Journal of Research in Personality 89 (2020) 104035

Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

A meta-analysis of linguistic markers of extraversion: Positive emotion and social process words

Jiayu Chen, Lin Qiu *, Moon-Ho Ringo Ho

School of Social Sciences, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 28 July 2020
Accepted 18 September 2020
Available online 8 October 2020

Keywords:
Extraversion
LIWC
Meta-analysis
Personality
Positive emotion
Social process
Language

ABSTRACT

Past literature has shown that extraversion is related to the use of positive emotion and social process words. However, the strength of the relationships varies substantially across studies. In this research, we conducted a meta-analysis (k = 37, N = 82,132) to estimate the overall effect size of the two linguistic correlates of extraversion. In addition, we tested potential moderators including demographic variables (e.g., age and gender) and communication contexts (e.g., synchronous vs. asynchronous, public vs. private). Our random effects models revealed a small correlation between extraversion and positive emotion words (r = 0.069, *95% CI* = [0.041, 0.096]), and a small correlation between extraversion and social process words (r = 0.077, *95% CI* = [0.044, 0.109]). In addition, the strength of the relationship between extraversion and positive emotion words varies across communication contexts, while the relationship between extraversion and social process words remains consistent across contexts. Our results suggest that positive emotion words and social process words are linguistic correlates of extraversion, but they are small in magnitude.

1. Introduction

Extraversion is a trait that represents the dispositional tendency to experience positive emotions (John & Srivastava, 1999) and enjoy social interactions (Costa & McCrae, 1992a; Eysenck, 1981; Tobin, Graziano, Vanman, & Tassinary, 2000; Wilt and Revelle, 2009). Aligning with this definition, research on personality and word use has shown that extraversion is related to two linguistic markers: positive emotion words and social process related words. For example, extraversion was found to be related to positive emotion words in tweets (Qiu, Lin, Ramsay, & Yang, 2012), Facebook status updates (Kern et al., 2014; Sumner, Byers, & Shearing, 2011), and blogs (Yarkoni, 2010; Gill et al., 2009). In addition, extraversion is positively correlated with the use of social process words in self-narratives (Hirsh & Peterson, 2009), personal essays (Pennebaker & King, 1999), and emails (Oberlander & Gill, 2006). However, the strength of the two relationships varies substantially across studies. Therefore, this study aims to conduct a meta-analysis to estimate the overall effect size regarding the relationships between extraversion and its linguistic markers. This will reveal how extraversion is associated with language use. Our findings will also have important practical implications. Given the unprecedented opportunity provided by Big Data, a growing number of studies have tried to predict personality based on linguistic markers in social media using machine learning approaches (e.g., Park et al., 2015; Youyou, Stillwell, Schwartz, & Kosinski, 2017). However, these approaches often lack structural validity and reliability (Bleidorn & Hopwood, 2019). Our study will provide strong basis to judge the content validity of machine learning approaches and guide the development of effective personality assessment tools.

In the following, we will review research on the characteristics of extraversion, and discuss why word choices may be behavioral indicators of extraversion.

1.1. The characteristics of extraversion

Extraversion has been found to be a fundamental and robust dimension of personality in either the Big Five model (Goldberg, 1990, 1992; McCrae & Costa, 1987) or the six-factor HEXACO model (Lee & Ashton, 2008). It is defined as the tendency to be sociable, assertive, and energetic based on factor analysis of trait descriptive adjectives (Goldberg, 1990; Lee & Ashton, 2008). Meanwhile, according to the two-factor motivational model of personality, extraversion, positive emotionality, and behavioral activation system form the approach temperament factor of personality,

* Corresponding author at: School of Social Sciences, Nanyang Technological University, 14 Nanyang Drive, Singapore 637332, Singapore.

E-mail address: linqiu@ntu.edu.sg (L. Qiu).

https://doi.org/10.1016/j.jrp.2020.104035
0092-6566/

while neuroticism, negative emotionality, and behavioral inhibition system form the avoidance temperament factor (Elliot & Thrash, 2002). This suggests that extraversion reflects vigilance and sensitivity to positive and desirable stimuli (Elliot & Thrash, 2002; Larsen & Ketelaar, 1991; Lucas & Baird, 2004). Extraversion has also been considered the tendency to experience positive emotion and engage in pleasurable activities (Tellegen, 1985; Tellegen et al., 1988; Watson & Clark, 1997), because it tends to have similar correlations with other variables as trait positive affectivity (Watson & Clark, 1997).

Despite the different accounts of extraversion, empirical evidence has shown reliable connection between extraversion and the experience of positive emotions (Argyle & Lu, 1990; Aziz, Mustaffa, Samah, & Yusof, 2014; Costa, McCrae, & Norris, 1981). The connection has been found to hold across cultures (Lucas, Diener, Grob, Suh, & Shao, 2000). Studies found that extraversion is positively related to individuals' general happiness (Tan & Lee, 2017), positive affect (Lin, 2014) and happiness after controlling for other four Big Five traits (Warner & Vroman, 2011). A meta-analysis showed that the correlation between extraversion and happiness is 0.27 based on 15 independent samples, and the relationship between extraversion and positive affect is 0.20 based on 39 independent samples (DeNeve & Cooper, 1998). These studies suggest that extraverts tend to experience more positive emotion than introverts.

Extraversion has also been found to be related to social interaction. Argyle and Lu (1990) measured how participants felt and how often they participated in 37 daily activities (e.g., chat with friend) and found that extraverts enjoyed and participated more in social activities than introverts. Oerlemans and Bakker (2014) asked participants to list the activities they engaged in the previous day and how they had felt during each activity, and found that extraverts experienced higher level of positive emotion during social interactions than introverts. Asendorpf and Wilpers (1998) found that extraverts spent more time interacting with others than introverts in a 21-day dairy study. In a recent study, Harari et al. (2019) used smartphone sensing and found that extraversion was positively correlated with the frequency of outgoing phone call, text messaging, messaging application usage, and social media application usage. These studies suggest that social interaction is an important characteristic of extraversion.

1.2. Personality expression and language use

A sizable amount of research has shown the connection between personality and verbal behavior. For example, Borkenau and Liebler (1992) videotaped and rated participants in various situations such as reading a standard text, and found that extraversion was positively correlated with powerful voice and negatively correlated with soft voice. Back, Schmukle, and Egloff (2009) videotaped and rated participants in tasks such as self-introduction and description of future vision, and found that extraversion was associated with verbal behaviors such as loudness of voice, number of words, and question-asking during small talks. Similarly, Cuperman and Ickes (2009) randomly paired participants in unstructured dyadic interaction and found that extraversion was negatively associated with the use of first-person singular pronouns while agreeableness was positively associated with verbal acknowledgements. While the above evidence illustrated personality expression in verbal behaviors such as acoustic features and linguistic styles, our meta-analysis focused on word use and its relationship with personality.

Word use unveils important psychological processes (Pennebaker, Mehl, & Niederhoffer, 2003). For example, the use of "I" indicates attention to oneself, and its frequent use predicts depression because excessive attention to the self is usually associated with negative emotion (Edwards & Holtzman, 2017; Tackman et al., 2018; Zimmermann, Brockmeyer, Hunn, Schauenburg, & Wolf, 2017). The pronoun use in "you and I can do this" vs. "we can do this" reveals subtle but critical distinction in the speaker's emotional ties to the other person (Chung & Pennebaker, 2007). Linguistic Inquiry and Word Count (LIWC; Tausczik & Pennebaker, 2010) is a software tool that has been developed and widely used to measure emotion, cognitive styles, and social processes based on word use (Pennebaker et al., 2003; Tausczik & Pennebaker, 2010). It counts word frequencies in around seventy pre-defined grammatical and psychologically meaningful categories. Words in these categories were selected and validated by independent judges (Pennebaker, Booth, & Francis, 2007) and shown to have profound psychological implications (Tausczik & Pennebaker, 2010).

The LIWC positive emotion word category includes words indicating positive feelings (e.g., happy, love) and positive valence (e.g., beautiful, nice). They have been found to be related to self-report positive emotion. Kahn, Tobin, Massey, and Anderson (2007) showed that LIWC positive emotion words positively correlated with self-reported amusement after participants watched a comedy film, and participants tended to use more positive emotion words after (vs. before) watching a comedy film. Tov, Ng, Lin, and Qiu (2013) collected participants' daily dairy and self-reported positive emotion for 3 weeks, and found that LIWC positive emotion words were related with self-reported positive emotion. Golder and Macy (2011) used the frequency of positive emotion words in tweets to measure people's positive affect and revealed diurnal and seasonal mood change around the world. Liu, Chan, Qiu, Tov, and Tong (2018) found that individuals from culturally tight U.S. states were more likely to use more positive emotion words and less negative emotion words on Facebook than those from culturally loose states, reflecting the effect of cultural tightness-looseness on emotional expression. These studies established the external validity of LIWC positive emotion word category for measuring positive affect.

The LIWC social process word category contains all personal pronouns except *I*, and words indicating social interactions (e.g., talk, share, meet), family, and friends. The inclusion of words related to social interaction provides face validity of using the category to measure thoughts and attention towards social interactions. The category has further been found to indicate social connections and closeness (Stone & Pennebaker, 2002). Sillars, Shellen, McIntosh, and Pomegranate (1997) showed that interdependent couples tended to use more "we" pronouns (included in the social process word category) in discussions about marital problems than independent couples, suggesting that the use of "we" pronouns indicates relationship closeness. Karan, Rosenthal, and Robbins (2019) conducted a meta-analysis and showed that the amount of we-talk predicted relationship satisfaction and functioning in romantic couples, suggesting that we-talk reflects interdependence between partners. Ritter, Preston, and Hernandez (2014) showed that Christians used more positive emotion words in tweets than atheists, partially mediated by use of social process words. This is consistent with previous results that religious people had stronger social relationships than less religious people, which led to greater life satisfaction (Salsman, Brown, Brechting, & Carlson, 2005). Tov et al. (2013) conducted a diary study and found that LIWC coding of positive family and friends related events (two categories included in social process word category) in the diary predicted self-report satisfaction towards family and friends respectively. These studies suggest that LIWC social process words are indicators of social interaction and connection.

Given that linguistic styles convey important information about the speaker (Pennebaker et al., 2003)), they have been proposed and shown to be valuable indicators of personality (Hirsh & Peterson, 2009; Pennebaker & King, 1999; Youyou et al., 2017). The act frequency approach to personality (Buss & Craik, 1983) posited that individuals higher on a particular trait would perform more acts that are prototypical of that trait. Therefore, it is likely that people tend to use words that reflect feelings, thoughts, and behaviors congruent with their personality characteristics. As extraversion is defined as the tendency to experience positive emotions and enjoy social interactions (Costa et al., 1981), studies have found that extraversion is related to the use of positive emotion and social process words. For example, Yarkoni, 2010 conducted a large-scale analysis of personality traits and language use in blogs, and found that extraversion was positively related with use of positive emotion and social process words. Pennebaker and King (1999) collected participants' personal essays and their personality traits, and found that extraversion predicted the use of positive emotion words and social process words. Qiu et al. (2012) collected participants' tweets over one-month, and found that extraversion was related with the use of positive emotion words and social process words. However, the effect size of the relationships varies substantially across studies. For example, the effect size of the correlation between extraversion and positive emotion words is 0.28 in a sample of tweets (Qiu et al., 2012) but only 0.05 in a sample of daily language use (Mehl, Gosling, & Pennebaker, 2006). The effect size of the correlation between extraversion and social process words is 0.22 in self-narratives samples (Hirsh & Peterson, 2009) but 0.09 in Facebook profile (Golbeck, Robles, & Turner, 2011). Thus, it is important to conduct a meta-analysis to estimate the overall effect size of these relationships.

1.3. The present study

In this study, we aim to conduct a meta-analysis of the published as well as unpublished studies regarding the correlation between extraversion and positive emotion words, as well as the correlation between extraversion social process words. We will also explore several moderators in our meta-analysis.

First, we will test the moderation effect of gender. Studies have shown gender differences in the link between extraversion and emotional experiences in daily life (Eaton & Funder, 2001). Meanwhile, Mehl et al. (2006) found that the use of present tense verbs was related with extraversion only among females while the use of words with more than 6 letters was negatively correlated with extraversion only among males in a sample of daily spoken language. This suggested that there might be gender differences in the relationship between word use and extraversion, although we do not have specific hypotheses regarding the gender effects.

Second, we will test the moderation effect of age. Studies have shown that older individuals tend to use more positive emotion and social process words (Pennebaker & Stone, 2003). Therefore, there might be age differences in the relationship between extraversion and word use. However, we do not have specific hypotheses regarding the moderation effect of age.

Third, we will test the moderation effect of word analysis tools. There are four versions of LIWC, including the original LIWC, LIWC2001, LIWC2007, and LIWC2015. The original LIWC dictionary (Pennebaker & Francis, 1996; Pennebaker, Mayne, & Francis, 1997) consists of over 2000 words and word stems, including 328 words for positive emotion. LIWC2001 dictionary extends the original version and contains 2300 words and word stems (Pennebaker, Francis, & Booth, 2001). It has 261 words for positive emotion, and 314 words for social process. LIWC2007 further expands its dictionary to include around 4500 words and word stems, with 406 words for positive emotion words and 455 words for social process (Pennebaker et al., 2007). LIWC2015 is the most recent version, consisting of about 6400 words, word stems, and emoticons (Pennebaker, Boyd, Jordan, & Blackburn, 2015).

Besides different versions of LIWC, we will include another two tools that have been used to measure emotions in texts. Text Analysis and Word Count program (TAWC) is an open-source program that can take a predefined dictionary (e.g., LIWC 2007 dictionary) to analyze social media texts (Kramer, 2010; Kramer, Fussell, & Setlock, 2004). It has been used to measure positive and negative emotion in Facebook status updates, and show their associations with self-reported life satisfaction (Kramer, 2010). Oedipus Text (Levenson, 1992) is a software program that counts the percentage of emotional words in texts based on a self-defined dictionary. The dictionary contains around 3500 emotion words, including positive emotion words (e.g., joy, love, amusement) and negative emotion words (e.g., anger, disgust, guilt). Oedipus Text has been used to measure schizophrenia patients' narratives of their positive and negative emotional life experiences (Gruber & Kring, 2008).

Fourth, we will test the moderation effect of the duration of language sample. According to the density-distribution theory of personality, states vary across time and trait personality is a summary of the entire distribution of states (Fleeson, 2001). McNiel and Fleeson (2006) found that state extraversion (i.e., acting extraverted) changed participants' emotion, leaving the possibility that state extraversion may influence emotional expression in language. However, it is unclear how state and trait extraversion may differ in their relationship to linguistic style. One possible indicator of trait or state expression could be the duration of the language sample. If the language sample is collected over long stretches of time (e.g., several days), it is likely that its linguistic styles reflect traits rather than states. Thus, we include duration of language sample as a moderator in our study.

Fifth, we will test the moderation effect of synchronous vs. asynchronous communication. Compared with synchronous communication, asynchronous communication allows individuals to have full control of what they want to communicate, and therefore easily manipulate their self-presentation for impression management purposes (Walther, 1996). It is reasonable to expect that the linguistic style in asynchronous communication may be less reflective of personality that those in synchronous communication.

Sixth, we will test the moderation effect of public vs. private context. Mehl, Robbins, and Holleran (2012) collected participants' daily conversations and stream-of-consciousness essays, and found that the personality-word use associations were highly context-dependent. Specifically, they found that extraversion was related with word count in daily conversations but not in stream-of-consciousness essays. Neuroticism was associated with positive and negative emotion words in stream-of-consciousness essays but not in daily conversations. They proposed that privacy could be a moderator because extraversion is a public trait and likely to be expressed in public (i.e., daily conversations) while neuroticism is a private trait and tends to be expressed in private (i.e., stream-of-consciousness essays). Therefore, we will examine the moderation effect of public vs. private context. We consider the situation where the utterance or text is accessible to many people as public context, and the situation where the utterance or text is only accessible to a few people as private context.

Seventh, we will test the moderation effect of real-life vs. lab setting. Compared to real-life settings, laboratory settings are well-controlled and decontextualized. They often give participants specific writing or speaking tasks which may influence their language styles and personality expression. Thus, we will investigate the moderation effect of real life vs. lab setting.

Eighth, we will test the moderation effect of online vs. offline communication. Online communication lacks nonverbal cues (e.g., facial expression) and immediate feedback from the communicating partner, and therefore is expected to enough more self-disclosure than offline communication (Nguyen, Bin, & Campbell, 2012). However, a systematic review of 15 studies found an equal number of studies supporting either online or offline self-disclosure (Nguyen et al., 2012). In addition, a review paper about emotion expression concluded that emotional expression is highly similar in both kinds of communication (Derks, Fischer, & Bos, 2008), likely because people may compensate the lack of nonverbal cues with linguistic cues. Given these mixed findings, we will explore whether online vs. offline communication moderates the associations between extraversion and language use.

Lastly, we will test the moderation effect of language mode (i.e., written vs. spoken). Compared to written language, spoken language has more grammatical errors (Bushnell, 1930), and is more spontaneous and less manipulated (Chafe & Tannen, 1987). Mehl et al. (2012) found that individuals used more words in spoken language than written language. Furthermore, previous studies only found that extraversion predicted positive emotion words and social process words in written languages (e.g., Pennebaker & King, 1999; Qiu, Lin, Ramsay, & Yang, 2012; Yarkoni, 2010), but not in spoken languages (e.g., Mehl et al., 2006). Thus, we will explore the moderation effect of language mode in our study.

2. Method

2.1. Selection of studies

We used the following criteria to include studies in our meta-analysis: (1) The study needs to measure extraversion and use text analysis tools such as LIWC, TAWC, and Oedipus to analyze language data. (2) The study contains at least one of the correlations (the correlation between extraversion and positive emotion words or the correlation between extraversion and social process words), or relevant data are provided to calculate the correlation. (3) The study uses a validated measure to assess extraversion, such as Big Five Inventory (John, Donahue, & Kentle, 1991), the International Personality Item Pool (Goldberg et al., 2006), BFI-2 (Soto & John, 2017), HEXACO-60 (Ashton & Lee, 2009), and Eysenck Personality Questionnaire-Revised (Eysenck, Barrett, & Eysenck, 1984).

Two researchers independently searched multiple databases including PsycARTICLES, PsycINFO, Web of Science, ProQuest Dissertations and Theses, and Google Scholar, using "extraver* AND (LIWC OR TAWC OR Oedipus)" as the search query. These five databases yielded a total of 2057 articles (ProQuest Dissertations and Theses: 1054; Google Scholar: 1000; PsycINFO: 8; Web of Science: 5; PsycARTICLES: 0). Then the two researchers scanned the title and abstract of each article to determine if it met criteria 1. An article was included if at least one researcher considered it to meet criteria 1. The inter-rater agreement was 98.86% for ProQuest Dissertations and Theses, 89.50% for Google Scholar, 75% for PsycINFO, and 80% for Web of Science. There were a total of 172 articles met criteria 1. For articles that meet criteria 1, their full texts were screened using criteria 2 and 3. This resulted in 29 articles. Disagreements during this process were resolved through discussion with a third researcher.

Next, we scanned the reference section of the 29 articles to identify additional studies that might meet our inclusion criteria. One additional article was added after this procedure, resulting in a total of 30 articles from literature search. When one dataset was used in multiple articles, we chose to include the article that provided more information. For instance, Kern et al. (2014) and Ateş (2014) used the same MyPersonality dataset, but Kern et al. (2014) provided age, gender, and reliability of extraversion while such information was missing in Ateş (2014). Therefore, we chose to include Kern et al. (2014) in our meta-analysis. Mehl (2004) and Mehl et al. (2006) used the same dataset, but only Mehl (2004) provided the correlation between extraversion and social process words. Thus, Mehl (2004) was chosen to be included in our analysis.

After the literature search, we posted this meta-analysis project on Researchgate.com, and added all the twenty articles available to the project to notify their authors about this project so that they could share unpublished data with us. We also emailed the corresponding and first author of these articles to request for unpublished data. Furthermore, we requested for unpublished data by emailing the corresponding and first author of the unselected articles in the 172 articles mentioned above due to the lack of information for the correlation between extraversion and the two word categories. This step resulted in an additional 11 samples being included for analysis.

2.2. Study coding

Two researchers independently coded each study according to 15 variables (see Table A in Appendix). Their disagreements were resolved through discussion with a third researcher. The effect size and direction of correlation were used to estimate the overall effect size. Nine variables were treated as moderators: percentage of female participants, mean age of participants, word analysis tools, duration of language sample, synchronous vs. asynchronous communication, public vs. private context, real life vs. lab setting, online vs. offline communication, and written vs. spoken language.

2.3. Strategy of data analysis

The correlation coefficients were converted to the Fisher's z score, averaged, and then converted back to the correlation coefficients based on the instructions from Borenstein, Hedges, Higgins, and Rothstein (2009). We used the random effects model and the metafor R package (Viechtbauer, 2010) to analyze the data. Publication bias was tested by using funnel plots and regression model (Egger, Smith, Schneider, & Minder, 1997). Heterogeneity was tested using Cochran's Q-test (Hedges & Olkin, 1985).

3. Results

3.1. Descriptive of studies

A total of 37 independent samples from 31 articles were included in our meta-analysis (see Table 1 for a descriptive summary). Among them, 34 samples provided the correlation between extraversion and social process words, and 35 samples provided the correlation between extraversion and positive emotion words. A total of 82,132 participants were involved in these samples. Except one sample with 69,792 participants (Kern et al., 2014), the sample size of the rest samples ranged from 8 to 2927 (M = 342.78, SD = 539.78). Age was reported in 31 samples[1] (M = 27.38, SD = 12.82), and the percentage of female participants was reported in 35 samples[2] (M = 59.40%, SD = 16.64%).

3.2. Overall effect size calculation and publication bias

Participants in four samples (sample 9, 15, 16, and 34 in Table 1) completed two language tasks. For example, Mehl (2004) and Mehl et al. (2012) shared the same participants. Mehl (2004) reported the correlation between extraversion and positive emotion words in daily conversations captured by electronically activated recorder (EAR), while Mehl et al. (2012) reported the same correlation in par-

[1] In some studies, some participants did not indicate their age. The mean of age was estimated based on the reported population.

[2] In some studies, some participants did not indicate their gender. The percentage of female was estimated based on the reported population.

Table 1
Characteristics of samples included in the meta-analysis.

Sample	Study	N	Publication status	Language tasks	Extraversion measure(s)	Extraversion reliability
1	Beukeboom et al. (2013)	40	Journal	Oral description of five photos	FFPI	0.93
2	Burdick et al. (2020) 2015 dataset	1353	Journal	Captions of five pictures expressing yourself	BFI	–
3	Burdick et al. (2020) 2016 dataset	705	Journal	Captions of five pictures expressing yourself	BFI	–
4	Gill et al. (2006) study 2 e-mail dataset	105	Journal	E-mail texts to a good friend	Short EPQ-R	–
5	Gill et al. (2006) study 2 blog dataset	71	Journal	Blog entries	IPIP FFPI	–
6	Golbeck, Robles, Edmondson, & Turner (2011)	50 and 260[b]	Conference paper	Twitter status updates	BFI-45	–
7	Golbeck, Robles, & Turner (2011)	167	Conference paper	Facebook profile	BFI-45	–
8	Hall and Caton (2014)	282	Conference paper	Facebook status updates	BFI	–
9[a]	Hawkins et al. (2017) recent dream	629	Journal	Reports about a recent dream	TIPI	–
	Hawkins et al. (2017) important dream			Reports about an important dream		–
10	Hirsh and Peterson (2009)	94	Journal	Writing assignment about past experiences and future goals	BFAS	–
11	Holtgraves (2011)	224	Journal	Latest 20 text messages	TDA	–
12	Kahn et al. (2007) study 3	66	Journal	Verbalizations about feelings following film clips	BFI	0.89
13	Kern et al. (2014)	69,792	Journal	Facebook status updates	IPIP & NEO-PI-R	0.93
14	Krieger (2016)	128	Thesis	Writing assignment about experiences reflecting socially skilled or not socially skilled	NEO-PI-R	–
15[a]	Li and Chignell (2010) blog: commentary journal	8	Journal	Blog entries in commentary journal style	TIPI	–
	Li and Chignell (2010) blog: personal journal			Blog entries in personal journal style		–
16[a]	Mehl (2004)	96	Thesis	Daily conversation captured by electronically activated recorder (EAR)	BFI	0.90
	Mehl et al. (2012)	90	Journal	Stream-of-conscious writing tasks (SOC)		0.90
17	Morgan (2014)	171	Thesis	Online 3-people group chatting in a business simulation game	BFI	–
18	Nowson (2006)	71	Thesis	Personal weblog	IPIP FFPI	–
19	Pennebaker and King (1999)	841	Journal	Class writing assignment	BFI	–
20	Pereira et al. (2019) adolescents with cancer	10	Journal	YouTube video narrative about experiences with cancer	TIPI	–
21	Pereira et al. (2019) adolescents without cancer	10	Journal	YouTube video narrative about experiences with hardship	TIPI	–
22	Pirzadeh and Pfaff (2012)	42	Conference paper	Online 6-members group chatting in a simulation	NEO PI-R™	–
23	Qiu et al. (2012)	142	Journal	Twitter status updates	BFI	0.83
24	Qiu et al. (2017) sample 1	470	Journal	Sina Weibo status updates	BFI	0.69
25	Qiu et al. (2017) sample 2	90	Journal	Sina Weibo status updates	BFI	–
26	Raess (2018)	62	Thesis	Twitter status updates	TIPI	–
27	Sandy (2013)	942	Thesis	Chatting with a stranger online	TIPI	0.74
28	Sumner, Byers, Boochever, and Park (2012)	2927	Conference paper	Twitter status updates	TIPI	–
29	Sumner et al. (2011)	537	Conference paper	Last 25 Facebook status updates	BFI	–
30	Tackman et al. (2020) sample 1	183	Journal	Daily conversation captured by EAR	BFI	0.85
31	Tackman et al. (2020) sample 2	52	Journal	Daily conversation captured by EAR	BFI	0.81
32	Tackman et al. (2020) sample 3	120	Journal	Daily conversation captured by EAR	BFI	0.86
33	Tackman et al. (2020) sample 4	107	Journal	Daily conversation captured by EAR	TIPI	0.72
34[a]	Tassone (2019) study 3 online action	185	Thesis	Writing of reasons about opinions towards four online actions	BFI	0.86
	Tassone (2019) study 3 offline action			Writing of reasons about opinions towards four offline actions		0.86
35	Tighe and Cheng (2018)	250	Conference paper	Twitter status updates	BFI	0.82
36	Williams et al. (2003)	206	Conference paper	One-page free description of reactions to the 9/11 attacks	BFI	–
37	Yarkoni (2010)	694	Journal	Blog status updates	NEO PI-R™ & IPIP-300	–

Note. **BFI-45** = 45-question version of the Big Five Personality Inventory (John, 1999); **BFAS** = Big Five Aspect Scales (DeYoung, Quilty, & Peterson, 2007); **BFI** = Big Five Inventory (John & Srivastava, 1999); **FFPI** = The Five-Factor Personality Inventory (Hendriks, Hofstee, De Raad, & Angleitner, 1995); **IPIP** = International Personality Item Pool (www.ipip.ori.org; Goldberg, 1999; Goldberg et al., 2006); **IPIP FFPI** = International Personality Item Pool Five Factor Personality Inventory (Buchanan, 2001); **IPIP-300** = 300-item IPIP representation of the NEO-PI-R; **NEO PI-R™** = 50-item IPIP (Goldberg et al., 2006); **NEO-PI-R** = NEO personality inventory (Costa & McCrae, 1992b) (240-item); **Short EPQ-R** = Eysenck Personality Questionnaire-Revised short version (Eysenck et al., 1984); **TDA** = 100-item measure using trait descriptive adjectives (Goldberg, 1992); **TIPI** = Ten Item Personality Inventory (Gosling, Rentfrow, & Swann, 2003).

[a] Each participant completed two language tasks.

[b] The sample size reported for the correlation between extraversion and social process words is 50 while that for the correlation between extraversion and positive emotion words is 260.

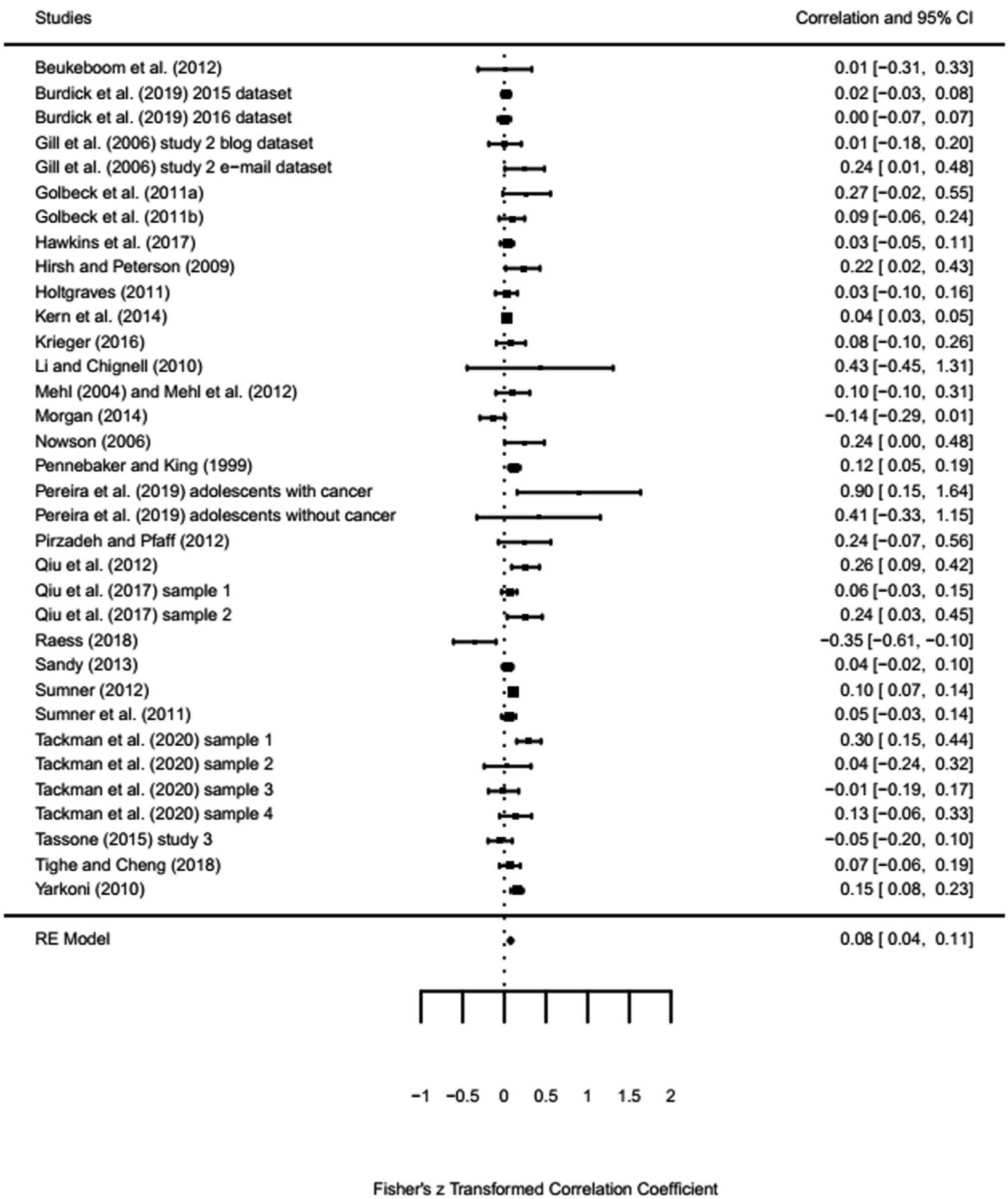

Fig. 1. Forest plot of estimating the overall effect size between extraversion and social process words. *Note.* RE model = random-effects model.

ticipants' stream-of-conscious (SOC) essays. These two correlations were dependent, and simultaneously including them into the meta-analysis might lead to an improper estimate of the overall effect size. Thus, following Borenstein et al. (2009), we computed the mean of correlations and included this synthetic score as the unit of analysis when estimating the overall effect size. For example, the correlation between extraversion and positive emotion words for the EAR sample (r = 0.05) and SOC sample (r = 0.03) were averaged, and 0.04 was included in our estimation of overall effect size for the correlation between extraversion and positive emotion words[3].

The random effects model (see Fig. 1) showed that the mean effect size of the correlation between extraversion and social process words was r = 0.077 (95% CI [0.044, 0.109], p < .001), suggesting a positive relationship between extraversion and social process words. The I^2 was 71.01%, showing a moderate heterogeneity (Higgins & Thompson, 2002), indicating large variation across studies' results. Similarly, the Cochran's Q-test revealed that there was significant heterogeneity, $Q(33)$ = 84.91, p = .001. Publication bias was tested by funnel plot (see Fig. 2) and Egger's regression model (Egger et al., 1997). The funnel plot showed a slightly asymmetrical pattern. Consistently, the Egger's regression model revealed a significant association between effect sizes and standard errors (z = 2.09, p = .04), suggesting the presence of publication bias. The trim-and-fill method (Duval & Tweedie, 2000a, 2000b) was

[3] The sample size of Mehl (2004)'s EAR sample is 96 while the one for Mehl et al., (2012)'s SOC sample is 90. When estimating the overall effect size, the sample size of the merged results was treated as 96.

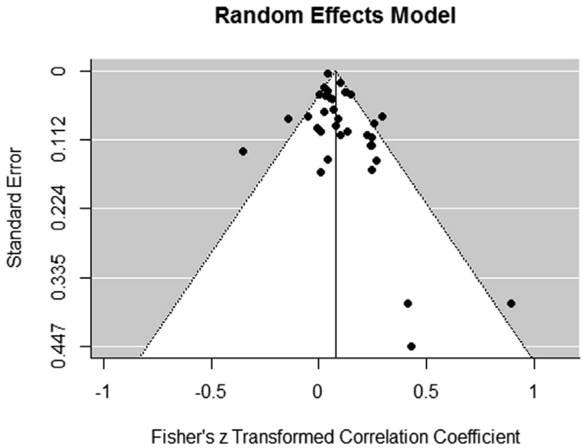

Fig. 2. Funnel plot for the random effects model of the relationship between extraversion and social process words.

used to estimate the number of studies missing from our meta-analysis and the corrected effect size. Results showed that the overall effect size became smaller but still significant (r = 0.060, 95% *CI* [0.022, 0.098]), p < .01), with six studies estimated to be missing on the left side of the funnel plot (see Fig. 3).

Similarly, we performed random effects model (see Fig. 4) to estimate the average effect size of the correlation between extraversion and positive emotion words, and found a small positive relationship between extraversion and positive emotion words (r = 0.069, 95% *CI* [0.041, 0.096], p < .001). The I^2 was 60.05%, suggesting the effects across studies have moderate heterogeneity. There was significant variability of effect size across studies, Q (34) = 93.83, p < .001. Publication bias was tested using funnel plot and the Egger test (Egger et al., 1997). The funnel plot revealed a symmetrical pattern (see Fig. 5) and the Egger test showed non-significant effects (z = −1.28, p = .20). They indicated that no publication bias was present.

3.3. Moderation effects

A series of exploratory analysis were conducted to test nine potential moderators, including gender (percentage of female participants), mean age of participants, duration of language sample, synchronous vs. asynchronous communication, public vs. private setting, real-life vs. lab setting, online vs. offline communication, and written vs. spoken language, and LIWC versions (we renamed "word analysis tools" because all samples used LIWC).

As mentioned above, there are four samples using within-subject design where participants completed two different language tasks (sample 9, 15, 16, and 34 in Table 1). Participants in sample 9, 15, and 34 completed two highly similar tasks. For example, participants in Hawkins, Raymond, and Boyd (2017) (sample 9) reported their recent and important dreams. For these samples, all the moderators were coded the same value for the two language tasks. To avoid inaccurate estimation caused by including dependent data in the analysis, the correlations between extraversion and its two linguistic markers were averaged across tasks for each of the three samples, as well as their moderators. Different from these samples, participants in Mehl (2004) and Mehl et al. (2012) (sample 16) completed two very different language tasks, EAR in Mehl (2004) and SOC in Mehl et al. (2012). EAR and SOC have different characteristics (e.g., EAR was in spoken language while SOC was in written language). It is not typical to average the scores of these moderators especially when they are categorical variables. Thus, we conducted two rounds of moderator testing, with each round including either EAR or SOC. They generated highly similar results (see Table 2). In the following, we only described the results from moderator testing including the EAR sample.

None of the nine moderators significantly moderated the relationship between extraversion and social process words. However, four out of nine moderators significantly moderated the relationship between extraversion and positive emotion words. The correlation was stronger in asynchronous communication (r = 0.084) than synchronous communication (r = −0.002), Q(1) = 6.08, p = .01, which is contrary to our expectation that words in synchronous communication would be more reflective of personality than in asynchronous communication. The moderation effect of public vs. private communication was also significant, Q (1) = 10.56, p = .001; the correlation was stronger in public communication (r = 0.119) than private communication (r = 0.036), consistent with previous findings that extraversion was better expressed in public than private communication (Mehl et al., 2012). The moderation effect of real-life vs. lab setting was significant as well, Q (1) = 5.91, p = .02, showing a stronger correlation in real-life setting (r = 0.107) than lab setting (r = 0.034). Lastly, LIWC version significantly moderated the correlation, Q(3) = 8.71, p = .03. Specifically, LIWC1999 (r = 0.130) showed significantly higher correlation than LIWC2007 (r = 0.062) and LIWC2015 (r = 0.043), and LIWC 2001 (r = 0.075) had significantly higher correlation than LIWC2015 (r = 0.043).

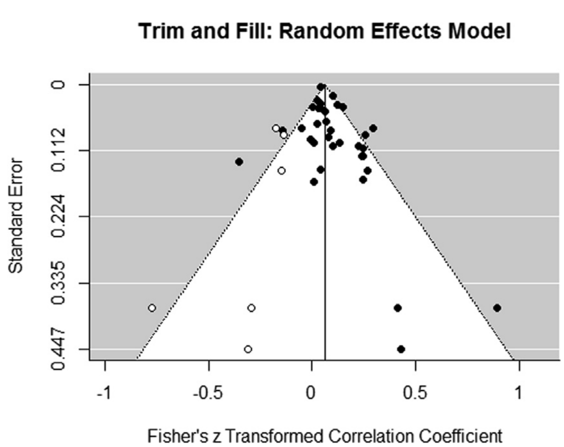

Fig. 3. Funnel plot with filled-in data based on the trim-and-fill method for the random effects model of the relationship between extraversion and social process words.

4. Discussion

Our meta-analysis found that extraversion is related to the use of social process words (r = 0.060) and positive emotion words (r = 0.069). These results show small effect sizes similar to those found in previous meta-analysis on the linguistic markers of individual differences. For example, the linguistic correlates of narcissism all yielded small effects (each $|r|$ < 0.10) (Holtzman et al., 2019). I-words are related to neuroticism with a small correlation of r = 0.10 (Tackman et al., 2018). These findings suggest that linguistic markers may not be strong predictors of personality traits.

Our findings are consistent with the conceptualization of extraversion as a trait that reflects the tendency to experience positive emotion and engage in social interactions (Harari et al., 2019; Watson & Clark, 1997). They show that extraverts tend to use more positive emotion and social process words than introverts. Our moderator analyses suggest that the relationship between extraversion and social process words is consistent across demographic characteristics and communication contexts. None of the

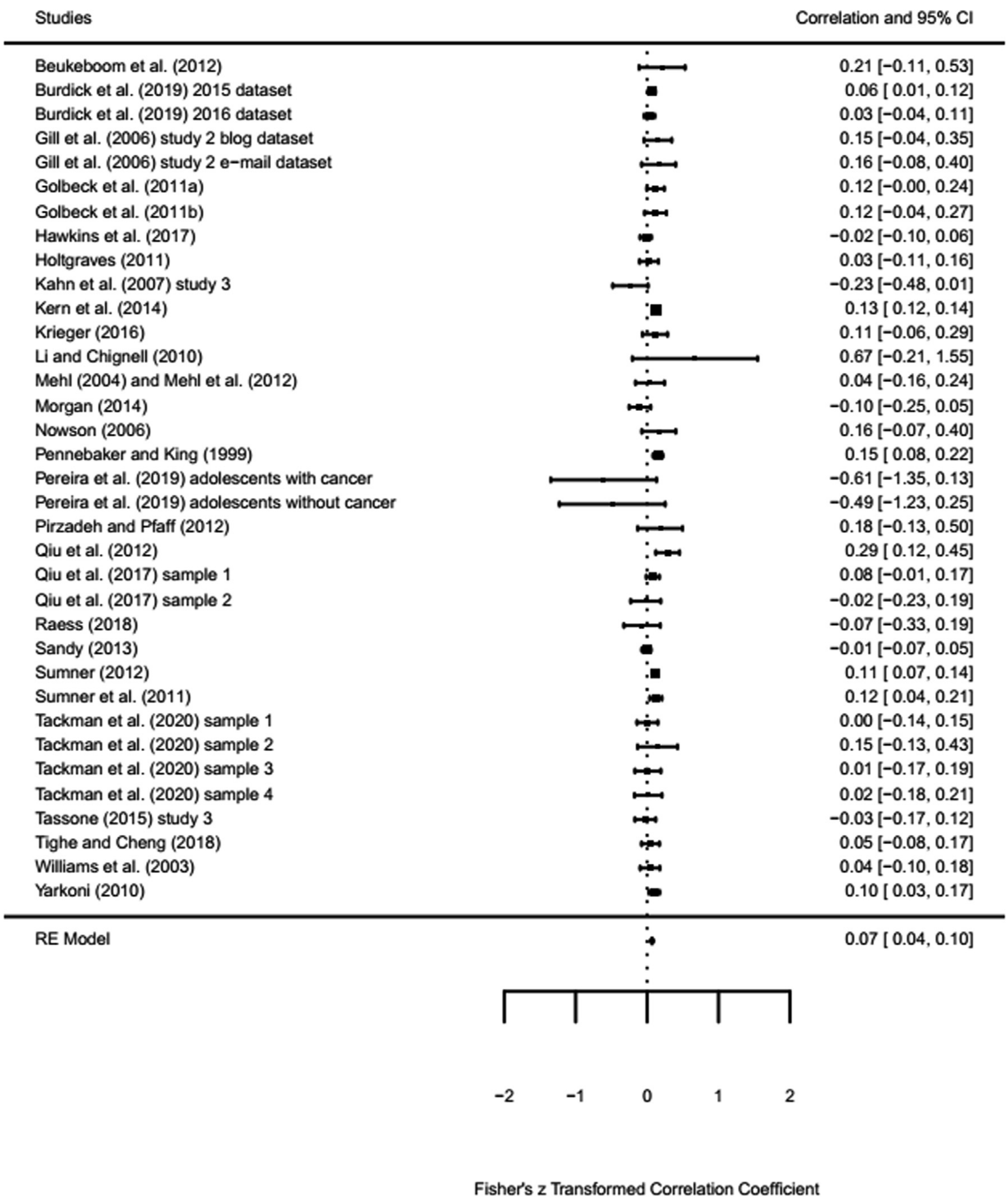

Fig. 4. Forest plot of estimating the overall effect size between extraversion and positive emotion words. *Note.* RE model = random-effects model.

nine moderators had significant effects. We made specific hypothesis for two moderators. The correlation between extraversion and social process words was expected to be stronger in public than private contexts, because traits are better expressed in trait-relevant situations (Tett & Guterman, 2000) and extraversion has been considered a public trait (Mehl et al., 2012). Given that communication in public contexts usually involve more social opportunities than private contexts, we predicted that extraverts are likely to use more social process words in public contexts. However, our results show that extraverts use more social words than introverts regardless of talking to others in private or public, suggesting their general tendency to perseverate about social relationships and activities. We also predicted that the relationship between extraversion and social process words would be stronger in synchronous communication than asynchronous communication, because synchronous communication would involve less impression management and therefore more personality expression. However, our results suggest extraverts tend to use more social process words than introverts in both synchronous and asynchronous communication. One explanation could be that impression management strategies does not affect the use of social process words and therefore does not affect the relationship between extraversion and social process words.

The moderation analyses for the relationship between extraversion and positive emotion words show that the relationship varies according to communication contexts and LIWC versions. The relationship is stronger in public than private contexts, suggesting that extraverts tend to experience and express positive emotions in

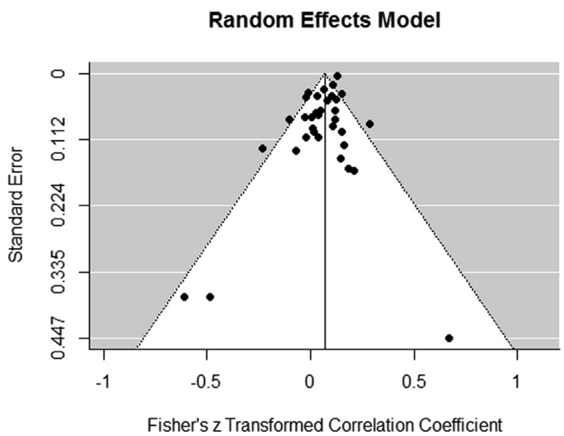

Fig. 5. Funnel plot for the random effects model of the relationship between extraversion and positive emotion words.

public contexts where opportunities for social interactions are common. This is consistent with previous findings that extravert enjoy social activities more than introverts (Oerlemans & Bakker, 2014). We also found that the relationship between extraversion and positive emotion words is stronger in asynchronous than synchronous communication. It could be possible that asynchronous communication allows extraverts to apply impression management strategies where they express more positive emotions to present a better social image. The relationship between extraversion and positive emotion words was also found to be stronger in real-life than lab settings, suggesting that language tasks in lab studies may not provide as many opportunities for extraverts to experience and express positive emotions as real-life communications. Overall, the above results suggest that the amount of positive emotions expressed by extraverts depends on the communication contexts. Finally, LIWC version moderated the relationship between extraversion and positive emotion words, highlighting the possibility that language analysis tools may differ in their

Table 2
Number of effect size (k), correlation (r), 95% confidence interval (95% CI), and significance of moderator analysis (Moderator).

	Extraversion-social process words				Extraversion-positive emotion words			
Variables	k	r	95% CI	Moderator	k	r	95% CI	Moderator
Continuous variables								
percentage of female participants	32			ns	34			ns
mean age of participants	31			ns	30			ns
duration of language sample	13			ns	14			ns
Categorical variables								
LIWC version	34			ns	35			sig
LIWC1999	5	0.093	[0.021, 0.163]	REF	5	0.130	[0.123, 0.138]	REF
	(5)	(0.093)	([0.021, 0.163])		(5)	(0.130)	([0.123, 0.138])	
LIWC2001	6	0.065	[−0.109, 0.235]		8	0.075	[0.025, 0.125]	
	(5)	(0.067)	([−0.161, 0.289])		(8)	(0.074)	([0.024, 0.124])	
LIWC2007	12	0.072	[0.027, 0.115]		11	0.062	[0.004, 0.119]	
	(13)	(0.074)	([0.032, 0.115])		(11)	(0.062)	([0.004, 0.119])	
LIWC2015	11	0.068	[0.000, 0.135]		11	0.043	[0.008, 0.079]	
synchronous vs. asynchronous	34			ns	35			sig
synchronous	8	0.073	[−0.029, 0.175]	REF	8	−0.002	[−0.050, 0.045]	REF
	(7)	(0.076)	([−0.041, 0.191])		(7)	(−0.005)	([−0.054, 0.044])	
asynchronous	26	0.074	[0.044, 0.104]		27	0.084	[0.057, 0.111]	
	(27)	(0.075)	([0.046, 0.105])		(28)	(0.083)	([0.056, 0.110])	
public vs. private	34			ns	35			sig
public	16	0.104	[0.056, 0.153]	REF	16	0.119	[0.100, 0.137]	REF
	(16)	(0.104)	([0.056, 0.153])		(16)	(0.119)	([0.100, 0.137])	
private	18	0.052	[0.009, 0.095]		19	0.036	[−0.001, 0.073]	
	(18)	(0.055)	([0.011, 0.099])		(19)	(0.035)	([−0.002, 0.072])	
real-life vs. lab	34			ns	35			sig
real-life	19	0.100	[0.056, 0.144]	REF	19	0.107	[0.084, 0.129]	REF
	(18)	(0.102)	([0.056, 0.147])		(18)	(0.108)	([0.085, 0.130])	
lab	15	0.041	[0.001, 0.080]		16	0.034	[−0.014, 0.081]	
	(16)	(0.044)	([0.005, 0.083])		(17)	(0.034)	([−0.012, 0.079])	
online vs. offline	34			ns	35			ns
online	19	0.078	[0.035, 0.122]	REF	19	0.089	[0.054, 0.123]	REF
	(19)	(0.078)	([0.035, 0.122])		(19)	(0.089)	([0.054, 0.123])	
offline	15	0.074	[0.021, 0.126]		16	0.040	[−0.003, 0.083]	
	(15)	(0.079)	([0.024, 0.132])		(16)	(0.039)	([−0.004, 0.082])	
language mode	34			ns	35			ns
written	26	0.067	[0.036, 0.097]	REF	26	0.077	[0.048, 0.106]	REF
	(27)	(0.068)	([0.038, 0.098])		(27)	(0.076)	([0.048, 0.105])	
spoken	8	0.131	[0.020, 0.239]		9	0.001	[−0.075, 0.078]	
	(7)	(0.146)	([0.015, 0.272])		(8)	(−0.007)	([−0.089, 0.076])	

Note. Two rounds of moderation testing showed very similar results. The mean correlation for the level of the moderators were presented separately for the two rounds. Specifically, the results including Mehl (2004)'s EAR sample were shown without parentheses while the results including Mehl et al. (2012)'s SOC sample were presented in parentheses.
ns = non-significant; sig = significant; REF = reference category.

implementations and therefore affect the measure of psychological constructs such as positive emption from texts.

In conclusion, our study found that positive emotion words and social process words are linguistic markers of extraversion. However, the two linguistic correlates of extraversion are small in magnitude. In addition, the strength of the relationship between extraversion and positive emotion words varies across communication contexts, while the relationship between extraversion and social process words remain consistent across contexts. With an increasing interest in using Big Data to predict personality (Qiu, Chan, & Chan, 2018), our findings call for future research to explore other linguistic analysis methods and find stronger linguistic predictors of extraversion.

Preregistration

This study was preregistered prior to submission.

Acknowledgements

This work was partial supported by NTU HASS Incentive Scheme Grant 2018 awarded to the second author. The authors would like to thank Pang Lu Woon for her help in the data collection and materials coding section.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrp.2020.104035. Data used for this study could be downloaded at https://researchdata.ntu.edu.sg/dataset.xhtml?persistentId=doi:10.21979/N9/AUAVFQ.

References

**References marked with asterisk are included in the meta-analysis.*

Argyle, M., & Lu, L. (1990). The happiness of extraverts. *Personality and Individual Differences, 11*(10), 1011–1017.

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*(6), 1531.

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. Journal of Personality Assessment, 91(4), 340-345.

Ateş, Ü. (2014). *Inference of personality using social media profiles*. Middle East Technical University.

Aziz, R., Mustaffa, S., Samah, N. A., & Yusof, R. (2014). Personality and happiness among academicians in Malaysia. *Procedia-Social and Behavioral Sciences, 116*, 4209–4212.

Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*(3), 533.

*Beukeboom, C. J., Tanis, M., & Vermeulen, I. E. (2013). The Language of Extraversion: Extraverted People Talk More Abstractly, Introverts Are More Concrete. Journal of Language and Social Psychology, 32(2), 191-201. (Correlation matrix obtained on request from first author)

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Meta-regression* (pp. 190–202). Introduction to meta-analysis. Chichester: Wiley.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*(4), 645.

Buchanan, T. (2001). Online implementation of an IPIP five factor personality 45inventory [web page].

*Burdick, L., Mihalcea, R., Boyd, R. L., & Pennebaker, J. W. (2020). Analyzing connections between user attributes, images, and text. Cognitive Computation, 1-20. (Correlation matrix obtained on request from first author)

Bushnell, P. P. (1930). An analytic contrast of oral with written language. *Teachers College Contributions to Education, 451*, 87–95.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychologica Review, 90*(2), 105.

Bleidorn, W., & Hopwood, C. J. (2019). Using machine learning to advance personality assessment and theory. *Personality and Social Psychology Review, 23*(2), 190–203.

Chafe, W., & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology, 16*(1), 383–407.

Chung, C., & Pennebaker, J. W. (2007). The psychological functions of function words. *Social Communication, 1*, 343–359.

Costa, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653–665.

Costa, P. T., & McCrae, R. R. (1992b). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEOFFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., McCrae, R. R., & Norris, A. H. (1981). Personal adjustment to aging: Longitudinal prediction from neuroticism and extraversion. *Journal of Gerontology, 36*(1), 78–85.

Cuperman, R., & Ickes, W. (2009). Big Five predictors of behavior and perceptions in initial dyadic interactions: Personality similarity helps extraverts and introverts, but hurts "disagreeables". *Journal of Personality and Social Psychology, 97*(4), 667.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124*(2), 197.

Derks, D., Fischer, A. H., & Bos, A. E. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior, 24*(3), 766–785.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89–98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463.

Eaton, L. G., & Funder, D. C. (2001). Emotional experience in daily life: Valence, variability, and rate of change. *Emotion, 1*(4), 413.

Edwards, T., & Holtzman, N. S. (2017). A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality, 68*, 63–68. https://doi.org/10.1016/j.jrp.2017.02.005.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634.

Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology, 82*(5), 804.

Eysenck, H. J. (Ed.). (1981). *A model for personality*. New York: Springer-Verlag.

Eysenck, S. B., Barrett, P., & Eysenck, H. J. (1984). *Eysenck personality questionnaire-revised*. Pergamon Press.

Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*(6), 1011.

*Gill, A. J., Nowson, S., & Oberlander, J. (2006). Language and Personality in Computer-Mediated Communication: A cross-genre comparison. *Journal of Computer Mediated Communication*.

Gill, A. J., Nowson, S., & Oberlander, J. (2009). What Are They Blogging About? Personality, Topic and Motivation in Blogs. Paper presented at the ICWSM.

*Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011). Predicting personality from twitter. Paper presented at the 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing. (The correlation between extraversion and positive emotion words is obtained on request from first author)

*Golbeck, J., Robles, C., & Turner, K. (2011, May). Predicting personality with social media. In CHI'11 extended abstracts on human factors in computing systems (pp. 253-262). ACM.

Goldberg, L. R. (1990). An alternative" description of personality": The big-five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe, 7*(1), 7–28.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*(1), 84–96.

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science, 333*(6051), 1878–1881.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr, (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528.

Gruber, J., & Kring, A. M. (2008). Narrating emotional events in schizophrenia. *Journal of Abnormal Psychology, 117*(3), 520.

*Hall, M. A., & Caton, S. (2014). A crowdsourcing approach to identify common method bias and self-representation.

*Hawkins, I., Raymond, C., & Boyd, R. L. (2017). Such stuff as dreams are made on: Dream language, LIWC norms, and personality correlates. Dreaming, 27(2), 102. (Correlation matrix obtained on request from first author).

Harari, G. M., Müller, S. R., Stachl, C., Wang, R., Wang, W., Bühner, M., & Gosling, S. D. (2019). Sensing sociability: Individual differences in young adults' conversation, calling, texting and app use behaviors in daily life. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000245.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hendriks, A. A. J., Hofstee, W. K. B., De Raad, B., & Angleitner, A. (1995). *The Five-Factor Personality Inventory (FFPI)*. Groningen, Netherlands: University of Groningen.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558.
*Hirsh, J. B., & Peterson, J. B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality, 43*(3), 524–527. https://doi.org/10.1016/j.jrp.2009.01.006.
*Holtgraves, T. (2011). Text messaging, personality, and the social context. *Journal of Research in Personality, 45*(1), 92–99.
Holtzman, N. S., Tackman, A. M., Carey, A. L., Brucks, M. S., Küfner, A. C., Deters, F. G., ... Mehl, M. R. (2019). Linguistic markers of grandiose narcissism: A LIWC analysis of 15 samples. *Journal of Language and Social Psychology, 38*(5–6), 773–786.
John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The big five inventory—versions 4a and 54. In: Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research, 2*(1999), 102–138.
*Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. *The American Journal of Psychology*, 263–286.
Karan, A., Rosenthal, R., & Robbins, M. L. (2019). Meta-analytic evidence that we-talk predicts relationship and personal functioning in romantic couples. *Journal of Social and Personal Relationships, 36*(9), 2624–2651.
*Kern, M. L., Eichstaedt, J. C., Schwartz, H. A., Dziurzynski, L., Ungar, L. H., Stillwell, D. J., ... Seligman, M. E. (2014). The online social self: An open vocabulary approach to personality. *Assessment, 21*(2), 158–169.
Kramer, A. D. (2010). An unobtrusive behavioral model of gross national happiness. *Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems*.
Kramer, A. D., Fussell, S. R., & Setlock, L. D. (2004). Text analysis as a tool for analyzing conversation in online support groups. Paper presented at the CHI'04 Extended Abstracts on Human Factors in Computing Systems.
*Krieger, K. (2016). Words of Well-Being: The Relation of an Individual's Word Choice to their Social Well-Being.
Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology, 61*(1), 132.
Lee, K., & Ashton, M. C. (2008). The HEXACO personality factors in the indigenous personality lexicons of English and 11 other languages. *Journal of Personality, 76*(5), 1001–1054.
*Li, J., & Chignell, M. (2010). Birds of a feather: How personality influences blog writing and reading. *International Journal of Human-Computer Studies, 68*(9), 589–602.
Levenson, R. (1992). *Oedipus Text [Computer software]*. Berkeley, CA: University of California.
Lin, C.-C. (2014). A higher-order gratitude uniquely predicts subjective well-being: Incremental validity above the personality and a single gratitude. *Social Indicators Research, 119*(2), 909–924.
Liu, P., Chan, D., Qiu, L., Tov, W., & Tong, V. J. C. (2018). Effects of cultural tightness–looseness and social network density on expression of positive and negative emotions: A large-scale study of impression management by Facebook users. *Personality and Social Psychology Bulletin, 44*(11), 1567–1581.
Lucas, R. E., & Baird, B. M. (2004). Extraversion and emotional reactivity. *Journal of Personality and Social Psychology, 86*(3), 473.
Lucas, R. E., Diener, E., Grob, A., Suh, E. M., & Shao, L. (2000). Cross-cultural evidence for the fundamental features of extraversion. *Journal of Personality and Social Psychology, 79*(3), 452.
McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81.
McNiel, J. M., & Fleeson, W. (2006). The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality, 40*(5), 529–550.
*Mehl, M. R. (2004). The sounds of social life: Exploring students' daily social environments and natural conversations.
Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90*(5), 862–877 https://doi:10.1037/0022-3514.90.5.862.
*Mehl, M. R., Robbins, M. L., & Holleran, S. E. (2012). How taking a word for a word can be problematic: Context-dependent linguistic markers of extraversion and neuroticism. Journal of Methods and Measurement in the Social Sciences, 3(2), 30-50. (The correlation between extraversion and social process words is obtained on request from first author).
*Morgan, J. L. (2014). *Effects of personality, communication, and cross-training on virtual team performance*. The University of Alabama in Huntsville.
Nguyen, M., Bin, Y. S., & Campbell, A. (2012). Comparing online and offline self-disclosure: A systematic review. *Cyberpsychology, Behavior, and Social Networking, 15*(2), 103–111.
*Nowson, S. (2006). The Language of Weblogs: A study of genre and individual differences.
Oberlander, J., & Gill, A. J. (2006). Language with character: A stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes, 42*(3), 239–270.
Oerlemans, W. G., & Bakker, A. B. (2014). Why extraverts are happier: A day reconstruction study. *Journal of Research in Personality, 50*, 11–22.
Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., ... Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*(6), 934.
Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *LIWC2007: Linguistic inquiry and word count*. Austin, Texas: liwc. net.
Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015.
Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition & Emotion, 10*(6), 601–626.
Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. Mahway: Lawrence Erlbaum Associates, 71(2001), 2001.
*Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*(6), 1296.
Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*(4), 863.
Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*(1), 547–577.
Pennebaker, J. W., & Stone, L. D. (2003). Words of wisdom: Language use over the life span. *Journal of Personality and Social Psychology, 85*(2), 291–301.
*Pereira, L., Sampson, J., & DiCola, K. (2019). Factors Related to Linguistic Content in Video Narrative of Adolescents with Cancer and Healthy Controls. *Journal of Psycholinguistic Research, 48*(5), 1185–1201.
*Pirzadeh, A., & Pfaff, M. S. (2012). Emotion expression under stress in instant messaging. *Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.
Qiu, L., Chan, S. H. M., & Chan, D. (2018). Big data in social and psychological science: Theoretical and methodological issues. *Journal of Computational Social Science, 1*(1), 59–66.
*Qiu, L., Lin, H., Ramsay, J., & Yang, F. (2012). You are what you tweet: Personality expression and perception on Twitter. *Journal of Research in Personality, 46*(6), 710–718.
*Qiu, L., Lu, J., Ramsay, J., Yang, S., Qu, W., & Zhu, T. (2017). Personality expression in Chinese language use. *International Journal of Psychology, 52*(6), 463–472.
*Raess, M. L. (2018). Interactions3: language, demographics and personality: an in-depth analysis of German tweets.
Ritter, R. S., Preston, J. L., & Hernandez, I. (2014). Happy tweets: Christians are happier, more socially connected, and less analytical than Atheists on Twitter. *Social Psychological and Personality Science, 5*(2), 243–249.
Salsman, J. M., Brown, T. L., Brechting, E. H., & Carlson, C. R. (2005). The link between religion and spirituality and psychological adjustment: The mediating role of optimism and social support. *Personality and Social Psychology Bulletin, 31*(4), 522–535.
*Sandy, C. J. (2013). Predicting accuracy in first impressions based on language use in computer-mediated communication environments.
Sillars, A., Shellen, W., McIntosh, A., & Pomegranate, M. (1997). Relational characteristics of language: Elaboration and differentiation in marital conversations. *Western Journal of Communication (includes Communication Reports), 61*(4), 403–422.
Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*(1), 117.
Stone, L. D., & Pennebaker, J. W. (2002). Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana. *Basic and Applied Social Psychology, 24*(3), 173–183.
*Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. *Paper presented at the 2012 11th International Conference on Machine Learning and Applications*.
*Sumner, C., Byers, A., & Shearing, M. (2011). Determining personality traits & privacy concerns from facebook activity. *Black Hat Briefings, 11*, 197–221.
Tackman, A. M., Sbarra, D. A., Carey, A. L., Donnellan, M. B., Horn, A. B., Holtzman, N. S., ... Mehl, M. R. (2018). Depression, negative emotionality, and self-referential language: A multi-lab, multi-measure, and multi-language-task research synthesis. *Journal of Personality and Social Psychology, 116*(5), 817.
*Tackman, A. M., Baranski, E. N., Danvers, A. F., Sbarra, D. A., Raison, C. L., Moseley, S. A., ... Mehl, M. R. (2020). 'Personality in Its Natural Habitat'Revisited: A Pooled, Multi-sample Examination of the Relationships Between the Big Five Personality Traits and Daily Behaviour and Language Use. *European Journal of Personality*.. https://doi.org/10.1002/per.2283.
Tan, C.-S., & Lee, Q.-W. (2017). The role of self-esteem and social support in the relationship between extraversion and happiness: A serial mediation model. *Current Psychology, 36*(3), 556–564.
Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24–54.
Tellegen, A. (1985). Structures of mood and personality and their rele-vance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. Mason (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.
Tellegen, A., Lykken, D. T., Bouchard, T. J., Wilcox, K. J., Segal, N. L., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology, 54*(6), 1031.

Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality, 34*(4), 397–423.

Tassone, A. (2019). The Relationship Between Characteristics of Collective Action, Introversion/Extraversion, and Collective Action Endorsement.

*Tighe, E., & Cheng, C. (2018). Modeling personality traits of filipino twitter users. Paper presented at the Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media. (Correlation matrix obtained on request from first author).

Tobin, R. M., Graziano, W. G., Vanman, E. J., & Tassinary, L. G. (2000). Personality, emotional experience, and efforts to control emotions. *Journal of Personality and Social Psychology, 79*(4), 656.

Tov, W., Ng, K. L., Lin, H., & Qiu, L. (2013). Detecting well-being via computerized content analysis of brief diary entries. Psychological Assessment, 25(4), 1069–1078.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23*(1), 3–43.

Warner, R. M., & Vroman, K. G. (2011). Happiness inducing behaviors in everyday life: An empirical assessment of "the how of happiness". *Journal of Happiness studies, 12*(6), 1063–1082.

Watson, D., & Clark, L. A. (1997). Extraversion and its positive emotional core. In *Handbook of Personality Psychology* (pp. 767–793). Elsevier.

*Williams, K. M., Paulhus, D. L., & Nathanson, C. (2003). Personality and Political Predictors of Emotional Reactions to 9/11. *Paper presented at the Poster presented at the 83rd annual convention of the Western Psychological Association*.

Wilt, J., & Revelle, W. (2009). Extraversion.

*Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44*(3), 363–373. https://doi.org/10.1016/j.jrp.2010.04.001.

Youyou, W., Stillwell, D., Schwartz, H. A., & Kosinski, M. (2017). Birds of a feather do flock together: Behavior-based personality-assessment method reveals personality similarity among couples and friends. *Psychological Science, 28*(3), 276–284.

Zimmermann, J., Brockmeyer, T., Hunn, M., Schauenburg, H., & Wolf, M. (2017). First person pronoun use in spoken language as a predictor of future depressive symptoms: Preliminary evidence from a clinical sample of depressed patients. *Clinical Psychology & Psychotherapy, 24*(2), 384–391.