Dean&Francis
ISSN 2959-6149

The impacts of gender, educational level, and personality on the reaction to and use frequency of offensive words on social media by Chinese young adults

Pengyu Chu[1,*,#],

Siqi Yuan[2,#],

Sichen Wu[3],

Ziyun Zhao[4]

[1]Department of Psychology, Lingnan University, Hong Kong 999077, China, pengyuchu@ln.hk
[2]The College, University of Chicago, Chicago, 60637, United States, siqiy@uchicago.edu
[3]Beijing Haidian Foreign Academy, Beijing 100000, China, 3326058365@qq.com
[4]Nanjing No.13 middle school, Nanjing 210000, China, 3416644934@qq.com
**corresponding author*
#co-first authors

Abstract:

This study examined the relationship between gender, educational level, and personality traits and the reaction to and use frequency of offensive words by Chinese young adults. The target group was 16 to 25 years old and binary gender. The educational level divided into two levels with university graduation as the cut-off point. Personality traits were measured using selected components of the Big Five personality, openness, extroversion, and neuroticism. Offensive words are categorized into vulgar, pornographic, and hateful language. Statistical methods adopted multiple linear regression, and factorial ANOVA was further conducted to explore the influence of educational level and personality traits after controlling for gender. The results suggested that gender and openness were significant predictors for negative reactions to offensive words, while extroversion and neuroticism significantly influence reaction to or frequency of use within a single gender. This research could help moral educators, web censors, and censorship managers anticipate the reaction and frequency of young Chinese adults and provide more focused instruction and judgments.

Keywords: offensive words; social media; Chinese young adults; gender difference; personality traits.

. Introduction

With the development of the internet, people are becoming increasingly aggressive online. Past research has pointed out that in online virtual areas, people are less likely to take the perspective of others and more willing to think individualistically [1]. Moreover, fewer of them thought about the moral and ethical implications of their spontaneous actions. Moral emotions such as compassion require slower information processing because they need deeper reflection, yet the internet environment represents the opposite side of careful pondering -- fast circulation and consumption [2]. The lack of consideration for unknown users and the anonymity of online communication hamper people's ability to empathize cog-

ISSN 2959-6149

nitively, leading to the aggressive expressions online [3, 4]. For young adults, these impacts are stronger on them due to presence of peer behaviors and social feedback [5]. This study will explore the relationship between young adult users the aggressive content in the Chinese social media, represented by offensive word use.

This study employs Jay and Janschewitz's definition of offensive words as vulgar, pornographic, and hateful language [6]. Vulgar language describes crude and insulting terms, such as overt references to sex or body parts; pornographic language a depicts explicit sexual content intended to arouse sexual desire and satisfy erotic desires; hateful language includes any statement that disparages an individual or a group based on a feature such as race, color, ethnicity, gender, sexual orientation, nationality, or religion that is not legally protected [6].

Offensive words spread widely across language families, exhibiting unique characteristics in phonetics, grammar, and other linguistic factors derived from the cultures [7]. To achieve quick conversations, the internet encourages the alternation and variation of offensive words to compensate for the lack of language expression in traditional communication [8]. Direct offensive language including homophobic words and indirect offensive language including sarcasm are two main categories of this evolution. Although online offensive words and their alternatives are neutral outcomes of the interaction between the language, platform, and society, people's reactions toward and habits of using these words are different. This study will explore the factors that affect young Chinese adults' reactions and use frequency of offensive words on social media.

2. Research Design

2.1 Independent Variables

2.1.1 Gender

Gender is considered as an important factor that effects how people perceive and use offensive words on online environment. For a long time, offensive language has been considered aggressive and masculine, discouraging females from using it. Some women would be pressured as they trespass on the expectation of using feminine language and avoid offensive language [9].

A recent study of gender and language provides an inspiring perspective on the new relationship between gender and offensive words in the digital era, claiming that both genders were willing and frequently using offensive words [9]. However, women are less likely to use "sexual anatomy" words than male, including words like "cunt/ fanny" , while some males deploy them frequently in their conversations. This study would perform the test of pornographic language to see how the relationship between gender and online offensive language use is presented in young Chinese adult's life. To expand the research area into offensive words, this study would do the same examination of vulgar language and hateful language, getting a deeper understanding on the impact of gender differences on various kinds of offensive words.

2.1.2 Educational level

Educational levels have an impact on the reaction to and use frequency of offensive words on the premise that education nurtures morality. A psychology paper exploring the empirical studies of psychology on morality over 77 years mentioned that while morality is the aggregative result of various factors, education is an important part of forming moral judgment.

Another paper examines how people are judged based on their language use [10, 11]. The result shows that speakers who use offensive language give poorer impressions to people, including making others underestimate their intelligence. This proves that society takes "using fewer offensive words" as a merited characteristic of highly educated people. Both papers present the correlation between education level and morality and motivate this study's exploration of whether there is a relationship between educational level and offensive language use and reaction.

Different from the eight categories that the International Standard Classification of Education (ISCED) divides, the categories of educational levels were modified to seven in this study – primary school graduate or below, junior high school graduate, high school or secondary school graduate, college graduation or associate degree, university graduation or bachelor's degree, and master's degree or above. In the Chinese context, college graduation or associate degree is specifically referred to as a person with a professional or technical background that received substantially less liberal arts training than university students. The localization of this general definition can help improve the credibility of the study and yield significant results.

2.1.3 Personality

Personality traits massively affect how a person uses language. For example, some researchers directly examine the in-person language habits of extroverts in comparison to introverts [12]. Extroverts display a bigger reliance on abstract interpretation, while introverts speak with concrete facts.

Assuming that people continue to perform their personality traits on social media, the Big Five personality was used as a measure in this study to explore the relationship

between personality traits and language choices online. The Five Personality Theory includes neuroticism, extraversion, openness, agreeableness, and conscientiousness. Roccas et al. defined: Openness refers to the degree of acceptance of new ideas, experiences, and cultural diversity; extroversion refers to the degree to which an individual is attentive to the outside world, socializes, and expresses emotions; neuroticism relates to an individual's emotional stability and ability to regulate emotions [13]. Since the five dimensions are independent of each other and relatively stable, this study ultimately only explores the potential impact of the above three aspects on offensive words, excluding agreeableness and conscientiousness [13].

2.2 Dependent Variables

2.2.1 Reaction to offensive words

Different groups have different attitudes toward the same offensive language. When seeing an offensive word online, some people find it impolite, while others treat it as a normal expression. Through quantification, the study simplified people's reactions to offensive words into five categories on a Likert scale and converted subjective feelings into measurable data. By examining people's response to offensive words, it's easier to deduce whether their attitudes to a certain type of offensive language is a shared perception among people with the same social identities, or a product of individual variations.

2.2.2 Use frequency of offensive words

As articulated in the previous section, people vary in language habits due to numerous factors, including gender, education level, and personality traits. This study measured the use frequency of offensive words to represent the language habit of the participants. By coding the participants' answers into number according to the what they chose, the experimenters would get an accumulated value for each participant. The more frequent the participant use the listed offensive words, the bigger the number value would be.

2.3 Research Questions

As mentioned, previous research has only strongly proved that gender could have an impact on offensive word use. Therefore, this study substantiates the past conclusion and introduces educational level and three aspects of the Big Five Personality to investigate their predictive powers. In terms of offensive words, negative reactions were considered in addition to frequency of use. The following are the research questions.

Research Question 1. Are there any differences regarding the negative reaction and use frequency of offensive words by Chinese young adult on social media?

Research Question 2. How does the education level affect the negative reaction and use frequency of offensive words by Chinese young adult on social media?

Research Question 3. How does the personality traits affect the negative reaction and use frequency of offensive words by Chinese young adult on social media, particularly openness, extraversion and neuroticism?

3. Methodology.

3.1 Questionnaire

The questionnaire consisted of three components. The first section asked about demographic information, including age, gender, educational level, and time spent on social media. To extract the necessary data, the respondents whose age wasn't in the range 16-25 or whose gender belonged to "non-binary" were not included in the analysis. For education level, participants were presented with seven levels from "primary school graduate or below" to "master's degree or above" (See Appendix A, Table A1). For data analysis, only two levels were used: above or below college graduation. The time of social media use was to test the reliability of the samples preliminarily.

In the second part, a simplified version of the Donnellan et al. personality scale was adopted for openness, extraversion, and neuroticism (See Appendix A, Table A2) [14]. Participants completed nine questions and their scores on the dimensions were calculated according to the answer they chose. Each personality dimension was represented by three questions which were mixed to avoid bias.

The last segment focused on the participants' reactions and frequency of use of offensive language, which was categorized into vulgar, pornographic, and hateful language. In the first stage, participants answered questions about how they felt. Subjects were presented with definitions of three offensive word categories, after which subjects reported their judgment of how they felt about a representative list of offensive words collected by the researcher. After finishing the report for the specific expressions, subjects were asked to reflect on their feelings about the corresponding category of these specific words. Responses were given in the order of vulgar language, pornographic language, and hateful language. In the second stage, the participants answered how often they use the offensive words. Although the design logic is the same as the first stage, to avoid repetition, the definitions of the offensive words are presented as additional information after the questions about the general categories.

There were 24 questions; 12 for each dependent variable (See Appendix A, Table A3 and A4). Each category had

ISSN 2959-6149

three questions for the three sets of specific words and one about the category. The items from Table A2 used a 5-point Likert scale with options from "very inaccurate" to "very accurate". Tables A3 and A4 also used a 5-point Likert scale with options from "not offensive at all" to "very offensive" and "never used" to "often used" respectively.

The specific offensive words were mainly collected from "Baidu Tieba", a Chinese Internet forum. The selection and categorization of terms was done by four experimenters, using both manual counting and limited crawler searches. In addition to the three categories of vulgar, pornographic, and hateful language, words with essentially the same meaning were grouped together as "a set of specific words" in the same question. They may be synonymous with each other, as in the case of "cunt" and "slut", or they may have the same intrinsic meaning, as in the case of "fairy disease" and "otaku", which essentially mock the state of life in which men or women are not accepted by social norms, or as in the case of "traitor" and "Jap sucker", which is a specific expression of the former in a particular historical period. In addition, Chinese netizens are also accustomed to using the initials of the Chinese pinyin or homophonic words to represent offensive words, so researchers have classified these as variants of a specific word. Some specific words originally in Chinese have been double translated by native Chinese speakers and native English speakers.

3.2 Sample method and participants

Overall, the sampling method was nonprobability sampling, including convenience sampling and snowball sampling. The researcher collected samples through online questionnaires and snowball sampling on social media. The target population in this study was 16 to 25 years old and binary gender. After excluding all non-compliant samples, the final 137 respondents consisted of 48.2% male and 51.8% female. Additionally, approximately 87% of the participants used social media for more than 3 hours per day, which ensured the sample reliability to a certain extent.

3.3 Analysis Method

Pre-processing was implemented before formal data analysis. First, all samples that were chosen under 16 and over 25 in the age question or non-binary in the gender question were excluded. After the initial analysis of education level, it could be noticed that the total sample was better divided into two parts using high school graduation as the node. Therefore, high school graduation and below was categorized as a low education level, while college graduation and above was a high education level, accounting for 52.55% and 47.45%, respectively. Third, the final scores for openness, extroversion, and neuroticism were calculated after reversing, summing, and averaging. Lastly but crucially, an independent sample t-test was conducted for coding the general types of offensive words. The null hypothesis was that there was no difference between the mean score of specific words and the score of that type of offensive word with only general description, with a significance level of 0.05.

Standard Multiple Regression was used as a statistical tool to address the research questions. With female and high educational level being the reference level, two predictors – gender and educational level – were transformed into two dummy variables respectively. Next, the correlation and predictive power of gender, educational level, and three aspects of personality were explored with the reaction and frequency of three couple of specific words and corresponding general types of words. Specifically, while the gender, educational level, and three aspects of personality remaining as predictors, the outcomes of regression model are the reaction of offensive words and using frequency of offensive words, regarding three sets of specific words and corresponding general types of offensive words. After multiple regression, gender was found to be the most significant predictor across different types of words. Consequently, factorial ANOVA was further conducted within male and female samples to probe the impact of other independent variables on the dependent variable after controlling for gender. All the hypotheses were investigated with Jamovi, and the significance level was set at 0.05.

4. Results

4.1 Coding for the General Types of Offensive Words

For reactions to vulgar, pornographic language and frequency of vulgar, hateful language (Table 1), there were no significant differences between the average scores of specific words and scores of corresponding types of offensive words with definitions (reaction to vulgar language: $p = .910$; reaction to pornographic language: $p = .469$; frequency of vulgar language: $p = .674$; frequency of hateful language: $p = .980$). However, only the reaction to the hateful language ($p = .035$) and the use frequency of pornographic language ($p = .002$) got significant results, indicating these two average scores of specific words could not represent the overall reaction or frequency towards corresponding types of offensive word (Table 2 & 3 and Table 4 & 5). While the reaction to the hateful language and the frequency of pornographic language maintained

using the general question with definition, the others were all using the average score to represent.

Table 1. Independent sample t-test of reactions to vulgar, pornographic language and frequency of vulgar, hateful language

t-test	Statistic	df	p	Cohen's d
Reactions to vulgar language	0.113	272	0.910	0.014
Reactions to pornographic language	0.725	272	0.469	0.088
Frequency of vulgar language	-0.421	272	0.674	-0.051
Frequency of hatful language	0.025	272	0.980	0.003

Table 2. Independent sample t-test for reactions to hateful language

Reactions to hateful language	Statistic	df	p	Cohen's d
Student's t	2.124[a]	272	0.035	0.257
Welch's t	2.124	258.356	0.035	0.257

Note. [a] Levene's test is significant (p <.05), suggesting that a violation of the assumption of equal variances. Therefore, use the Welch's t-test instead.

Table 3. Group descriptives for reactions to hateful language

Group Descriptives	N	Mean	SD	SE
General	137	3.927	0.967	0.083
Specific	137	3.703	0.766	0.065

Table 4. Independent sample t-test for using frequency of pornographic language

Frequency of pornographic language	Statistic	df	p	Cohen's d
Student's t	-3.116	272	0.002	-0.377

Table 5. Group descriptives for using frequency of pornographic language

Group Descriptives	N	Mean	SD	SE
General	137	1.504	1.000	0.075
Specific	137	1.864	1.333	0.088

4.2 Gender as the most significant predictor

Among all these results, gender could significantly predict the negative reaction to pornographic language and the use frequency of all three types of offensive words.

4.2.1 Females feel more offended towards pornographic language

For the pornographic language, the overall model statistically significantly predicted the negative reaction of young adults, F(5, 131) = 2.860, p = .017, R2 = .098, indicating that 9.8% of the variance was explained (Table 6). The RMSE for this model equaled 0.901, which indicated the predicted negative reaction given by this regression model is off by a margin of 0.901 units of the reaction on average. However, only gender was the significant predictor of the reaction over and above the effects of the other factors, t(131) = -3.690, p < .001 (Table 7). The estimated regression coefficient of gender is - 0.587, which indicates that the predicted negative reaction for males is 0.587 units less than that for females, with the other factors constant (Table 7).

ISSN 2959-6149

Table 6. F-test for the overall regression model of the reaction to pornographic language

		Overall Model Test			
R^2	RMSE	F	df_1	df_2	p
0.098	0.901	2.860	5	131	0.017

Table 7. Regression coefficients and statistic tests for gender of the reaction to pornographic language

Predictor	Outcome	Estimate	df	t	p
Gender Male-Female	Reaction to pornographic language	-0.587	131	-3.690	< .001

Additionally, among these cited words, only "ejaculation/masturbation" could be significantly predicted by gender ($t(131) = -3.808$, $p < .001$) with a significant overall model, $p = .010$ (Table 8, 9). Although gender could predict "prostitute" ($t(131) = -2.939$, $p = .004$) and "cunt/slut" ($t(131) = -2.239$, $p = .027$) significantly, the overall models for these words are not significant ("prostitute": $p = .095$; "cunt/slut": $p = .141$) (Table 8, 9).

Table 8. F-test for the overall regression models of the reaction to "prostitute", "cunt/slut" and "ejaculation / masturbation"

Outcome			Overall Model Test			
	R^2	RMSE	F	df_1	df_2	p
Reaction to "prostitute"	0.068	1.109	1.922	5	131	0.095
Reaction to "cunt/slut"	0.061	0.907	1.692	5	131	0.141
Reaction to "ejaculation / masturbation"	0.107	1.231	3.138	5	131	0.010

Table 9. Regression coefficients and statistic tests for gender of the reaction to "prostitute", "cunt/slut" and "ejaculation / masturbation"

Predictor	Outcome	Estimate	df	t	p
Gender Male-Female 1. 1.	Reaction to "prostitute"	-0.575	131	-2.939	0.004
	Reaction to "cunt/slut"	-0.358	131	-2.239	0.027
	Reaction to "ejaculation/masturbation"	-0.827	131	-3.808	< .001

4.2.2 Males use more regarding all types of offensive words

For the vulgar, pornographic, and hateful language, the overall models could all statistically significantly predict the use frequency of young adults (vulgar: p = .004; pornographic: p = .004; hateful: p = .001) (Table 10). The estimate regression coefficients of gender are all positive, suggesting that the predicted using frequency for males is more than that for females regarding three types of words, with the other factors constant (vulgar: Estimate = 0.573, t(131) = 3.341, p = .001; pornographic: Estimate = 0.599, t(131) = 4.155, p < .001; hateful: Estimate = 0.556, t(131) = 4.265, p < .001) (Table 11).

Table 10. F-test for the overall regression models of the frequency of vulgar, pornographic and hateful language

Outcome			Overall Model Test			
	R^2	RMSE	F	df_1	df_2	p
Frequency of vulgar language	0.124	0.973	3.716	5	131	0.004
Frequency of pornographic language	0.123	0.817	3.660	5	131	0.004
Frequency of hateful language	0.139	0.739	4.235	5	131	0.001

Table 11. Regression coefficients and statistic tests for gender of the frequency of vulgar, pornographic and hateful language

Predictor	Outcome	Estimate	df	t	p
1. Gender Male-Female 1. 1.	Frequency of vulgar language	0.573	131	3.341	0.001
	Frequency of pornographic language	0.599	131	4.155	< .001
	Frequency of hateful language	0.556	131	4.265	< .001

For the frequency of vulgar language, all cited words could significantly predict by gender with corresponding significant models ("fuck with variants": p = .008; "dick head with variants": p = .047; "moron/retard with variants": p = .009). And the estimate regression coefficients of gender are all positive ("fuck with variants": *Estimate* = 0.577, $t(131) = 2.755$, p = .007; "dick head with variants": *Estimate* = 0.495, $t(131) = 2.597$, p = .010; "moron/retard with variants": *Estimate* = 0.649, $t(131) = 3.442$, $p < .001$) (Table 12).

Table 12. F-test for the overall regression model and regression coefficients for gender of the frequency of "fuck with variants", "dick head with variants" and "moron / retard with variants"

Overall Model Test		Outcome	Predictor: gender		
F	p		Estimate	t	p
3.259	0.008	Frequency of "fuck with variant"	0.577	2.755	0.007
2.321	0.047	Frequency of "dick head with variant"	0.495	2.597	0.010
3.232	0.009	Frequency of "moron / retard with variant"	0.649	3.442	< .001

For the frequency of pornographic language, except for "prostitute" (*Estimate* = 0.257, $t(131) = 1.802$, p = .074), both "cunt/slut" (*Estimate* = 0.298, $t(131) = 2.528$, p = .013) and "ejaculation/masturbation" (*Estimate* = 0.504, $t(131) = 3.334$, p = .001) could be significantly predicted by gender with corresponding models ("prostitute": p = .308; "cunt/slut": p = .015; "ejaculation/masturbation": p = .023) (Table 13).

Table 13. F-test for the overall regression model and regression coefficients for gender of the frequency of "prostitute", "cunt/slut" and "ejaculation/masturbation"

Overall Model Test		Outcome	Predictor: gender		
F	p		Estimate	t	p
1.211	0.308	Frequency of "prostitute"	0.257	1.802	0.074
2.941	0.015	Frequency of "cunt/slut"	0.298	2.528	0.013
2.704	0.023	Frequency of "ejaculation/masturbation"	0.504	3.334	0.001

Additionally, for the frequency of hateful language, except for "fairy disease or otaku" (*Estimate* = 0.250, $t(131) = 1.445$, p = .151), both "nigger or chino" (*Estimate* = 0.878, $t(131) = 5.321$, $p < .001$) and "traitor or Jap sucker" (*Estimate* = 0.540, $t(131) = 3.731$, $p < .001$) could be significantly predicted by gender with corresponding models ("fairy disease or otaku": p = .170; "nigger or chino": p < .001; "traitor or Jap sucker": p = .011) (Table 14).

Table 14. F-test for the overall regression model and regression coefficients for gender of the frequency of "fairy disease or otaku", "nigger or chino" and "traitor or Jap sucker"

Overall Model Test		Outcome	Predictor: gender		
F	p		Estimate	t	p
1.580	0.170	Frequency of "fairy disease or otaku"	0.250	1.445	0.151
6.209	< .001	Frequency of "nigger or chino"	0.878	5.321	< .001
3.127	0.011	Frequency of "traitor or Jap sucker"	0.540	3.731	< .001

ISSN 2959-6149

4.3. High openness people use less vulgar and certain pornographic language

Among all these results, the unique contribution of openness in explaining the use frequency for certain kinds of offensive words was significant. They are the vulgar language in general (Estimate = -0.464, t(131) = -2.372, p = .019), "fuck words with its synonyms" (Estimate = -0.612, t(131) = -2.568, p = .011) and "cunt/slut" (Estimate = -0.319, t(131) = -2.375, p = .019). All of them have a significant overall model (vulgar language: p = .004; "fuck words with its synonyms": p = .008; "cunt/slut": p = .015) (Table 15).

The estimated regression coefficients of openness for predicting the using frequency of corresponding words are all negative, indicating that a high score in openness possibly will lead to less frequency in using vulgar language in general, fuck words with variant, and cunt/slut. To be more specific, for each score increase in openness, people using frequency for vulgar language is predicted to decrease by 0.464 units, with other factors constant. Similarly, for fuck words and cunt/slut, their frequencies are predicted to be decreased by 0.612 units and 0.319 units respectively (Table 15).

Table 15. F-test for the overall regression model and regression coefficients for gender of the frequency of the vulgar language in general, "fuck words with variants" and "cunt/slut"

Overall Model Test		Outcome	Predictor: gender		
F	p		Estimate	t	p
3.716	0.004	Frequency of vulgar language	-0.464	-2.372	0.019
3.259	0.008	Frequency of "fuck words with variants" (vulgar language)	-0.612	-2.568	0.011
2.941	0.015	Frequency of "cunt/slut" (pornographic language)	-0.319	-2.375	0.019

4.4 While high extroversion for males leads to respectively high negative reaction, high extroversion for females leads to respectively low negative reaction

It is found that the main effects of extroversion on the reaction to "nigger or chino" were both significant within the male and female group after controlling the gender with the factorial ANOVA (Male: $F(1, 50) = 7.448$, $p = .009$, $\eta^2 = 0.106$; Female: $F(1, 55) = 5.017$, $p = .029$, $\eta^2 = 0.068$) (Table 16).

However, within the male group, the mean difference between high and low extroversion was positive, indicating that the negative reaction to "nigger or chino" in high extroversion was significantly higher than that in low extroversion ($t(50) = 2.729$, $p = .009$). While, within the female group, the mean difference between high and low extroversion was negative, indicating that the negative reaction to "nigger or chino" in high extroversion was significantly lower than that in low extroversion ($t(55) = -2.240$, $p = .029$) (Table 17).

Table 16. Factorial ANOVA for the main effect of extroversion on reaction to "nigger or chino" for both genders

Sample	Independent variable	Dependent variable	F	P	η^2
Male	Extroversion	"nigger or chino"	7.448	0.009	0.106
Female			5.017	0.029	0.068

Table 17. Post Hoc comparison for the main effect of extroversion on reaction to "nigger or chino" for both genders

Sample	Comparison		Dependent variable	Mean difference	df	t	P_{tukey}
	Extroversion						
Male	High	Low	"nigger or chino"	0.944	50	2.729	0.009
Female	High	Low		-0.447	55	-2.240	0.029

4.5. High neuroticism leads to respectively high using frequency of offensive words in both genders

In both male and female samples, the main effects of neuroticism were significant for certain kinds of offensive words. For males, it is "fairy disease or otaku" ($F(1, 50) = 4.853$, $p = .032$, $\eta 2 = 0.073$). And for females, they are the vulgar language in general ($F(1, 55) = 4.864$, $p = .032$, $\eta 2 = 0.063$), "fuck words with variant" ($F(1, 55) = 5.285$, $p = .025$, $\eta 2 = 0.068$) and "dick head with variant" ($F(1, 55) = 5.128$, $p = .028$, $\eta 2 = 0.065$) (Table 18).

Besides, for both genders, the mean differences between high and low neuroticism were positive, indicating that the using frequency of these certain kinds of offensive words in high neuroticism was significantly higher than that in low neuroticism ("fairy disease or otaku": $t(50) = 2.203$, $p = .032$; vulgar language: $t(55) = 2.206$, $p = .032$; "fuck words with variant": $t(55) = 2.299$, $p = .025$; "dick head with variant": $t(55) = 2.264$, $p = .028$) (Table 19).

Table 18. Factorial ANOVA for the main effect of neuroticism on male and female using frequency in terms of "fairy disease or otaku", the vulgar language in general, "fuck words with variant" and "dick head with variant" respectively

Sample	Independent variable	Dependent variable	F	P	η^2
Male	Neuroticism	"fairy disease or otaku"	4.853	0.032	0.073
Female		the vulgar language in general	4.864	0.032	0.063
Female		"fuck words with variant"	5.285	0.025	0.068
Female		"dick head with variant"	5.128	0.028	0.065

Table 19. Post Hoc comparison for the main effect of neuroticism on male and female using frequency in terms of "fairy disease or otaku", the vulgar language in general, "fuck words with variant" and "dick head with variant" respectively

Sample	Comparison		Dependent variable	Mean difference	df	t	P_{tukey}
	Neuroticism						
Male	High	Low	"fairy disease or otaku"	0.628	50	2.203	0.032
Female	High	Low	the vulgar language in general	0.462	55	2.206	0.032
Female	High	Low	"fuck words with variant"	0.640	55	2.299	0.025
Female	High	Low	"dick head with variant"	0.515	55	2.264	0.028

5. Discussion

5.1 Females feel more offended towards pornographic language

Tables 2, 3, 4, and 5 present the difference between female and male respondents' reactions and use frequency of pornographic language. The result very much resembles the findings in Stapleton's work [9]. The pornographic languages emphasize the version of humour expressed by the dominant side of the sexual relationship and trivialize the suffering of the submissive side. Words such as "cunt," and "slut" are explicitly pointed toward women, emphasizing their insecurity due to their weaker control of sexual relationships. For some internet users, the seriousness of the words was diluted by alternating them into variations or using them extensively on social media. For some male respondents, the insult in the words was not felt because it was hard to empathize with the other gender, especially on something that is assumed to be exclusive to women. For some female users, however, the words were indecent attacks on their gender, humiliating the whole group, and seeking cruel entertainment from belittling others. Pornographic words subordinate and silence women casually and constitutionally [15]. Thus, females are more likely to feel offended by especially pornographic language.

Research shows that females tend to use offensive language less due to factors including gender stereotypes and expectations [9]. Since the contentious nature of offensive words lay in traditional masculinity, females were not allowed to use words that didn't fit their role. Although the freedom and opportunity of using offensive words among women skyrocketed as internet communities were formed, some females would avoid using these words due to edu-

ISSN 2959-6149

cation and social cues. One of the female participants of the questionnaire claimed that she was taught by her family to speak politely because using offensive words "doesn't sound like a girl". Family and social education are vital to females' significantly less frequent use of offensive words.

5.2 Females use offensive language less frequently

To explain the result in Tables 6, 7, 8, and 9, the group interviewed three female respondents A, B, and C through WeChat and questioned them about reasons not to use the words frequently. Respondent A mentioned that hateful language reminded her of regional discrimination, which she strongly disdained. She also thought words like "traitor" were too harsh to use in daily scenarios as they have serious connotations. Respondent B who took offensive language as a way of catharsis answered that she didn't consider using offensive words as a way to get rid of her negative emotions because she was more "unemotional than males." Respondent C mentioned that she didn't want to link individual actions with the whole group, which was why she disliked using hateful language. She believed that ethnic conflicts should have nothing to do with one's behavior, no matter how provoking it is. At the same time, she mentioned that most pornographic words are humiliating women, so she avoided using them.

The respondents show four possible reasons for young female internet users for their rejection of frequently using offensive words. First, females have a better ability to spot the hidden inequality among the offensive words due to inter-minority empathy. Having the experience of being the second sex, a female understands the discrimination behind these commonly used languages better [16, 17]. Both respondents A and C refuse to use the offensive words frequently due to the underlying meanings of the words, thinking that they are inappropriate and disrespectful to the targeted people. From another perspective, females could be explained as simply more empathetic than males, irrespective of the debate of whether the reason is biological or sociological. Research shows that the notion that "women are more empathetic" is a result of systematic bias and gender stereotyping because females tend to report a stronger empathetic response than their actual feelings [18]. Living under the huge social indication of females' role in emotional abilities, the respondents can perform their indoctrinated characteristics online.

Third, women don't use offensive words as a way to express negative feelings. Research shows that males find directly expressing emotions online, positive or negative, less acceptable than females [19]. Thus, offensive words become a popular medium to hoard their condensed emotions. However, it's reasonable to argue that females don't use offensive words much because they are less willing to engage in conflicts. Without the respondent answering whether she learned to be reserved in potential conflicts naturally or socially, the study could not come up with a precise explanation for the argument. Further exploration is needed to gain a deeper understanding of their choices, and more feedback should be collected to increase the rationality of the deduction. The group did not get a chance to interview any male respondents, thus the group could not compare their replies and examine the differences in their attitudes.

5.3 Impact of openness on how both genders use offensive language

Table 15 explains the role of openness on males' and females' use frequency of offensive words respectively. Openness measures one's acceptance of different mindsets, various experiences, and distinct cultures. People with high openness tend to have more creativity and curiosity [20]. Table 15 shows that higher openness leads to less frequent use of offensive words, which is counterintuitive. Previous research mentioned that using offensive words on social media could help to expand the range of expressions [8]. A possible explanation for the unexpected result is that people of high openness prefer to use various ways of sarcasm online instead of attacking others through insults straightforwardly. Research showed that sarcasm is commonly used on social media, to convey criticism and negative emotions in many languages, including Chinese [21]. Another valid explanation is that they do use offensive words frequently, but due to internet censorship, they need to seek alternatives and use uncommon words to represent the same meaning. The way of evading detection and expanding interpretability influenced the development of slang and memes and underlines the intelligence of the human mind [22]. Even though the common offensive words were not used by high-openness people, they took advantage of the internet culture and internet censorship, creating alternative expressions from traditional offensive words.

5.4 Impact of extroversion on how both genders react to offensive language

Table 16 explores how extroversion impacts people's reaction to offensive words online in both genders. The result showed that the level of extroversion and the use frequency of offensive words in females is positive. This means in female respondents, the more one is extrovert, the more often she uses offensive words. On the other hand, Table 16 shows that male with high extroversion use

offensive words less often. Past research gained complex outcomes on the relationship between extroversion and aggressive actions [23, 24]. Although high extroversion leads to a small possibility of high physical aggression, other behaviors such as the use of offensive language did not show any direct correlation with high extroversion [25]. It could be deduced that the result in the females' case shows the effect of assertiveness in people of high extroversion, making them more resilient and accepting of offensive words. The result of the male group seeks support from the fundamental trait of extroverted people: high adaptation to the environment and high sociability [26]. They are sensitive to and care about the relationship between people, and use offensive words less frequently as it might harm social harmony. Despite the possible reasons, the group needs further research to find an accurate explanation for this inconsistency between genders.

5.5 Impact of neuroticism on how both genders use offensive language

Table 18 showed that high neuroticism results in a higher use frequency of offensive words, regardless of gender. In the book "The Theories of Personality", neuroticism is related to the inclination to psychological distress and emotional instability [20]. People with such characteristics have a greater tendency to use words with negative sentiments online, and in extreme cases, redirect their negative emotions to attack other users for catharsis [27]. Further research found that neuroticism, combined with high extraversion, contributes to delinquency, which explains the tendency of some neurotic people who also possess high extraversion to perform aggressive actions [20]. Although using offensive words isn't considered illegal, under strict circumstances it is immoral. It is then reasonable to claim that people with higher scores on neuroticism use offensive words more frequently no matter their gender.

5.6 Limitations

There are several limitations that inevitably undermine the credibility of the research. The biggest issue is spotted in the offensive word choice. First, the study selected the words from Chinese social media Tieba. There might be chances that the words are exclusive to Tieba, and people who don't use the platform couldn't recognize some words. The group received feedback from some respondents saying that they did not know the meaning of some words, which added errors to the final results. Secondly, due to the time and scope of the research, the study limited the selection to straightforward offensive words. The study didn't include sarcasm, which is the most popular way of expressing criticism and negative feelings on Chinese social media, resulting in a bigger uncertainty in data.

Another drawback of the research was that the findings didn't answer all of the research questions. Educational level is absent from the analysis, either it does not affect people's reaction to and frequency of using offensive words at all, or it wasn't fully investigated. By mainly distributing the questionnaire on WeChat through non-probability sampling, the researchers only had access to people in their social network. As high school and university students who receive good education and monetary investment for doing research, the researchers mostly interact with peers who share similar experiences, especially in schooling. It could be suspected that the sample was biased because it didn't include young adults from diverse backgrounds with different levels of education.

6. Conclusion

In the 21st century, when the Internet is highly developed, social media has become widely accessible. However, this has been accompanied by people's use of offensive words to attack and abuse others online, both consciously and unconsciously. Using Jay and Janschewitz's definition of offensive words as the theoretical framework, this study investigates the impacts of demographic information and three aspects of Big Five personality traits on subjective reaction and use frequency of offensive words [6].

According to the findings, gender is the most significant predictor across different types of offensive words, both general and specific. While females feel more offended towards pornographic language due to the patriarchal essence behind offensive terms and the humiliation it inflicts on women, males use more offensive words across the board, potentially due to systematic bias and expectations and the acceptance of expressing emotions through such language as a form of catharsis. Besides, openness is also a unique predictor which provides new insights into this field. The findings emphasize the crucial role that multiple sarcastic approaches and internet censorship play in the use of more offensive words by high-openness people. Further analysis indicates that while high extroversion for males leads to respectively high negative reactions, high-extroversion females have respectively low negative reactions. The former could be explained by the assertive or dominant nature of extroversion. However, the latter is suspected to be the result of positive social skills and adaptability; the current findings are a strong encouragement for future research in this field. Finally, high neuroticism leads to a higher frequency of offensive word usage in both genders due to its essence of emotional instability, following the tendency towards negative sentiment ex-

ISSN 2959-6149

pression as a form of catharsis.

Overall, our research substantiates the impacts of gender on the reaction and frequency of offensive words and demonstrates people with different levels of openness, extroversion, and neuroticism have diverse perceptions and usage. However, subsequent in-depth studies need to be implemented to further confirm the irrelevance of education level to offensive vocabularies.

Meanwhile, more authoritative vocabulary selection, including ironic words and more elaborate contexts should be considered in further experiments. From a practical perspective, understanding of driving or deterring factors of negative reactions and using the frequency of offensive words can better inform moral educators, web censors, and censorship managers in mainland China for target guidance and supervision. Expanding the scope of research could help mitigate the stigmatization of online language and help combat the hostility of social media commentary which is also a call for further research.

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Acknowledgement

Pengyu Chu and Siqi Yuan contributed equally to this work and should be considered co-first authors.

References

[1] Flores, A., & James, C. (2012). Morality and ethics behind the screen: Young people's perspectives on digital life. New Media & Society, 15(6), 834–852. https://doi.org/10.1177/1461444812462842

[2] Immordino-Yang, M. H., McColl, A., Damasio, H., & Damasio, A. (2009). Neural correlates of admiration and compassion. Proceedings of the National Academy of Sciences, 106(19), 8021–8026. https://doi.org/10.1073/pnas.0810363106

[3] Corkum, M., & N. Will Shead. (2023). Online Moral Disengagement: An Examination of the Relationships Between Electronic Communication, Cognitive Empathy, and Antisocial Behavior on the Internet. Psychological Reports. https://doi.org/10.1177/00332941231216415

[4] Parlangeli, O., Marchigiani, E., Bracci, M., Duguid, A. M., Palmitesta, P., & Marti, P. (2019). Offensive acts and helping behavior on the internet: An analysis of the relationships between moral disengagement, empathy and use of social media in a sample of Italian students. Work, 63(3), 469–477. https://doi.org/10.3233/wor-192935

[5] Armstrong-Carter, E., Garrett, S. L., Nick, E. A., Prinstein, M. J., & Telzer, E. H. (2022). Momentary links between adolescents' social media use and social experiences and motivations: Individual differences by peer susceptibility. Developmental Psychology, 59(4). https://doi.org/10.1037/dev0001503

[6] Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. Journal of Politeness Research. Language, Behaviour, Culture, 4(2). https://doi.org/10.1515/jplr.2008.013

[7] Liu, J., Zhang, X., & Li, H. (2023). Analysis of Language Phenomena in Internet Slang: A Case Study of Internet Dirty Language. OAlib, 10(08), 1–12. https://doi.org/10.4236/oalib.1110484

[8] Qu, R. (2020). Research on the Diffusion of Network Buzzwords from the Perspective of Communication. Probe - Media and Communication Studies, 2(2), 40. https://doi.org/10.18686/mcs.v2i2.1304

[9] Stapleton, K. (2003). Gender and Swearing: A Community Practice*: WL. Women and Language, 26(2), 22-33. https://www.proquest.com/scholarly-journals/gender-swearing-community-practice/docview/198813729/se-2

[10] Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The Psychology of Morality: A Review and Analysis of Empirical Studies Published From 1940 Through 2017. Personality and Social Psychology Review, 23(4), 332–366. https://doi.org/10.1177/1088868318811759

[11] DeFrank, M., & Kahlbaugh, P. (2018). Language Choice Matters: When Profanity Affects How People Are Judged. Journal of Language and Social Psychology, 38(1), 126–141. https://doi.org/10.1177/0261927x18758143

[12] Beukeboom, C. J., Tanis, M., & Vermeulen, I. E. (2012). The Language of Extraversion. Journal of Language and Social Psychology, 32(2), 191–201. https://doi.org/10.1177/0261927x12460844

[13] Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The big five personality factors and personal values. Personality and Social Psychology Bulletin, 28(6), 789–801. https://doi.org/10.1177/0146167202289008

[14] Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five Factors of Personality. Psychological Assessment, 18(2), 192–203. https://doi.org/10.1037/1040-3590.18.2.192

[15] Mikkola, M. (2019). Does Pornography Silence Women? Pornography, 51–87. https://doi.org/10.1093/oso/9780190640064.003.0003

[16] Kasai, M., Toda, Y., & Russell, S. (2023). SOGI Minority and School Life in Asian Contexts. Taylor & Francis.

[17] de Beauvoir, S. (1949). The Second Sex. Oxford University Press.

[18] Löffler, C. S., & Greitemeyer, T. (2021). Are women the more empathetic gender? The effects of gender role expectations. Current Psychology, 42(1). https://doi.org/10.1007/s12144-020-01260-8

[19] Waterloo, S. F., Baumgartner, S. E., Peter, J., & Valkenburg, P. M. (2017). Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. New Media & Society, 20(5), 1813–1831. https://doi.

org/10.1177/1461444817707349
[20] Schultz, D. P., & Schultz, S. E. (2017). Theories of Personality (11th ed., pp. 231–238). Cengage Learning Asia Pte Ltd.
[21] Zhu, Y. (2022). Open Chinese Internet Sarcasm Corpus Construction: An Approach. Frontiers in Computing and Intelligent Systems, 2(1), 7–9. https://doi.org/10.54097/fcis.v2i1.2484
[22] Ye, W., & Zhao, L. (2023). "I know it's sensitive": Internet censorship, recoding, and the sensitive word culture in China. Discourse, Context & Media, 51, 100666. https://doi.org/10.1016/j.dcm.2022.100666
[23] Sharpe, J. P., & Desai, S. (2001). The Revised Neo Personality Inventory and the MMPI-2 Psychopathology Five in the prediction of aggression. Personality and Individual Differences, 31(4), 505–518. https://doi.org/10.1016/s0191-8869(00)00155-0
[24] CAVALCANTI, J. G., & PIMENTEL, C. E. (2016). Personality and aggression: A contribution of the General Aggression Model. Estudos de Psicologia (Campinas), 33(3), 443–451. https://doi.org/10.1590/1982-02752016000300008
[25] Barlett, C. P., & Anderson, C. A. (2012). Direct and indirect relations between the Big 5 personality traits and aggressive and violent behavior. Personality and Individual Differences, 52(8), 870–875. https://doi.org/10.1016/j.paid.2012.01.029
[26] Evans, A. M., Meyers, M. C., De Calseyde, P. P. F. M. V., & Stavrova, O. (2021). Extroversion and Conscientiousness Predict Deteriorating Job Outcomes During the COVID-19 Transition to Enforced Remote Work. Social Psychological and Personality Science, 13(3), 194855062110390. https://doi.org/10.1177/19485506211039092
[27] Shen, J., Brdiczka, O., & Liu, J. (2015). A study of Facebook behavior: What does it tell about your Neuroticism and Extraversion? Computers in Human Behavior, 45, 32–38. https://doi.org/10.1016/j.chb.2014.11.067

Appendix A

Table A1. Questions about educational level

	Question	Choices
Educational Level	What is your highest level of education?	A. Primary school graduation or below
		B. Junior high school graduation
		C. High school or secondary school graduation
		D. College graduation (associate degree)
		E. University graduation (bachelor's degree)
		F. Master's degree or above

Table A2. Questions accessing openness, extroversion and neuroticism

Construct / Sub-Construct		Question
Big five Personality traits	Neuroticism	I have frequent mood swings.
		I am relaxed most of the time. (R)
		I get upset easily.
	Openness	I have a vivid imagination.
		I am not interested in abstract ideas. (R)
		I have difficulty understanding abstract ideas.
	Extroversion	I don't talk a lot. (R)
		I talk to a lot of different people at parties.
		I keep in the background. (R)

ISSN 2959-6149

Table A3. Question accessing the negative reaction to offensive words.

Construct / Sub-Construct		Question	
Negative reaction to offensive words	Vulgar language	How offensive do you feel when facing … on social media?	"fuck with variants"
			"dick head with variants"
			"moron/retard with variants"
			vulgar language in general
	Pornographic language		"prostitute"
			"cunt/slut"
			"ejaculation/masturbation"
			pornographic language in general
	Hateful language		"fairy disease or otaku"
			"nigger or chino"
			"traitor or Jap sucker"
			hateful language in general

Table A4. Question accessing the using frequency of offensive words.

Construct / Sub-Construct		Question	
Using frequency of offensive words	Vulgar language	How often do you use … on social media?	"fuck with variants"
			"dick head with variants"
			"moron/retard with variants"
			vulgar language in general
	Pornographic language		"prostitute"
			"cunt/slut"
			"ejaculation/masturbation"
			pornographic language in general
	Hateful language		"fairy disease or otaku"
			"nigger or chino"
			"traitor or Jap sucker"
			hateful language in general

Note: R = reversed coded items.

Remark 1: The items from Table A2 use 5-point Likert scales ranging from "very inaccurate" to "very accurate", with the instruction "To what extent do the following sentences accurately describe you?". Table A3 and A4 also use 5-point Likert scales. Answers range from "not offensive at all" to "very offensive" and "never used" to "often used" respectively.

Remark 2: For Table A2, the items are categorized for ease of understanding. On the contrary, the actual questionnaire has disorganized the order of the questions to increase validity. The five dimensions of the Big Five personality are included in each of the five questions.

Remark 3: For Table A3 and A4, the items are categorized for ease of understanding. On the contrary, the actual questionnaire has the definitions of each types of offensive words before these clusters of questions begin.

Remark 4: For the sake of flow when answering the questionnaire, the Chinese version was prepared while ensuring consistency of content. The actual questionnaire was displayed in both Chinese and English. For the Chinese version of the questionnaire, please contact the corresponding author.