation and as a fundamental tool to project their personal identity and manage social relationships. Taking the opportunity to reveal psychological features by analyzing personal profiles, this study examined the relationship between linguistic style, self-presentation, and other activities on Facebook over the course of one year. Through the linguistic analysis of 50 adolescents’ profiles, a new model titled “The Excessive Online Self-Presentation Model” was developed by considering only the contents of their public posts. Results showed how excessive self-presentation consisted of disclosing personal information, and with a higher number of words per post. Moreover, it was discovered that the frequent use of sexual language was associated with attention-seeking behaviors. The findings of this study were in line with prior research on the content differences in online self-presentation, confirming how linguistic analysis can be a useful tool to analyze cognitive dynamics within virtual environments, and providing a new assessment model to reveal fundamental psychological characteristics of adolescent online behavior.

Keywords—Facebook Use; Self-Presentation; Narcissism; Adolescent; Linguistic Inquiry and Word Count

I. INTRODUCTION

A. Self-Presentation on Social Networks and Adolescent Users

Social network sites (SNSs) are web tools that allow the self-presentation of users on public or private personal internet pages and facilitate interaction among users of the same social network [1].

The opportunity to share personal content in a mass communication highlights the implementation of specific self-presentation strategies in the presence of multiple audiences [1, 2].

Self-presentation, according to Goffman’s theory of identity and social performance, is a process of impression management [3]. Recently, Kim, et al. defined online self-presentation as a strategy for people to manage and introduce themselves to a virtual community [4]. Some studies suggest that self-expression and impression management are the most relevant reasons people create personal profiles [5, 6]. SNSs allow users to control their personal webpages for self-presentation by updating profile information, and posting photos and messages [7, 8].

Recent literature has repeatedly focused on online self-presentation, and a number of studies have highlighted the fact that people tend to disclose more personal information on SNSs compared to other means of communication [9]. Researchers began investigating the origins of excessive online self-presentation, wondering whether frequent posting might be a sign of narcissism [10-13]. As suggested by Kosinski, Stillwell, and Graepel [14], personality traits are identifiable from digital records of human behavior [14]. Researchers studying narcissism have generally emphasized a positive relation with social networks usage; studies have found that the narcissistic trait is related to an excessive amount of activity on Facebook, including connecting to a large number of friends and frequently publishing posts and photos [15, 16]. Furthermore, narcissistic people spend at least one hour a day on Facebook and use more self-promoting photos than other users [6].

Interestingly, there are sex-related differences in self-presenting behaviors, such as the fact that men appear to disclose more basic, impersonal information (i.e., sports) and more contact information than females on Facebook [17]. Moreover, it appears that men and women use SNSs for different reasons; women seem to self-present with more personal photos than males, while men publish more brand-related posts [18-22]. There are conflicting beliefs regarding gender-specific narcissism on social networks. Some studies suggest males are more narcissistic and post more personal photos on Facebook [10, 11, 23, 24]. On the contrary, other research shows females tend to be more narcissistic by using Facebook more frequently and displaying more self-promoting photos [6, 7, 25].
In a short period of time, Facebook has become the most popular SNS, especially among adolescents, with an average of over 936 million daily users in March 2015 [26-28]. Thus, the proliferation of SNSs has distinctly changed communication between people, the expression of emotions, the broadcast of information, and “socio-cultural revolution” mainly concerning adolescents, who were born in an “open cyber-world” [15, 29, 30]. With the ability to share information on personal pages and self-present, SNSs allow people to build an online social identity, which is crucial in adolescence [5, 7].

Previous research has demonstrated that adolescents use Facebook for identity building and expansion [27, 31]. Also, they appreciate the ability to share basic information and connect with people. In another study on teenage online habits, Madden, et al. reported that boys and girls post similar content, but boys tend to share their own phone number on SNSs and have more public profiles than girls [32].

Despite these interesting data, few studies have investigated the self-presentational style on SNSs in adolescents [15]. Because this demographic frequently uses Facebook and is more likely to publish posts with the purpose of building their identity, the concern is that SNSs may reinforce, or even create, narcissistic tendencies [7]. Furthermore, excessive online self-presentation might be misused, as in the context of cyberbullying, or have a negative effect on a user’s reputation, future education, and job opportunities [33, 34].

B. Self-Presentation and Language Use on Facebook

Researchers have frequently analyzed linguistic style and use of words to better understand human psychology [35, 36]. For example, some studies have shown that agreeable people use less swear words and negative emotion words, while extroverts use more personal pronouns, females low in conscientiousness use more second person pronouns, and finally those high in conscientiousness use less swear words [35, 36].

Literature shows that specific word categories can disclose a surprising amount of information about people and their behavior. Studies on text analysis suggest that function words (i.e., pronouns, prepositions, articles, and auxiliary verbs) appear to underline an attentional focus [37, 38]. Particularly, the use of first-person singular pronouns suggests a focus on one’s self, while second- and third-person pronouns indicate social engagement or awareness [39]. Another study revealed that function words such as first-person singular pronouns, short words, and conditional, present tense verbs, and articles can reflect psychological intimacy and an engagement in social interactions [40]. Since the analysis of post content is relevant to understanding the connection between activity on SNSs and self-presentation style, recent research has investigated the relationship between peculiarities in linguistic style and self-presentation on SNSs [41, 42].

In an analysis of self-presentation management, Bazarova, et al. found differences in the expression of negative emotions in status updates compared to wall posts or private messages [42]. It seems positive emotion words are correlated with self-presentational concerns in status updates, which suggests a specific strategy in sharing positive emotions through public communication. In another study that focused on linguistic styles on Facebook and its relation to gender [43], females appear to use more emotional words (e.g., ‘excited’), first-person singular pronouns, and psychological and social processes (e.g., ‘love you’), while males tend to use more swear words, and object references (e.g., ‘xbox’). Additional research has focused on linguistic markers of narcissism on SNSs. For example, DeWall, et al. investigated how narcissists spread information about themselves on social media [44]. Across two studies, narcissistic individuals who did not use first-person singular pronouns, and assumed implicit markers of narcissistic self-focus, compensated by publishing provocative online photos or using more profanity and verbal aggression in their self-presentation. Another study revealed an association between narcissism and the use of angry words, swear words, and sexual language [24].

Since previous research has provided little information on the relationship between linguistic style and frequency, and typology of activity on Facebook, this study approached the theme of online self-presentation using the Linguistic Inquiry and Word Count (LIWC) text analysis program to explore the correlation between narcissism and Facebook habits [45].

C. Aims

The purpose of this study was to analyze the self-presentation styles of 50 Italian adolescents on Facebook. Thus, a new model titled “The Excessive Online Self-Presentation Model” was defined to directly assess this tendency through the linguistic analysis of public contents of personal pages.

In particular, this study intended to accomplish the following:

1. Verify the relationship between linguistic style, self-presentation, and activity on Facebook.
2. Provide an operative framework to formalize, compare, and merge existing narcissistic models.
3. Determine which LIWC dimensions could be used to evaluate excessive online self-presentation and possible underlying gender differences.
4. Define a more effective model to automatically assess excessive online self-presentation using only public linguistic dimensions.
Within these contexts, the following specific hypotheses were considered:

H1: Adolescents who publish more wall posts have a higher level of activity on Facebook and receive a higher number of likes and comments.

H2: Different linguistic styles affect the kind of published posts on Facebook and the number of likes and comments received.

H3: Excessive self-presentation is associated with a higher level of activity on Facebook and a higher number of likes and comments received.

H4: Excessive self-presentation is related to the publishing of longer posts and the expression of negative emotions and sexual language.

H5: Excessive self-presentation reveals no gender differences but shows different linguistic styles between males and females.

II. METHOD

A. Participants

Fifty adolescents (50% female, 50% male; Aged M=16.95, SD=1.08) participated in this study. All participants had an active Facebook account. The subjects were recruited from a Tuscan high school, and were involved in the “ARCA Project” patronized by a Tuscan municipality.

The agreement of the school to participate was obtained from the head of the Institute. Teachers and a researcher explained the aim of the survey and confidentiality issues to the students. Parental and adolescent consent was obtained. The participants consisted only of Italian mother tongue citizens in order to conduct an accurate linguistic analysis. All subjects were involved as volunteers and could withdraw from the research at any time. To be included in the study, each participant had to meet the following requirements:

- Between 14-19 years old (i.e., age range of Italian high school students);
- Owned a Facebook profile for at least one year.

Hence, the exclusion criteria were as follow: students outside the specified age range, and having a recent Facebook profile, which could suggest an inexperienced Facebook user. A total of 50 students were contacted for the study, and all decided to participate for an entire year of data collection.

B. Design and Procedure

To analyze the Facebook usage on the participants’ pages, a Facebook account was created with a research logo as the profile picture. When a participant’s profile was located, a message was sent inviting him/her to participate in the study, and he/she was befriended. Participants were advised that the researchers would have access to their Facebook pages, and the data would be stored anonymously.

Data collection took place over a period of one year. Socio-demographic features of the participants were collected by taking only those provided on their public Facebook profile (i.e., hometown, current city, contact information, relationship status, friends, followed people, visited places), following a recent study that explored some of those variables [32].

Each Facebook profile was coded by means of an observation grid, which was structured using objective criteria regarding the activity on Facebook, extracting from each profile the dimensions of interest (i.e., complete activity, wall posts, profile picture edits, personal photos, photos, videos, quotes, likes, activities with likes, comments, wall post length, wall post average length). The grid aimed to increase the reliability and validity of the data mining by making the procedures more precise and reproducible. Moreover, the tool allowed a more effective and immediate comparison between different data collectors, as well as a robust way to allow for their synchronization and operative standardization.

One year of the participants’ activities on Facebook were analyzed, with a total of 32,368 activities (28,878 were wall posts) and 62,083 comments.

C. Data Analysis

A preliminary description of each sample was carried out by adopting standard descriptive statistics, and producing the psychometric variables for all the dimensions taken into consideration.

The inferential data analysis procedure began with the assessment of the preconditions required by the statistics adopted, assessing in particular the minimum sample size, balance of sub-samples, and their normality distribution parameters. In order to evaluate the differences between the sub-samples defined by the operative factors of the study, the student’s t-test statistics were adopted. Next, the Pearson r linear correlation coefficient was used to evaluate the degree of association regarding the
continuous variables. And finally, a linear regression analysis was employed to estimate the best model fitting the experimental data.

1) Linguistic inquiry word count computer-program (LIWC): Consistent with the need to study the linguistic use and self-presentation on SNSs, the linguistic content of the Facebook profiles was analyzed by applying the widely used Linguistic Inquiry Word Count computer-program (LIWC) [42, 45, 46]. All the profiles were considered separate elements, and the linguistic analysis concerned the entire production (i.e., the posts) as a single narration. The LIWC software analyzed transcripts on a word-by-word basis and compared words with a dictionary divided into 80 linguistic dimensions, including affective terms, cognitive terms, and social and communicative processes. For a complete overview, refer to literature [47].

2) Defining criteria for the convergent validity of the new model (“The Excessive Online Self-Presentation Model”): Following the link between excessive online self-presentation and narcissism, six models from literature describing narcissistic features on Facebook were taken into account [7, 10, 12, 13, 24, 41, 44].

Each model was developed by mapping the dimensions inferred from the empirical models on the Facebook activity variables and LIWC dimensions considered in this study.

The Buffardi model was defined with the following dimensions: wall posts, comments, personal information, quotes, profile picture edits, and personal photos [7].

The Bergman model presented the following variables: friends, photos, places visited, wall posts, profile picture edit, personal photos, and photos with other people [10].

The Carpenter model selected the following variables: wall posts, personal photos, profile picture edits, followed people, anger words, and swear words [11].

The DeWall model presented the following dimensions: first-person singular pronouns, reflexive pronouns, first-person singular verbs, anger and swear words, and personal photos [44].

The Holtzman model revealed the following variables: wall posts, wall post length, comments, friends, friends’ words, anger words, school words, sexual words, and swear words [24].

Finally, the Panek model presented the following variables: complete activity, wall posts, and profile picture edits [41].

All the final model scores were computed as the sum of the zeta-scores of each variable composing them. The resulting distribution reported an average of 0, and a peculiar standard deviation. The averages of the six models were acceptably distributed in a Gaussian way, showing different standard deviations that suggested their different sensitivity.

Since these six literature models were not designed as exhaustive models, it was decided to integrate them by creating an index (i.e., Centroid Model), which was defined as the average of the six models under scrutiny. In this way, a more precise and wider index was obtained to assess narcissism as a continuous trait [48]. The score of the Centroid Model was assumed as the “external” parameter estimating the subjects’ excessive self-presentation, and was used as external criterion to build and validate the new model on excessive online self-presentation, and the well-defined relationship between this behavior and narcissism [10-13].

The Centroid Model reported satisfactory parameters of normality for inferential analysis.

To construct the new model, first the correlations between each literature model and all the LIWC categories were calculated, except for those already contained in the models [7, 10, 11, 24, 41, 44]. Then, all the LIWC variables that did not correlate with any model were discarded, and a list of LIWC categories characterized by a total computed score was obtained as follows: each LIWC variable received one point for each model significantly correlated and one point if the LIWC variable already belonged to literature models, except for those containing such a LIWC variable. After this process, only four variables (i.e., Comma, Physical, Sexual, and Word Count) reached the maximum score (6), while only 23 variables out of 80 dimensions proposed by LIWC appeared to have a score above 3. All the possible combinations of the selected variables were considered as potential models and compared by means of linear regression analysis.

III. RESULTS

A. Descriptive Statistics of the Facebook Profiles Personal Information and Facebook Activities

Concerning the variables of the Facebook profiles, the participants disclosed a large amount of personal information. Particularly, only 16% of participants did not include a birth place, and only 12% did not declare their current city. Regarding contact information, only 14% of participants showed their mobile phone number, 10% their home address, 54% their Facebook email, and 8% their personal email address. With regard to relationship status, 48% of the participants reported being single, while the others specified they were in a relationship. The average number of Facebook friends was 3085.87 (SD=1089.118; range 642-4970), and 64% of participants reported an average high number of places they had been (M = 64.03; range 1-644). Moreover, 62% of participants indicated how many people they followed (M = 11.87; range 1-46).
In reference to H1, a relationship between variables regarding the Facebook activity of the participating adolescents was investigated. Descriptive statistics and correlations are presented in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
</table>

**TABLE 1 DESCRIPTIVE STATISTICS AND CORRELATIONS OF FACEBOOK ACTIVITIES**

Indeed, publishing more wall posts was strongly related to a higher overall activity level on Facebook, and was also related to a higher number of published photos and videos. Moreover, adolescents who published more wall posts also received more likes and comments on their activities. Since the Facebook variables presented many correlations, these variables were checked for when testing the aims.

As H2 focused on the relationship between linguistic analysis and the variety of activity on Facebook, Tables 2 and 3 show such results. The relationship between the Facebook activity variables and the LIWC categories was investigated using the Pearson correlation coefficient. Several significant associations are shown in Tables 2 and 3.

<table>
<thead>
<tr>
<th>LIWC Linguistic Categories</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
</table>

**TABLE 2 CORRELATIONS BETWEEN FACEBOOK ACTIVITY VARIABLES AND LIWC LINGUISTIC CATEGORIES**

Note: Correlations smaller than ±.27 were not significant; Correlations with * were significant at p. < .05; Correlations with ** were significant at p. < .01.

These seventeen Facebook variables are the same as those displayed in Table 1.

**TABLE 3 CORRELATIONS BETWEEN FACEBOOK ACTIVITY VARIABLES AND LIWC PSYCHOLOGICAL CATEGORIES**

Note: Correlations smaller than ±.27 were not significant; Correlations with * were significant at p. < .05; Correlations with ** were significant at p. < .01.

The percentage variables presented a mean and standard deviation because they were computed for each subject separately.
Table 2 shows that the Facebook profiles whose wall posts presented higher word counts and more commas had a higher number of wall posts, particularly longer wall posts and more quotes, and published more activities with likes, which received both likes and comments. The participants, whose wall posts showed more pronouns, particularly first-person singular pronouns, published more personal photos and quotes, received more likes, presented more activities with likes, and had longer wall posts. As main results, Table 3 shows the participants who spoke more about physical, body, and sex, and negative emotions presented a higher complete activity, particularly more and longer wall posts, more personal photos, quotes, and activities with likes. The participants who spoke more about friends exhibited longer wall posts and received more likes. Ultimately, speaking more about anger and sadness was related to receiving more likes and presenting more quotes.

B. Excessive Online Self-Presentation and Its Features about Facebook Activities and LIWC Categories

To verify H3 and H4 regarding the excessive online self-presentation’s features, a correlation analysis between Facebook activity variables, LIWC categories, and the Centroid Model was conducted. The results are shown in Table 4.

<table>
<thead>
<tr>
<th>Activity on Facebook Variable</th>
<th>LIWC – Linguistic Processes Variable</th>
<th>Pearson Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall Posts</td>
<td>Word Count</td>
<td>.84**</td>
</tr>
<tr>
<td>Activities with Likes</td>
<td>Comma</td>
<td>.84**</td>
</tr>
<tr>
<td>Complete Activity</td>
<td>Personal Pronouns</td>
<td>.83**</td>
</tr>
<tr>
<td>Personal Photos</td>
<td>Dictionary Words</td>
<td>.79**</td>
</tr>
<tr>
<td>Wall Posts Length</td>
<td>1st Pers. Pronouns Singular</td>
<td>.79**</td>
</tr>
<tr>
<td>Comments</td>
<td>If</td>
<td>.72**</td>
</tr>
<tr>
<td>Likes</td>
<td>Present Tense</td>
<td>.70**</td>
</tr>
<tr>
<td>Quotes</td>
<td>Verb First Person Singular</td>
<td>.67**</td>
</tr>
<tr>
<td>Photos</td>
<td>Numerals</td>
<td>.66**</td>
</tr>
<tr>
<td>Profile Picture Edit</td>
<td>Articles</td>
<td>.65**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Facebook Profile Variable</th>
<th>LIWC – Cognitive Processes</th>
<th>Pearson Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map</td>
<td>Discrepancy</td>
<td>.52**</td>
</tr>
<tr>
<td>Videos</td>
<td>Exclusive</td>
<td>.47**</td>
</tr>
<tr>
<td>Friends</td>
<td>Cognitive Processes</td>
<td>.34*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LIWC – Affective Processes</th>
<th>LIWC - Biological Processes</th>
<th>Pearson Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger Words</td>
<td>Physical</td>
<td>.42**</td>
</tr>
<tr>
<td>Sadness Words</td>
<td>Sexual</td>
<td>.41**</td>
</tr>
<tr>
<td>Affective Processes</td>
<td>Body</td>
<td>.41**</td>
</tr>
</tbody>
</table>

Note: Correlations with * were significant at p < .05; Correlations with ** were significant at p < .01.

Adolescents showing a higher tendency toward excessive self-presentation appeared to exhibit a higher level of activity on Facebook, had more friends, published more posts, particularly longer posts, had more activities receiving likes, and posted more photos. Consequently, they received more likes and comments. Regarding the correlations between the Centroid Model and the LIWC categories, the linguistic category indicated in general that excessive online self-presentation involved more structured, longer, and richer posts (i.e., higher word count and vocabulary, more commas, pronouns, articles, prepositions, and less numerals). Particularly, a higher number of first-person singular pronouns and first-person pronouns verbs was revealed. The affective category suggested that the prevalent emotions were usually negative.

The cognitive processes presented in excessive online self-presentation profiles were linked to concepts of discrimination (i.e. exclusive), and, at the same time, frequent usage of verbal immediacy (i.e., discrepancy, certainty) was shown. Finally, the biological category showed how the physical, sexual, and body words were used more often in excessive online self-presentation.

C. New Model: a Development and a First Implementation

To refine a new model to reveal the participants’ excessive online self-presentation using only publically accessible posts, the observables derived by the LIWC analysis were investigated as a potential criterion. Only four LIWC variables (i.e., word count, sexual, commas, and physically-related words) were found to be significantly correlated with all of the six literature models taken into account. Such criteria were considered as potential parameters of the new model, and compared by means of standard regression analysis.

Linear regression analysis was adopted to refine the proposed model (F = 61, p < 0.001), which considered two variables including word count and sexuality, explaining 71% of the variance of the data (Table 5).
The relations among the six models, the Centroid Model, and the proposed model were verified using the Pearson correlation coefficient (Table 6), and the resulting matrix was characterized by highly positive correlations. The proposed model reported a strong correlation with the Centroid Model (r = .83, p < .01), confirming its convergent validity with the independent models that were considered. Moreover, it showed large positive correlations with the six literature models (i.e., r > .51).

TABLE 6 DESCRIPTIVE STATISTICS AND CORRELATIONS AMONG THE SIX MODELS, CENTROID, AND NEW MODEL

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) - Holtzman Model</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) - DeWall Model</td>
<td>.69</td>
<td>.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) - Buffardi Model</td>
<td>.65</td>
<td>.41</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) - Carpenter Model</td>
<td>.76</td>
<td>.70</td>
<td>.74</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) - Bergman Model</td>
<td>.61</td>
<td>.37</td>
<td>.84</td>
<td>.79</td>
<td>.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) - Centroid Model</td>
<td>.86</td>
<td>.70</td>
<td>.90</td>
<td>.90</td>
<td>.91</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) - New Model</td>
<td>.83</td>
<td>.51</td>
<td>.78</td>
<td>.82</td>
<td>.67</td>
<td>.64</td>
<td>.83</td>
<td></td>
</tr>
</tbody>
</table>

Mean         - .0107  - .0001  .0001  .0001  - .0001  - .0001  .0001  .341
St. Dev.     4.92     4.05   4.01   2.62   3.45   3.98   .85   1.06
Skewness     .17     -.24   -.59   1.28   .19   .87   .61   1.4
Kurtosis     -.60    -.13   -.41   1.19   -.30   1.39  -.38  1.8
Minimum      -10.49  -10.38  -7.15  -3.26  -7.68  -6.43  -1.42  -1.07

Note: All reported correlations were significant at the level of p < .01. The average score of each model was 0 because its nature of zeta-scores sum.

Finally, to verify H5, student t-tests were conducted to explore gender differences in the linguistic style and in the kind of published Facebook posts (Table 7). Findings showed that females, compared to males, exhibited significantly higher average scores in almost all the LIWC categories. Particularly, females used more negations, more commas, and more personal pronouns, and principally reflexive and second-person pronouns. Moreover, they used a wider vocabulary, and posted more often about physicality, body, sex, negative emotions, and anxiety. On the contrary, males revealed a higher number of words per sentence and spoke more often about sports. Concerning Facebook activity variables, a few significant gender differences were found, with females posting longer posts and a higher number of quotes and photos. Independent-sample t-tests were conducted to compare the scores of males and females in all six literature models, the Centroid Model, and the proposed model. Only the Holtzman model (t(48)=-2.90, p < .05), the Bergman model (t(48)=-2.75, p < .05), and the Centroid Model (t(48)=-2.23, p < .05) showed significant differences, with females displaying higher scores. Nevertheless, as supposed in H5, the proposed model did not show a statistically significant difference between females and males.

TABLE 7 GENDER DIFFERENCES IN LIWC CATEGORIES AND FACEBOOK ACTIVITIES

<table>
<thead>
<tr>
<th>LIWC Categories</th>
<th>Male M</th>
<th>SD</th>
<th>Female M</th>
<th>SD</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linguistic Processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negation</td>
<td>2.40</td>
<td>.90</td>
<td>3.26</td>
<td>.52</td>
<td>-4.16**</td>
</tr>
<tr>
<td>Comma</td>
<td>2.97</td>
<td>1.67</td>
<td>4.49</td>
<td>1.03</td>
<td>-3.88**</td>
</tr>
<tr>
<td>Conditional</td>
<td>.54</td>
<td>.29</td>
<td>.77</td>
<td>.18</td>
<td>-3.219**</td>
</tr>
<tr>
<td>Words/Sentence</td>
<td>16.66</td>
<td>8.79</td>
<td>11.31</td>
<td>3.45</td>
<td>2.83**</td>
</tr>
<tr>
<td>Personal Pronoun</td>
<td>6.30</td>
<td>1.58</td>
<td>7.11</td>
<td>1.00</td>
<td>-2.16*</td>
</tr>
<tr>
<td>Reflexive Pronoun</td>
<td>1.31</td>
<td>.45</td>
<td>1.59</td>
<td>.39</td>
<td>-2.37*</td>
</tr>
<tr>
<td>Dictionary</td>
<td>51.53</td>
<td>6.91</td>
<td>55.04</td>
<td>4.54</td>
<td>-2.26*</td>
</tr>
<tr>
<td>Apostrophe</td>
<td>1.92</td>
<td>.82</td>
<td>2.52</td>
<td>1.10</td>
<td>-2.22*</td>
</tr>
</tbody>
</table>

Psychological Processes |       |    |          |    |         |
| Physical         | 1.14  | .44| 1.66     | .32| -4.74** |
| Body             | .67   | .34| .99      | .19| -4.7** |
| Sexual           | .30   | .17| .39      | .12| -2.17*  |
| Negative Emotions| 1.77  | .90| 2.06     | .54| -2.01*  |
| Anxiety          | .111  | .07| .185     | .12| -2.59*  |
Among the many high correlations and the results obtained by linear regression analysis, H2 was confirmed, as specific linguistic peculiarities were revealed in this study. It was emphasized that people who used more words (e.g., word count), signaled better communication and appeared to publish a higher number of wall posts and quotes [47]. Perhaps it suggested a relationship between publishing more posts and a better strategy in interacting with other people online. Moreover, the results showed how negative emotions correlated with a higher activity on Facebook, and a higher number of posts and receiving more likes. This latter result was also found to be related to anger words and sadness. Such outcomes contrasted previous literature, which asserted that negative emotions were less employed than positive emotions in status updates [42]. On the contrary, this study revealed that negative emotions not only were frequent in posts, but they were even a suitable strategy to interact with a larger audience. Furthermore, positive correlations were shown between the use of first-person singular pronouns and a higher number of longer wall posts, personal photos, and activities that received likes. Since the use of the first-person singular pronouns points to a focus on one’s self, it could be congruent with the purpose to publish more wall posts on personal pages, such as more personal photos, in order to receive more likes [39]. Moreover, according to Cegala, such function words also reflect a psychological intimacy and an engagement in social interactions, which could underline the intention to publish more posts to interact with other users [40].

Since this work was dedicated to enhance the comprehension of adolescents’ SNSs usage, the literature models were taken into account, and the predicting value of the considered variables were weighted. All six models revealed high mutual correlations and a strong correlation with the Centroid Model, confirming how such models measure similar theoretical perspectives. The models considered different implementations of the same construct, increasing the sensitivity and entirety of the model defined by their sum. The correlation analysis regarding the Centroid Model confirmed its representativeness (i.e., good convergent validity), and effectively integrated the six models. Many variables on Facebook activities were strongly and positively related to the Centroid Model, confirming H3 and supporting literature about how excessive self-presentation in SNSs displays more frequent activity and a publication of more personal data [7, 15]. According to H4, in order to examine which linguistic dimensions were more central to assess an excessive self-presentation style, the correlations between the LIWC categories scores and the Centroid Model score were computed.

This study confirmed the hypothesis in that a relationship between this kind of self-presentation and the use of negative (i.e., use of sexual words and anger words) was found, and also endorsed literature [24, 44].

The results also highlighted a correlation between the use of first-person singular pronouns and excessive self-presentation, enhancing how such individuals use implicit markers of self-focus [44].

Among the many high correlations and the results obtained by linear regression analysis, the proposed model selected two main dimensions (i.e., word count and sexuality), again confirming H4. A strong association was reported between the proposed model with all six literature models.

As already shown in previous studies, this study’s data described how excessive self-presentation consists of both positive behaviors, such as sociable behaviors (i.e., talking about friends), and negative behaviors (i.e., swear words and sexual language) [49]. Disclosing lots of information to online friends, highlighted by the presence of a higher number of words per
post (i.e., word count), appeared to increase the feeling of connectedness with others and reduce loneliness [50]. Moreover, they might have learnt strategies to better express themselves, such as constructing clearer and more detailed sentences, which revealed word count as being an important marker to assess such a trait, and confirming these people as sociable and prone to keep effective relationships [49]. Nevertheless, Holtzman, Vazire, and Meh discovered frequent use of sexual language in narcissistic profiles, supporting the analysis of sexual markers to understand both self-promotional and attention-seeking behaviors [24].

As reported in H5, this study was also devoted to exploring gender differences in online self-presentation. The results confirmed previous literature about females publishing more photos, particularly personal photos, quotes, and longer wall posts for self-presenting [6, 7, 19]. Although our results revealed gender differences in the LIWC categories, any gender difference was found in the Excessive Online Self-Presentation Model, confirming H5. This appeared to be particularly relevant, since it may reveal the same risk, for both males and females, in engaging in excessive self-presentation, but with different styles. Indeed, a diverse linguistic style was shown between the sexes, and the use of Facebook with different purposes for males and females [18-22]. Indeed, males disclosed more basic information (i.e. sports), which confirmed literature, while females were confirmed to use more emotional words, and more psychological processes in their posts [17, 43]. On the contrary, they employed more reflexive and second-person singular pronouns than males, instead of a higher number of first-person singular pronouns [43]. Noteworthy, while literature revealed that males historically used more swear words than females, which was not shown in this study, the results displayed that females used a higher number of physical, body, and sexual words. Since swear words and sexual language were both considered as negative language, these findings appeared to show a reversed such tendency [49].

The content analysis of Facebook profiles enabled this study to investigate original text material and to assess excessive online self-presentational styles by analyzing personal Facebook pages [16]. The results confirmed how linguistic analysis is a useful tool to investigate self-presentation on SNSs, and can be applied to study emotional and cognitive aspects [24, 29, 44].

V. CONCLUSIONS

In this study, efforts were first made to plan and carry out research that would organize knowledge about self-presentation on Facebook by adolescents, since they appreciate Facebook and use such a social network to spread and build their identity [27]. Exploring one year of contents from their personal Facebook pages gave the opportunity to understand their online social and psychological behavior [14].

Particularly, excessive online self-presentation was explored to understand how it is characterized in adolescence, since this tendency might be signal of a risk situation and has a correlation with narcissism [12, 33].

The main limitation of the study was the nature of the samples, which included 50 high school students from the same institution, although generalizability was balanced by the advantages of studying language use in naturally occurring environments [37]. Moreover, another limit worthy of attention was represented by the complex nature of the relationships between personality and psychological features in general with the cognitive dynamics within social networks. Such a limit should be controlled by enlarging the participant pool and adopting a multivariate approach to the data analysis, in particular controlling the combined effects of the individual features on digital life dynamics.

Future research is recommended to examine how these findings generalize to other Facebook populations, such as college students and older people. The absence of a self-report measure was a partial limitation, although the content analysis of personal wall posts might be a valid cue to accurately assess self-presentation, as suggested by Deters, et al., since his study underlined the appropriateness of non-self-report measures to explore online behaviors and the possibility of detecting self-presentational strategies by just watching personal profiles [12]. Linguistic analysis is objective and quantifiable behavioral data, and unlike surveys and questionnaires, it allows a “free” self-presentation in the users’ own words [43]. Moreover, a self-report scale may sometimes encounter several difficulties (i.e. people not answering all questions, social desirability bias, etc.), and the implementation of a model based on SNSs contents represents an ecological strategy to avoid well-known disturbing effects such as the Rosenthal and Hawthorne effects [51, 52].

The good convergent validity of the model, and the high correlations with all six literature models, appeared to confirm the potential of this tool to correctly assess excessive online self-presentation. Future studies are recommended with diverse samples to validate such an instrument and increase knowledge about online self-presentation in adolescents.

Finally, from a practical point of view, this paper suggested how excessive self-presentation can be assessed by analyzing public posts on a social network. On the other hand, from a theoretical perspective, the results suggested how the self-presentation dynamics were transformed by the virtual environment constraints.

ACKNOWLEDGMENTS

This work was partially funded by the European Commission, under the FP7 EINS Open Call Project FOCAL (Grant number: FP7-288021). We thank the Prato district for giving us the opportunity to collect data during the “ARCA Project”. We thank the Prato high school “Gramsci-Keynes” for the support and availability for the data collection, as well as for the
promotion of the research activities partially object of this study. Thanks are extended to Ms. Serena Coppolino Perfumi for English proofreading.

REFERENCES


[27] Z. Jordán-Conde, B. Mennecke, and A. Townsend, “Late adolescent identity definition and intimate disclosure on Facebook,”


Andrea Guazzini was born in Massa Marittima (Italy) on January 6th, 1976. He received his Master’s degree in Psychology and his Ph.D. in Complex Systems and Non-linear Dynamics from the University of Florence.

He is currently Researcher of the Department of Education and Psychology, and member of the Centre for the Study of Complex Dynamics of the University of Florence.

Dr. Guazzini has published research on the Nature open access journal Scientific Reports, Physical Review B, and Physica A. He has a keen interest in sociophysics and social psychology as well as a passion for cognitive modeling. He is Member of the Scientific Board for the PhD Curriculum of Complex Systems and Nonlinear Dynamics, within the PhD School on Information Engineering, Faculty of Engineering, University of Florence, Italy, and Responsible for the VirtHuLab, “Laboratory for the Study of Virtual Human Dynamics”, at the University of Florence.
**Cristina Cecchini** holds a Bachelor’s Degree in Science and Techniques of Developmental and Educational Psychology and a Master’s degree in Psychology from University of Florence.

She is currently Ph.D. student at the Department of Information Engineering, University of Florence, curriculum in complex system and non-linear dynamics. She has been nominated as an expert on the field of “Community and Groups Psychology” from Department of Education and Psychology at the University of Florence.

Ms. Cecchini is Member of Scientific Associations as the Italian Society of Community Psychology (SIPCO). Her research interests are risk behaviors in adolescence, virtual dynamics and community dynamics modeling.

**Elisa Guidi** holds a Bachelor’s Degree in Science and Techniques of Clinical and Community Psychology and a Master’s degree (cum laude) in Psychology from University of Florence.

She is currently a Ph.D. student with scholarship of the Department of Information Engineering at the University of Florence (curriculum in complex system and non-linear dynamics). She has been nominated as an expert on the field of “Empowerment of Community and Quality Methods of Research” from Department of Education and Psychology at the University of Florence. She has international collaborations with the Department of Social and Work Psychology (University of Sao Paulo – USP - Brazil) and the Department of Sexology (University of Québec at Montréal – UQAM - Canada).

Ms. Guidi is Member of Scientific Associations as the Italian Society of Community Psychology (SIPCO) and Society and the Society of Community Research and Action (SCRA). Her research interests are gender, family violence, dating violence and virtual community dynamics modeling.

**Monica Milani** got her Master’s degree in Sociology from the University of Urbino, and her Ph.D. in Telematics and Information Society from the University of Florence.

She is currently the School manager of the Gramsci-Keynes High school of Prato (Tuscany, Italy).

Dr. Milani is interested about the study of adolescents’ online dynamics.

**Patrizia Meringolo** graduated with honors in Education with Psychological Specialization in University of Florence.

She is Full Professor In Social and Community Psychology at Department of Education and Psychology of University of Florence, where is part of Research Unit in research and Action for Psychosocial well-being and of Research Unit in Gender Pedagogy and Psychology. She attended the School for Systemic-relational Psychotherapy, Rome for Training in Systemic-relational Psychotherapy.

Prof. Meringolo is Member of Scientific Associations. She is the President of the Italian Society of Community Psychology (SIPCO). She is member of the European Community Psychology Association (ECPA), the European Health Psychology Society and the Society of Community Research and Action (SCRA). She is in charge for International Agreements between University of Florence and other University in European and extra European countries. She has carried out researches and action researches, asked by Local Authorities, Health Public Services and NGO, and Projects funded by European Social Fund. Her research in last years concerned issues about health promotion, gender differences, and psychosocial problems related to migration, risky behavior and substance use. She has produced, individually or within a research group, more than 250 essays.