PERSONALITY-DRIVEN SOCIAL MEDIA CURATION:

HOW PERSONALITY TRAITS AFFECT

FOLLOWING DECISIONS ON TWITTER

by

CIANNA BEDFORD-PETERSEN

A DISSERTATION

Presented to the Department of Psychology
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2022

DISSERTATION APPROVAL PAGE

Student: Cianna Bedford-Petersen

Title: Personality-Driven Social Media Curation: How Personality Traits Affect Following Decisions on Twitter

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Psychology by:

Sanjay Srivastava	Chairperson
Sara J. Weston	Core Member
Elliot Berkman	Core Member
Ryan Light	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded December 2022

DISSERTATION ABSTRACT

Cianna Bedford-Petersen

Doctor of Philosophy

Department of Psychology

December 2022

Title: Personality-Driven Social Media Curation: How Personality Traits Affect Following Decisions on Twitter

As social media occupies an increasingly important place in people's lives, new opportunities are presented for people to select and modify their online environments. On many platforms, users have significant control over what kind of information and experiences they are exposed to. For example, on Twitter, virtually everything users see is a function of their decisions about what accounts to follow. What drives those decisions? My dissertation explores the extent to which personality is reflected in our social media environment by examining the relationship between personality traits and the accounts that users follow on Twitter. Particularly, what features of accounts influence following decisions and how personality traits of users align with characteristics of Twitter accounts. Exploring the relationship between who we are and the decisions we make online provides a better understanding of how characteristics, such as personality traits, drive the curation of our social environments.

Overall, findings indicate that personality does influence the decisions we make about which Twitter accounts to follow and in turn, how our social media environment is curated. The strength and stability of this relationship shows some heterogeneity across traits, though is generally comparable to the effect of some commonly used demographic variables. Personality traits of users also align with characteristics of Twitter accounts and moderate the effect of different Twitter profile features on our following decisions, highlighting potential psychological

processes that drive following decisions. For example, extraverts want to feel connected to popular accounts and seek content on topics that lots of other people care about while Neuroticism is associated with following accounts that conform to gender and age norms. Perhaps most notably, these relationships demonstrate remarkable generalizability when tested on a set of real-world followed accounts. Though this research is a first step in exploring the influence of personality on the vast number behaviors that occur on social media, these findings establish foundational knowledge and inform future research.

CURRICULUM VITAE

NAME OF AUTHOR: Cianna Bedford-Petersen

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
St. Olaf College, Northfield, MN

DEGREES AWARDED:

Doctor of Philosophy, Psychology, 2022, University of Oregon
Master of Science, Psychology, 2019, University of Oregon
Bachelor of Arts, Psychology, 2014, St. Olaf College

AREAS OF SPECIAL INTEREST:

Online Social Networks
Personality & Health
Data Science & R Programming
Earned *Specialization in Educational Data Science* from the Educational Leadership Department, University of Oregon, Eugene, Oregon
Reproducible Research

PROFESSIONAL EXPERIENCE:

Quantitative User Experience Research Intern, Meta, 2021
R Programming Consultant, NWEA, 2021
Graduate Employee, University of Oregon, 2017-2022
Laboratory Manager, University of Minnesota, 2016-2017
Digital Analyst, Collective Measures Agency, Minneapolis, MN, 2014-2015

GRANTS, AWARDS, AND HONORS:

Graduate Registration Award, Society for Social & Personality Psychology, 2020
Harvey E Lee Graduate Scholarship, University of Oregon, 2020
Graduate Travel Award, Society for Social & Personality Psychology, 2018
First Year Graduate Research Fellowship, University of Oregon, 2017
Magna Cum Laude, St. Olaf College, 2014
Distinction in Psychology, St. Olaf College, 2014
NHLBI Scholarship, Summer Training in Biostatistics, Boston University, 2013
Allport Award in Psychology Finalist, St. Olaf College, 2013
Presidential Academic Scholarship, St. Olaf College, 2010
Cassler Music Scholarship, St. Olaf College, 2010

PUBLICATIONS:

Bedford-Petersen, C., & Weston, S. J. (2021). Mapping Individual Differences on the Internet: Case Study of the Type 1 Diabetes Community. *JMIR diabetes*, *6*(4), e30756.

Hughes, B. T., Costello, C. K., Pearman, J., Razavi, P., Bedford-Petersen, C., Ludwig, R. M., & Srivastava, S. (2021). The Big Five Across Socioeconomic Status: Measurement Invariance, Relationships, and Age Trends. *Collabra: Psychology*, *7*(1), 24431.

Bedford-Petersen, C., & Saucier, G. (2021). Identifying contrasting themes that orchestrate personality expression across situations. *Personality and Individual Differences*, *171*, 110495.

Bedford-Petersen, C., DeYoung, C. G., Tiberius, V., & Syed, M. (2019). Integrating philosophical and psychological approaches to well-being: The role of success in personal projects. *Journal of Moral Education*, *48*(1), 84-97.

ACKNOWLEDGMENTS

I'd like to start by thanking my advisor, Sanjay Srivastava. Without your support during my time at the University of Oregon, I wouldn't be where I am today. I am grateful for your keen guidance as I navigated graduate school and open-mindedness as I explored career paths. I would like to thank my committee members Sara Weston, Elliot Berkman, and Ryan Light for their helpful feedback on this dissertation. Particularly Sara for her statistical and R guidance over the past few years which helped prepare me for the challenges I faced in this research.

I'd also like to thank the members of the Personality and Social Dynamics Lab for their support over the years, listening to countless practice talks and providing thoughtful feedback. Especially Cory Costello and Pooya Razavi, who were instrumental in data collection for Study 3 and the graduate student coders who helped evaluate Twitter profiles - Alexis Adams-Clark, Andrew Fridman, Bradley Hughes, Rachel Jacobson, and Pooya Razavi.

A huge thank you to my cohort mates and friends. From weekly trivia nights to just checking in, your support got me through the highs and lows of graduate school. Thank you to my parents, David and Shilon Bedford, for fostering my curious mind and channeling that strong-willed nature into something productive. Thank you to Bret and Karen Petersen for your support from afar and sending coffee money to fuel my writing.

Most importantly, I could not have finished the dissertation without the support of my little family. To my dog Fritz and my cat Freyja for providing comic relief, reminding me to take breaks, and being my late-night study buddies. Finally, my biggest thank you to my husband Connor. From your willingness to drop everything and move across the country for this degree to your extra help this year keeping our household together while I wrote, your support over the past 10 years has gotten us where we are today. I can't wait for our next adventure together.

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

Figure | Page

LIST OF TABLES

Table Page

Table Page

I. INTRODUCTION

The last decade has seen explosive growth of online social networks, which are now a pervasive part of everyday life for many people. While in 2005, just 5% of American adults had used a social media platform, that number has risen to almost 72% of the public using some form of social media in 2021. For many, checking social media is part of a daily routine with 70% of Facebook users checking their feed every day and around 50% of Twitter, Instagram, Snapchat, and YouTube users reporting daily use (Pew Research Center, 2021). This has led to a growing number of social interactions and behaviors occurring in online environments, increasing the digitization of our social world.

A primary goal of social media companies is to keep users continually engaged with their platform. As a result, the underlying architecture of these sites is designed to learn from how users engage and adapt accordingly (e.g., Likes, follows, comments). This design presents a unique opportunity to disentangle the relationship between who we are and our online behaviors. In particular, personality psychologists are interested in exploring how people select and modify their environment. Better understanding the relationship between who we are and the decisions we make online can provide insight into how characteristics, such as personality traits, drive the curation of our social environments.

Twitter provides an ideal online laboratory to study these interactions because a user's behavior on the site directly influences their experience on the platform. Specifically, the action of following accounts shapes what a user sees on their feed, which serves as the center of social interaction on Twitter. Additionally, unlike other social media platforms, Twitter users can directly select the accounts they wish to have incorporated into their feed without the requirement of a reciprocal follow. In turn, the social environment created by following accounts is a direct product of user decisions, providing a more robust reflection of user preference than other platforms.

Beyond the interests of personality psychologists, research connecting personality to the following decisions people make could provide valuable findings relevant to the aims of social media companies. As a part of keeping users engaged with their platform, these companies are interested in learning more about the varied experiences of their users. For example, they may seek an explanation for why some users feel positive emotions when viewing content on their platform while others experience social comparison and negative body image. Research within social media companies frequently focuses on the contribution of demographic categories such as age or gender. However, better understanding user personality in relation to curated social environments could be a helpful tool for combating negative experiences users have on their platform.

To what extent are our personalities reflected in the environments we create for ourselves online? In this dissertation, I will take an exploratory approach to better understanding the relationship between personality and the accounts that we choose to follow on Twitter by examining the following three questions: How strong is the association between personality and interest in following Twitter accounts (Studies 1 and 2)? What features of Twitter accounts are people with different personality traits drawn to (Study 2)? And do those associations between personality and account features generalize from the lab to real-world following decisions (Study 3)?

Psychological Characteristics and Online Behaviors

A unique byproduct of increasing online social interactions is that much of our behavior is recorded. This can include deliberately shared information such as social media posts in addition to more passive actions such as "Liking" or saving content. Taken together, the collection of our online actions makes up a relatively permanent record known as a *digital footprint*. The data from these footprints has led to new insights on both population-wide patterns of social interaction as well as detailed characteristics of individuals (Lambiotte & Kosinski, 2014).

One of the major insights offered by online social data research relates to the strength of associations between individuals' psychological traits and their digital footprints. This research has uncovered notable relationships between personality traits, sexual orientation, or mental health and social media behaviors such as "Likes" on Facebook (Kosinski et al., 2013; Youyou et al., 2015), language use on Facebook (Park et al., 2015; Schwartz et al., 2013) and Twitter (Costello, 2020; Arnoux et al., 2017; Qiu et al., 2012), and network metrics on Twitter (Golbeck et al., 2011; Quercia et al., 2011). The magnitude of relationships between personality and these online behaviors typically ranges from $\bar{r} = .29$ for Agreeableness to $\bar{r} = .40$ for Extraversion (Azucar et al., 2018), demonstrating a large effect size in the context of psychological research (Funder & Ozer, 2019).

Associations have also been found between who we are and who we follow on Twitter. Costello et al. (2021) examined the extent to which the accounts a person follows on Twitter can be used to predict self-reported anxiety, depression, post-traumatic stress, and anger. The results suggested that each of the four constructs can be predicted with modest accuracy (r*'s* ≈ .2). Costello et. al (2022) studied the extent to which personality is revealed in online environments using several machine learning techniques to predict personality traits based on followed Twitter accounts. The findings indicate that personality can be inferred with some degree of accuracy (*r's* = .07 for Agreeableness to .45 for Openness) from followed accounts, suggesting personality is reflected in network connections.

Though an association between personality and Twitter followed accounts has been established, it is unknown whether personality is prospectively driving following decisions rather than being shaped by them. Studies thus far examining relationships between psychological characteristics and social media behavior rely on already existing digital records, preventing conclusions about directionality from being drawn. While naturalistic studies allow us to see broad associations, these designs limit our understanding of the extent to which personality is reflected in the environments we construct for ourselves online. An aim of the present research is to shift exploration of this relationship from association to prospective prediction, using more controlled laboratory methods to uncover the extent to which personality shapes the environments people construct for themselves online. Specifically, I will examine whether personality traits predict interest in following different accounts (e.g., whether some accounts appeal more to extraverts than to introverts, to emotionally stable people more than neurotic people, etc.). Collecting personality metrics prior to hypothetical decisions about following Twitter accounts will allow me to distill the directional influence of personality on those social media preferences. If these relationships are found, I will further explore what features of the followed Twitter accounts could explain those associations. As a relatively new area of research, there is a lack of strong theory to guide testing of specific hypotheses. Instead, descriptive work in this area will provide important and foundational information for development of theory

(Rozin, 2009). To do this, I will conduct data-driven and exploratory analyses to look for patterns in findings that will help future research move in a more focused direction.

Personality and the Social Environment

An interest of personality psychologists is understanding how people select and modify their environment. Theories of personality development such as the corresponsive principle assert that life experiences are at least partly predictable by characteristics of the individual person, such as personality traits. A component of this theory is personality-based selection of environments, indicating that people actively seek out situations and roles based on their personality traits (Harms, 2020; Roberts & Nickel, 2021). For example, Roberts et al. (2003) found that those high in social potency (dominance and persuasiveness) tended to select into jobs that involve power and leadership.

A similar process that also focuses on selection effects is known as person-situation transactions. In this transaction, individuals selectively engage with situations that encourage and reinforce the expression of their own traits and attributes (Buss, 1987; Ickes et al., 1997). As early as 1937, Allport noted that people can play an active role in seeking out environments that correspond with their characteristics (Allport, 1937). Specifically, research on this process has demonstrated notable relationships between personality traits and choices related to our environment. For example, it has been shown that people high in Extraversion tend to seek out more stimulating situations while those low in Extraversion prefer more passive recreational situations. Similarly, people high in Extraversion were more likely to select situations that involve competition and assertiveness (Furnham, 1981).

Twitter as a Social Environment. The social media platform Twitter provides a unique environment in which to study these person-situation transactions. When users log into Twitter, an endless stream of Tweets appear known as the "feed" or "timeline." Users are able to scroll through Tweets posted by other Twitter accounts and engage with those posts in the form of likes, comments, and Retweets. This feed serves as Twitter's primary social center and the majority of user interaction occurs here. The content that a user sees on their feed is largely driven by the accounts they have selected to follow. Relative to other online experiences, the user has a lot of control to curate their social environment, as the feed they see is a direct result of their behavioral decisions.

A distinguishing element of Twitter is that network connections are directional. This implies that a user can initiate an outgoing tie (known as a "follow" or "friend") and can receive an incoming tie (known as a "follower"). However, these are separate actions and follows are not inherently reciprocal (unlike Facebook and some other platforms). This means that a user can freely select the accounts they wish have incorporated into their feed without the need for a reciprocal follow[1]. As a result, the social environment created by following accounts is a direct product of user choice, providing an unrestricted reflection of user preference in comparison to other platforms.

The accounts users choose to follow can reflect experiences people are seeking to have on the platform. A person may follow celebrity accounts if they are interested in hearing their thoughts and opinions, news accounts if they want to learn more about current world events, or beauty influencer accounts if they want to keep up with makeup trends. Choosing which accounts to follow in turn creates a personalized, curated experience on Twitter reflecting a

[1] Twitter accounts can be set to private, which requires account owner approval to follow, though this makes up only around 11% of Twitter accounts (McClain et al., 2021).

user's interests, goals, and broader values. Past research has shown that these attributes are associated with broad personality domains like the Big Five (Rentfrow & Gosling, 2003; Roberts & Robins, 2000; Roccas et al., 2002). For example, someone high in Openness may seek out intellectually stimulating content, while a person low in Agreeableness may follow accounts that post about controversial topics or stimulate a lot of discussion. Better understanding the relationship between traits and the types of accounts followed can provide insight into how personality drives the construction of our social media environments.

Variation Across Traits

In exploring the relationship between personality traits and Twitter following decisions, it is likely that some traits will be better predictors of followed accounts than others. A meta-analysis from Azucar et al. (2018), used 16 independent studies to examine the association between social media digital footprints and each of the Big 5 personality traits. The estimated meta-analytic correlations were 0.40 for Extraversion, 0.39 for Openness, 0.35 for Conscientiousness, 0.33 for Neuroticism, and 0.29 for Agreeableness. Costello et. al's (2022) examination of the relationship between personality traits and Twitter followed accounts found the strongest association with Openness (r = .45) while moderate relationships were found with Neuroticism (r = .30), Extraversion (r = .28), and Honesty (r = .24). However, little to no relationship was found with the traits of Conscientiousness (r = .13) and Agreeableness (r = .07). One possible explanation for this variation is that differing reasons for using Twitter may make some traits stronger predictors of online behaviors than others. For example, if someone wants to follow accounts that are consistent with their own emotions, then Extraversion (positive affect) and Neuroticism (negative affect) may be strong predictors. Alternatively, if they're using twitter to reduce negative emotions, you might also expect associations with Neuroticism, but different ones (e.g., neuroticism being associated with soothing or comforting content, rather than affectively negative).

Personality Traits Relative to Other Variables. Another important consideration when exploring the effect of personality on Twitter followed accounts, is how the influence of personality compares to other analogous variables. Previously, the ability of personality traits to predict important life outcomes has been called into question due to small effect sizes (Mischel, 1968). This idea was explored in Roberts et. al (2007) where the power of personality to predict mortality, divorce, and occupational attainment was compared to the predictive influence of socioeconomic status (SES) and cognitive ability. Their findings showed that the effect of personality traits on these important life outcomes was of the same magnitude as the effects of SES and cognitive ability. These results demonstrate the predictive power of personality and give us reason not only to explore the impact of personality on followed Twitter accounts, but also to compare this effect to other similar variables.

To consider what variables might afford an appropriate comparison it is important to think about the level of measurement being made with Big 5 personality traits. One way to conceptualize this is considering the tradeoff between bandwidth and fidelity (Cronbach & Gleser, 1965; Shannon & Weaver, 1949). This concept states that broad, global constructs can be used to predict many different outcomes with moderate validity; and narrow, specific constructs can be used to predict a few outcomes with higher validity. Personality traits as measured by the domain level of the Big 5 are considered to be a measurement high in bandwidth and lower in fidelity. As a measure, it is not the best way to predict every single behavior a person will do, but it can predict a wide range of things, at a very general level. As this is an exploratory investigation, measuring personality at the domain level of the Big Five provides broad coverage with a relatively manageable number of variables. Likewise, I can consider similarly broad and high bandwidth demographic variables such as gender, socioeconomic status, and political orientation appropriate comparison variables when considering influence of personality on Twitter followed accounts.

Twitter Account Following Decisions

Why are people drawn to follow certain accounts? Beyond establishing the strength of the relationship between personality and Twitter followed accounts, a major goal of this dissertation is to better understand what drives those following decisions in relation to personality traits. In taking an exploratory approach to understanding broad patterns of interest in following accounts, it is useful to draw inspiration from theories and findings from both social media and personality research.

Social media scholars frequently draw on qualitative data to identify motivators of social media behaviors, framing motives in the language of participants' own understandings rather than in psychological theories of underlying processes. Whiting and Williams (2013) conducted 25 in-depth interviews to gain a better understanding of why people use social media. Their findings identified 10 motivators for using social media which included: social interaction, information seeking, pass time, entertainment, relaxation, communicatory utility, convenience utility, expression of opinion, information sharing, and knowledge about others. While these findings do not directly address Twitter following behaviors, they provide some clues for potential motivators that intersect with psychological research.

From psychology and other behavioral sciences, there are a number of theories about what motivates us to pursue behaviors and how they relate to our personality. These theories address reasons or ways we might pursue interest and seek information but haven't been thoroughly applied to social media contexts. For example, Twitter offers vast opportunities to follow accounts related to personal interests. Informed by theory and research on person-situation transactions, several studies have examined associations between interests and personality in other settings. Rentfrow & Gosling (2003) found that people who are more creative and seek out intellectual stimulation prefer music that is unconventional and complex, and those who are sociable and enthusiastic prefer styles of music that are energetic. The connection between personality and interests has also been demonstrated in clothing preferences (Rosenfeld & Plax, 1977; Sharma, 1980), room decoration (Gosling et al., 2002), and even vocational interests (Tokar et al., 1998). To better understand whether or how personality is associated with interests that drive following decisions on Twitter, I will examine the primary topics of Twitter profiles that people want to follow and their relationship to personality traits.

Social connection and interaction are additional primary functions of social media. How might personality be associated with the ways people seek social connection? Research on homophily suggests that people prefer and seek to connect with those who are similar to themselves. Homophily has been demonstrated for individual differences in emotion (Anderson et al., 2003; Watson et al., 2000), as well as social network connections. Youyou et al. (2015) used both Facebook Likes and language to measure personality and study similarity among romantic partners and friends. Strong similarity was detected between romantic partners and between friends in both Openness and Extraversion, indicating that personality as expressed through social media behaviors may attract followers with similar features. People seek these similarities not only in human social connections but also in brand preferences as well. Mulyanegara et al. (2007) found that some Big 5 personality traits are significantly related to preferences for particular dimensions of brand personality. For example, consumers high in Conscientiousness show a preference towards “trusted” brands while those high in Extraversion preferred “sociable” brands. These findings indicate that perceiving the personality of a Twitter account may be a part of deciding whether or not to follow that account. Thus, I will code for the perceived personality of Twitter accounts and examine their relation to the personality of those interested in following them. This could reveal if people differ in how they are drawn to accounts with different personalities, and if so, whether those associations are suggestive of homophily or some other systematic pattern.

Social media content can evoke intense emotions, suggesting that emotion regulation may play into our interactions with social media. Generally, emotion regulation refers to the ability to control one’s own emotional state, particularly through the use of behaviors and strategies, including situation selection and modification (Gross, 1998). Choosing which Twitter accounts to follow can be seen as a form of selecting one’s situation as these behavioral choices directly modify the content shown in the Twitter feed. It is possible that personality traits of a user may influence the kinds of emotional experiences they want to have or avoid when on twitter, and accordingly, they may be drawn to accounts based on their affective content. To that end, examining the sentiment (positive and negative affect) of profiles, as well as emotion categories (e.g., anger, sadness, joy) could provide insight as to how users with different personality traits may be employing these regulation strategies to modify their social media environment.

Features of Twitter Profiles

A possibility that will be explored in this dissertation is that certain features of Twitter profiles themselves may influence following behaviors. Additionally, the effect of those features on following decisions may vary based on user personality traits. When presented with the opportunity to follow an account on Twitter a user is likely to go through a process of evaluating elements of that account to determine if they would like to follow that account. Clicking into a Twitter account displays that account's profile, which includes customized information about the account (e.g., biography, location, number of friends and followers), a profile and banner photo, and a timeline of that user's Tweets and Retweets. This profile page provides a rich set of features for a user to draw information about the account from, even with just a quick glance. An important component of exploring the effect of Twitter profile features on following decisions is deciding how to measure those features. Better understanding which categories of profile features influence following decisions can provide insight into the process that users take when evaluating potential accounts to include in their curated social media environment. In determining which types of features to measure, it is useful to draw from prior research on expression and perception of attributes.

Twitter Metrics. The Twitter API (application programming interface) enables unique access to public Twitter account data. Through this interface, historical Tweets and user information can be collected. From this data, a number of metrics can be extracted to provide information about Twitter accounts of interest. These metrics can signal information about a Twitter profile that may entice or deter a potential follower. For example, a count of an account's followers can indicate the general popularity of an account while a count of friends can indicate an account's willingness to reciprocate following and potentially engage with their community.

Average word count of Tweets and can be calculated from extracted Tweets. While Tweet text is limited to 280 characters, average Tweet length and frequency can still imply communication style and potential characteristics about the account owner. For example, extraverts have been previously shown to produce text with more words than introverts (Gill & Oberlander, 2019). The ratio of Tweets to Retweets can also provide information about the type of experience a user will have on their feed if they follow this account. Users may be drawn to accounts that produce more original content, while other users may appreciate a high number of Retweets to stay in the loop with a particular community or topic. Relatedly, frequency of Tweets (average number of Tweets per day) can provide additional information about what it would be like to have an account incorporated into a user's feed. Some users may seek out accounts that are very active, while others may feel overwhelmed by frequent Tweets.

Language. Language is inherently social and serves as a tool for communication and expression of oneself (Baldwin & Meyer, 2007; Tomasello, 2010). Particularly, Tweets serve as a very declarative form of communication where a user can explicitly share their beliefs, views, and interests. When a user examines the most recent Tweets of a profile they are considering following, they make judgments about characteristics of the account owner based on the language used. Previous research has shown that people can use online text to accurately perceive user characteristics such as Big 5 personality traits (Qiu et al., 2012) as well as mental health characteristics like depression (Rodriguez et al., 2010). Extracting and analyzing linguistic features from Tweets could uncover categories and topics that drive following decisions.

Approaches to distilling and characterizing the impact of Tweet language on following decisions broadly fall under two categories: dictionary-based approaches and open-vocabulary approaches. The general idea of dictionary approaches is that they pair a person's text with content categories based on common words or phrases. The most frequently used examples of this are the Linguistic Inquiry Word Count software (LIWC; Boyd et al., 2022) and sentiment analysis (Mohammad & Kiritchenko, 2015). LIWC calculates the percentage of words in a user's Tweets that fall into one or more of over 100 categories indicating social and psychological states. This includes categories such as words referencing the self (e.g., I, me) and affiliation words (e.g., community, together) reflecting a person's need to connect with others. In sentiment analysis, words are scored for their relative positivity or negativity based on a pre-trained dictionary (Mohammad & Kiritchenko, 2015). These scores can be average across a user's Tweets to give a sense of the overall affect of a profile. While these methods are great for assessing particular linguistic categories of interest, they are bound to features and words set in the a priori dictionary.

In opposition, an open-vocabulary approach to text analysis examines data-driven topics that are extracted from the text. A common analysis technique is Latent Dirichlet Allocation (LDA; Blei et al., 2003) which is an unsupervised machine learning algorithm that allows sets of observations to be explained by latent groups. Similar to other data reduction methods (e.g., factor analysis), researchers must choose the number of latent topics to fit and there is human interpretation involved in labeling the categories. These methods typically require larger sets of data than dictionary approaches but have the potential to discover important features in the text that may not be captured by dictionary-based approaches.

Perceived Account Characteristics. Judgments about humans' characteristics are made in response to social information in the environment in a process known as social perception (Bruner & Tagiuri, 1954; Hall et al., 2016). This can occur with basic demographic variables such as age, race, and sex (Macrae & Martin, 2007), but also can include other social attributes such as socioeconomic status, political ideology, and even sexual orientation (Bjornsdottir & Rule, 2017; Kraus et al., 2019; Tskhay & Rule, 2014). These perceptions occur in a variety of settings including online social environments (Qiu et al., 2012; Vazire & Gosling, 2004; Waggoner et al., 2009). Profiles are a very public component of Twitter and are set up to be consumed by an audience of perceivers. When a potential follower views a profile, they perceive characteristics of the account. These judgements seek to identify the psychological attributes of people that explain past behavior and help us predict future behavior (Funder, 1995). In the context of Twitter, this can help a user predict the kind of content that will be posted on that account and evaluate if they would like to incorporate that account's Tweets into their feed.

Previous evidence has also established that personality can be perceived with at least moderate accuracy from online profiles. Hall et al. (2014) examined observer accuracy of personality based on Facebook profiles. One hundred participants took a self-report of personality, and 35 zero-acquaintance observers estimated the personality of participants based on PDFs of their Facebook profiles. Observers could estimate profile owners' Agreeableness (r = .32), Extraversion (r = .23), and Conscientiousness (r = .20) with moderate accuracy, but these judgments were less accurate for Neuroticism (r = .16) or Openness (r = .15). These results provide evidence that the content users are displaying on their social media profiles can be utilized to accurately assess characteristics of the account owner. Accuracy in judgements of characteristics may help a potential follower better predict the type of content an account will post in the future. While accuracy in these judgments may provide longer term satisfaction in following, it is important to note that perceived characteristics are likely to be far more important for understanding how someone follows an account than their actual characteristics.

The Current Project

Broadly, my dissertation will explore the extent to which personality is reflected in the environments we create for ourselves online. While prior research and theory support this association, they don't provide a clear path to developing testable hypotheses. Instead, I will take an exploratory and data-driven approach to look for associations that are consistent with past theories or connections that inspire new theories. This exploratory approach will focus future research towards more specific hypotheses, which will allow for controlled experiments, such as manipulating features of accounts, to further narrow in on causal support for associations.

The three studies outlined below will examine the relationship between personality and following Twitter accounts, using behavioral data, as well as self-reports of personality traits. In Study 1, I will use an existing dataset, originally collected for other purposes, to explore possible analysis approaches to better understanding these interconnected relationships and generate preliminary findings. These analyses will aim to not only establish the relationship between personality and interest in Twitter accounts but will also begin to explore features of Twitter profiles that may drive following decisions such as account topics and account metrics. The influence of these features on following accounts will also be considered in relation to personality traits to examine possible interactions.

In Study 2, I will collect new data to further examine the relationship between personality and Twitter account preferences. This study will use an updated set of stimuli that are more relevant to the sampled population. Additionally, when analyzing features of stimuli Twitter profiles, I will also incorporate analysis of Tweet text and human coding of profile characteristics to further explore if specific elements of Twitter profiles drive participant following decisions and if personality moderates that relationship.

Finally, in Study 3 I will explore if Study 2 results can be extended to a naturalistic dataset. Scalable features from Study 2 will be tested on a set of real-world following decisions to examine if results can be generalized beyond controlled laboratory methods. Taken together, the results from these studies will elucidate how people curate their online experiences and help disentangle the complex relationship between social media usage and who we are.

Study Aims

1. Examine if personality traits influence Twitter account following decisions

2a. Explore what features of Twitter profiles drive account following decisions

2b. Analyze if personality moderates this relationship

3. Test generalizability of results beyond hypothetical following decisions

R Analyses

Unless otherwise noted, all analyses were conducted in R (Version 2022.02.2 +485; RStudio Team, 2022) with the R-packages *corrplot* (Version 0.92; Wei et al., 2021), *effects* (Version 4.2-2; Fox et al., 2022), *flextable* (Version 0.7.0; Gohel et al., 2022), *furrr* (Version 0.3.0; Vaughan et al., 2022), *future* (Version 1.25.0; Bengtsson, 2022), *glue* (Version 1.2.6; Hester et al., 2022), *GPArotation* (Version 2022.4.1; Jennrich, 2022), *here* (Version 1.0.1; Müller & Bryan, 2020), *Hmisc* (Version 4.4.2; Harrell Jr, 2020), *imputeTS* (Version 3.2; Moritz et al., 2021), *irr* (Version 0.84.1; Gamer et al., 2019), *janitor* (Version 2.1.0; Firke et al., 2021), *lmerTest* (Version 3.1.3; Kuznetsova et al., 2020), *nFactors* (Version 2.4.1; Liege, 2020), *psych* (Version 2.2.3; Revelle, 2022), *RColorBrewer* (Version 1.1.3; 2022), *reshape2* (Version 1.4.4; Wickham, 2020), *rio* (Version 0.5.29; Becker et al., 2021), *rtweet* (Version 0.7.0; Kearney et al., 2022), *tictoc* (Version 1.0.1; Izrailev, 2021), *tidymodels* (Version 0.2.0; Kuhn et al., 2022), *tidytext* (Version 0.3.3; Queiroz et al., 2022), *tidyverse* (Version 1.3.1; Wickham, 2022), *topicmodels* (Version 0.2.12; Grün et al., 2021), *wordcloud* (Version 2.6; Fellows, 2018).

II. STUDY 1: ANALYSIS APPROACHES TO EXPLORING THE RELATIONSHIP BETWEEN PERSONALITY AND TWITTER ACCOUNT INTEREST

The purpose of Study 1 is to test the feasibility of several analysis approaches as well as generate preliminary findings about how who we are is reflected in online environments. I will use correlations and a cross-validated machine learning approach to quantify the relationship between personality traits and interest in Twitter following Twitter accounts (Aim 1). I will also use multilevel modeling to look at which features of Twitter profiles drive following decisions (Aim 2a) and if personality traits moderate that relationship (Aim 2b). To assess the feasibility of these analyses, I will look at convergence of models and interpretability of results. This will allow for selection and preregistration of the appropriate analyses for Study 2.

The data used in this study was originally collected with the primary aim to assess the relationship between several mental health variables (anxiety, depression, anger, and PTSD) and interest in Twitter accounts. As such, the Twitter accounts used as stimuli in this study were chosen because following them was correlated with mental health variables (Costello et al., 2021). However, in addition to measuring mental health for the original purpose of the study, participant personality traits were also collected. Exploratory analyses of these data revealed some surprising and interesting connections between personality and interest in this particular set of Twitter profiles. Specifically, the traits of Neuroticism and Openness had significant positive associations with interest in accounts previously positively associated with these mental health variables. The findings from this exploration serve as the motivation for applying data-driven analyses to this data set with two goals: (1) to further explore the potential relationship between personality and Twitter account choices in this dataset, and (2) to evaluate the feasibility of planned analyses for Study 2.

Materials and Methods

Data Collection Procedure

Recruitment for Study 1 took place online within the University of Oregon's Department of Psychology subject-pool website. Data collection was open for 8 weeks in Spring 2020, and students could participate at their convenience. Participants who read and agreed with the informed consent were redirected to the Qualtrics questionnaire. They completed the questionnaire as described below and provided demographic information. Participants were then shown screenshots of 100 Twitter profiles and asked how interested they were in following each account. Finally, participants responded to open-ended questions about their Twitter account preferences, provided their Twitter handle, and responded to a series of questions about their Twitter usage. After survey completion, participants' Twitter accounts were scraped to collect their Tweets, a list of their followers, and a list of their friends (accounts they are following on Twitter) using the Rtweet package (M. Kearney, 2019). Students were compensated with course credit for their participation.

Participants. N = 196 participants were recruited through the University of Oregon human-subjects pool, which consisted of students from introductory psychology and linguistics courses. All participants met the prescreening criteria of having a currently active Twitter account and following at least 25 other Twitter accounts. Survey responses were screened for anomalous responding and excessive missingness by a blinded analyst. The analyst determined that all participants' data should be included, and no exclusions were recommended.

After data collection was complete, I discovered that 1 Twitter profile (ESPN) was accidentally duplicated in the set of stimuli that participants viewed. I removed the duplicated profile from the dataset and will conduct the analyses on the 99 remaining accounts that participants viewed. In the cleaned data set, 196 participants each rated 99 different Twitter profiles for a possible 19,404 account interest ratings. Of this, there were only 322 skipped ratings, giving us less than 2% missing data and 19,082 account interest ratings. Demographics for Study 1 participants are shown in Tables 1-3. Participants ranged in age from 18-39 years old with an average age of 19.7 years old.

Table 1

Participant Gender for Study 1

Gender	n
Male	60 (31%)
Female	133 (68%)
Another Identity	3 (1%)

Table 2

Participant Race for Study 1

Race	n
American Indian or Alaska Native	5 (2%)
Asian	17 (9%)
Black or African American	11 (6%)
Native Hawaiian or Pacific Islander	3 (1%)
White	123 (63%)
Other	15 (8%)
More than one race	19 (10%)
Not reported	3 (1%)

Table 3

Participant Ethnicity for Study 1

Ethnicity	n
Hispanic or Latino	39 (20%)
Not Hispanic or Latino	157 (80%)

Sample size. The original aim was to collect data from 300 participants. A sample of 300 and .05 error probability would provide .95 power to detect an effect size of .2, the average effect size in social psychology (Richard et al., 2003). However, this was limited by the number of participants that met the prescreening criteria in the human-subjects pool that term. With a sample of 196 participants there is still .88 power.

Measures

Self-report Measures. Participants completed self-reports of personality traits using a combination of two measures. The Big Five traits (Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Openness) were measured using the Big Five Inventory 2 (BFI 2; Soto & John, 2017), consisting of 60 short statements rated on a scale from 1 (Disagree Strongly) to 5 (Agree Strongly) with a neutral point of 3 (Neither Agree nor Disagree). Eight items from Questionnaire Big Six measure were used to capture the sixth domain, Honesty-Propriety (Thalmayer et al., 2011). These measures showed expected and adequate internal consistency with alpha coefficients for the BFI-2 scales ranging from .78 for Agreeableness to .89 for Neuroticism and an alpha coefficient for Honesty-Propriety at .65. Though not analyzed as a part of this dissertation, participants also completed the following self-report measures: Satisfaction with Life Scale (SWLS; Diener et al., 1985), Scale of Positive and

Negative Experiences (SPANE; Diener et al., 2009), PROMIS Depression, Anger, and Anxiety scales (PROMIS; Pilkonis et al., 2011), Trauma Symptom Questionnaire (TSQ; Brewin et al., 2002).

Twitter Account Stimuli and Ratings. After completing the self-report measures, participants viewed screenshots of 100 Twitter profiles in a randomized order (after data collection was complete, 1 duplicate stimulus profile was found and removed which left 99 stimuli accounts to be analyzed). Each Twitter stimulus screenshot included the banner, profile picture, account name, account bio, account's number of friends and followers, and about 3-6 of the account's most recent Tweets. Participants could scroll to view all components of the screenshot, simulating the experience of viewing a profile page on the Twitter platform. At the end of the screenshot, participants were asked how interested they were in following that account. The study was originally piloted with just a binary response (Follow or Not Follow). However, a floor effect was found where participants were not willing to follow many of the accounts used in the study. To allow for more nuanced responding, the measure was modified to a 4-point scale (Not at All, A Little, Moderately, Very). Participants could also choose to skip the profile. For the purposes of this study, this measure will be referred to as an *account interest rating*.

The original purpose of creating the set of Twitter stimuli used in Study 1 was to test if current mental health variables could prospectively predict interest in following those accounts in a new sample of participants. However, in addition to measuring mental health, participant personality traits were also collected. This has presented the opportunity to apply additional data-driven analyses to this data set to explore possible relationships between personality variables and Twitter account choices. The accounts used as stimuli were taken from a previous naturalistic study (Costello et al., 2021), where the authors scraped the actual accounts that a group of active Twitter users followed on their real accounts to see whether their mental health status correlated with following specific Twitter accounts. The stimuli that were selected for the present study are the 50 accounts that were most positively correlated and the 50 accounts that were most negatively correlated with a general pathology metric in the Costello et al. dataset (the average of anger, depression, anxiety, and PTSD). It was also required that these accounts were followed by at least 3 participants in the Costello et al. study, so they tended to be accounts with a lot of followers such as celebrities or companies.

Qualitative Responses. After viewing and rating interest in all 100 Twitter profile stimuli, participants were asked to reflect on the Twitter accounts they were shown and consider why they were interested or not interested in following those accounts. I will read open-ended text responses and extract broad themes to inform the data collection procedures and stimuli for Study 2.

Features of Twitter Profile Stimuli

In order to analyze what features of the stimuli Twitter accounts drove participants' interest, I extracted features of each of the profiles. As Study 1 is being used to test out the feasibility of analyses, I collected only a select few easily accessible features, with the intention of expanding the set of features in Study 2.

Account Metadata. Several features providing information about the Twitter account itself were extracted from the Twitter API (application programming interface) using the RTweet package (M. Kearney, 2019). *Count of Friends* refers to the number of accounts that a given stimuli Twitter profile follows. *Count of Followers* refers to the number of accounts that follow a given stimuli Twitter account.

Perceived Account Characteristic. Some profile features of interest had potential for subjectivity based on the perception of the viewer. *Human Account Owner* indicates that the Twitter account represents a single person rather than a brand, organization, or group of people. To determine this metric, I reviewed each stimuli Twitter account and coded them as either a human or non-human account.

Principal Component Analysis Dimensions. To uncover latent categories of Twitter accounts within the set of Study 1 stimuli, I used the dimensionality-reducing technique of principal components analysis (PCA) on account interest ratings, treating each of the 99 accounts as a variable. The scree plot testing a number of components indicated potential solutions between 4 and 6 components. I examined solutions with 4, 5, and 6 components with varimax rotation to enhance interpretability and found the 4-component solution to be the most interpretable (see Appendix A). This solution was additionally tested with an oblimin rotation, but I found that the solution was not notably altered by this rotation. The loadings from these PCA components will be used as features to represent content categories present in this set of stimuli Twitter accounts.

I reviewed the stimuli Twitter accounts with the highest loadings for each of the four components to determine the associated latent categories of accounts. Component 1 was determined to be related to *Sports.* Accounts that loaded highly on this topic tended to be sports news accounts and sports commentators, such as Sports Center and Jim Rome. There were also several professional athletes and organizations represented in this component including Kris Bryant, Rickie Fowler, and the National Football League (NFL). Football, baseball, and golf were the most represented sports in the component. Component 2 was determined to be related to *Gaming.* Accounts that loaded highly on this topic tended to be Twitch streamers and video game content creators like Northernlion. There were also several companies that sell gaming and streaming equipment such as Logitech and Elgato Gaming. Component 3 was determined to be related to *Actors.* Accounts that loaded highly on this component were primarily movie and television stars with a particular focus on comedic actors such as Melissa McCarthy and Steve Carrell. Finally, component 4 was determined to be related to *Mom Bloggers.* Accounts that loaded highly on this topic tended to be lifestyle blogs that were primarily run by stay-at-home moms such as, Life with Heidi and Two Kids with a Coupon. Within these accounts, topics generally centered around family, recipes, and saving money.

Analytic Procedure

Aim 1. The first aim of this study is to examine if personality traits influence interest in following Twitter accounts. I will use a multi-faceted approach to examine the strength of the relationship between these two variables. First, I will use Pearson product-moment correlations between participant personality traits and Twitter account interest ratings to examine the strength of this relationship for individual accounts. Then, I will calculate a mean of absolute correlations, indicating a typical level of interest in an account. To aid in effect size interpretation and provide context, I will run the same analyses using several demographic variables including gender, age, and socioeconomic status in place of personality traits.

To quantify how much personality is associated with following interest in the aggregate, I will use the machine learning technique of random forests to further examine the strength of association between Twitter account interest and personality. This technique provides the advantage of cross-validation to prevent overfitting in these models. Additionally, random forests models are well-equipped to handle high dimensionality data where there are a large number of predictors. All models in this study will be trained using k-fold cross-validation in order to reduce overfitting. This procedure splits the data into k random subsets called folds, trains the data with k-1 folds, and tests the model's performance on the kth fold. This method is repeated until each fold has been used as the test fold. In this study k is set to 10, which is commonly recommended and the default setting in the Tidymodels package. Though hyperparameters can be tuned in the training set prior to testing the holdout sample to refine the model parameters, evidence thus far has not demonstrated notable improvements in model performance from hyperparameter tuning (Probst & Boulesteix, 2017; Tang et al., 2018). Therefore, default hyperparameters will be used.

Although the random forests approach is often described as a 'predictive' modeling technique, in the present context it will simply be used to quantify the strength of an association between a personality trait on the one hand, and a set of account interest ratings on the other (cf. set correlation; Cohen, 1982). I will use these models not only to examine the association between personality and account interest ratings, but also other demographic variables such as gender, age, and socioeconomic status. These analyses will demonstrate the effect of personality on following decisions relative to the impact of other demographic variables.

Aim 2a and 2b. The second aim of this study is to examine how account features relate to participants' interest in following them. In aim 2a, I will examine what features of Twitter profiles drive interest in following Twitter accounts. To address this question, I will use multilevel modeling to predict account ratings from the selected Twitter profile features. These data require multilevel modeling because each participant is rating the same set of 99 profiles, giving us both effects of participants and effects of profiles.

In aim 2b I will look at whether there are features of Twitter accounts that appeal to people more or less as a function of personality traits. Moderation analysis within the multilevel models will be used to indicate if Twitter account interest ratings are a main effect of features (everyone wants to follow accounts with the same features) or if the effect of a feature depends on personality traits.

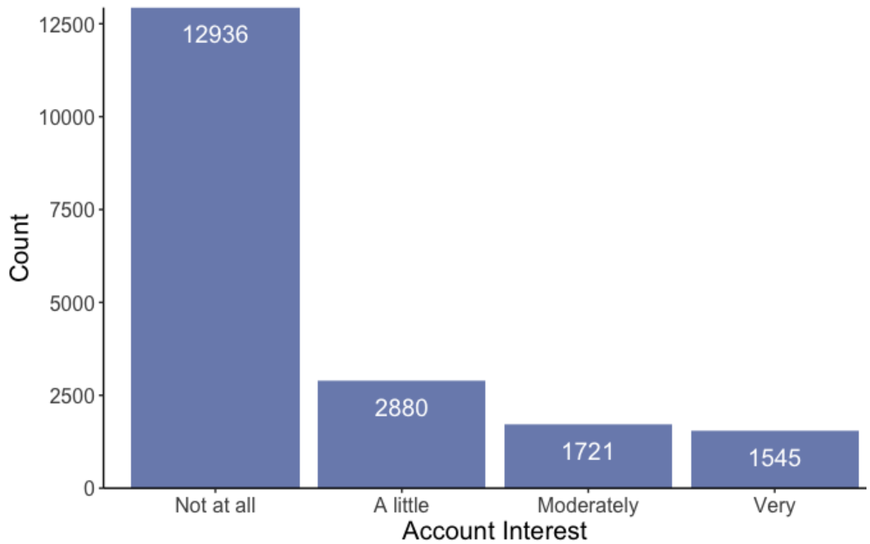
Results

Aim 1: Associations Between Personality Traits and Interest in Following Twitter Accounts

Stimuli Account Ratings. As a preliminary analysis, I examined the distribution of participant interest ratings across all 99 Twitter stimuli profiles. As shown in Figure 1, the distribution is positively skewed, with 68% percent of ratings indicating no interest in following a given account and 32% indicating at least a little interest in following a given account. The mean interest rating across all participants and profiles was 1.57 (*SD* = .95) out of a 4-point scale.

Figure 1

Distribution of Study 1 Account Interest Ratings



Note. There were 19,082 total ratings.

Correlations. To examine the strength of the relationship between participant personality traits and Twitter profile account interest ratings, I calculated Pearson product-moment correlations (Table 4). There were a number of significant relationships between personality traits and account interest ratings at the individual account level. Figures 2-7 visualize the twenty highest correlated accounts with each personality trait and the direction of those correlations, identifying the strongest relationships with individual accounts. These figures can also help us compare the strength of these relationships across traits. For example, we can see the highest correlated accounts with Neuroticism and Extraversion generally show stronger correlations than the highest correlated accounts with Conscientiousness and Agreeableness.

To further summarize this data, I calculated a mean of absolute correlations for each trait across the 99 stimuli profiles (Table 5). This metric provides further information about the typical level of interest in an account by personality trait. Mean absolute correlations showed the strongest relationship with Extraversion (r = .12) and the weakest with Conscientiousness and

Agreeableness (r = .08). The same correlation summary was calculated for the demographic variables of gender (binary male = 1 and female = 2, non-binary was removed from these analyses due a low number of responses), age (numeric, range = 18-39 years), and socioeconomic status (a numeric self-report on a ten-point scale) (Table 6). Gender had the highest mean absolute correlation (r = .17), self-reported SES showed a relationship similar to Conscientiousness and Agreeableness (r = .08), and age had the smallest relationship (r = .05). The mean absolute correlations for personality traits were larger than age and at least as big as SES, but none were quite as large as gender. This indicates that the effect of personality traits on Twitter account interest falls generally in the same range as demographic variables.

One question that arises from these results is whether these associations occur because some personality traits are associated with generally wanting to follow more accounts, as opposed to interest in the specific account content. To investigate this relationship, I standardized account interest ratings within participants (ipsatized) and recalculated correlations with this data (Table 7 & 8). The mean of absolute correlations did not change notably between regular and ipsatized data, demonstrating that even when controlling for a person's individual propensity to follow accounts, there are still notable relationships between personality traits and account interest ratings. To look at this relationship more specifically, I calculated correlations between an individual's mean account interest rating and their personality traits and demographic variables. While many of these correlations were close to zero, Extraversion showed a notable positive correlation (r = .2) indicating that those higher in Extraversion generally showed more interest in following accounts. Notable negative relationships were demonstrated with Neuroticism (r = -.16) and gender (r = -.28) indicating that those high in Neuroticism and Male participants generally showed less interest in following accounts.

Table 4

Correlations Between Study 1 Twitter Account Interest Rating and Personality Traits

Twitter Account	Account Description	Neuroticism	Extraversion	Openness	Honesty	Conscientiousness	Agreeableness
SportsCenter	Sports News Show	**-.43*** **	.25***	-.24***	-.15*	.02	-.04
KrisBryant_23	Baseball Player	**-.39*** **	.31***	-.25***	-.12	.07	-.01
NFL	Football League	**-.35*** **	.25***	-.29***	-.12	.09	-.08
McShay13	Sports Analyst	**-.31*** **	.22**	-.12	-.09	.02	-.06
obj	Football Player	**-.29*** **	.29***	-.19**	-.08	.07	.06
mortreport	Sports Reporter	**-.28*** **	.27***	-.12	-.01	.12	.00
Rotoworld_FB	Fantasy Football	**-.27*** **	.18*	-.10	-.09	.02	-.11
MelKiperESPN_2	Sports Analyst	**-.27*** **	.14	-.08	-.18*	-.05	-.10
Arrieta34	Baseball Player	**-.27*** **	.25***	-.21**	.02	.10	.10
AdamSchefter	Sportswriter	**-.26*** **	.17*	-.07	-.19**	-.06	-.08
PatMcAfeeShow	Sports Analyst	**-.25*** **	.14*	-.17*	-.17*	-.04	-.09
StuartScott	Sportscaster	**-.25*** **	.12	-.05	-.07	.04	.00
YahooSportsNBA	Sports News	**-.24*** **	.23***	-.17*	-.07	.03	-.02
Ken_Rosenthal	Sportswriter	**-.23** **	.20**	-.12	-.04	.03	-.02
InternetHippo	Humor, Twitch Streamer	**.23** **	-.13	.17*	-.12	-.20**	.05
jimrome	Sports Radio Host	**-.20** **	.10	-.13	-.15*	-.14	-.13
PFTCommenter	Fictional Sports Host	**-.17* **	.09	-.04	-.17*	-.13	-.12
Totalbiscuit	Gaming Commentator	**-.13**	.04	.05	-.01	.01	.07
Chrisspymakeup	Beauty Influencer	**.12**	.02	.12	-.01	-.06	-.02

Table 4 Continued

Twitter Account	Account Description	Neuroticism	Extraversion	Openness	Honesty	Conscientiousness	Agreeableness
PBnWhine	Lifestyle Blog	**-.09**	.07	-.03	-.02	.01	-.01
TeamUSA	Olympic Organization	-.29***	**.34*****	-.11	-.01	.23**	.06
JLo	Singer, Actress	-.24***	**.33*****	-.13	.03	.15*	.09
NFL_Memes	Humor, Memes	-.17*	**.31*****	-.16*	-.10	.03	.02
ElizabethBanks	Actor	-.10	**.25*****	.05	-.02	.00	.05
RobLowe	Actor	-.05	**.25*****	.16*	.03	.07	.01
DangeRussWilson	Football Player	-.21**	**.24*****	-.11	-.10	.01	.03
RickieFowler	Golfer	-.21**	**.24****	-.10	-.09	.05	-.05
melissamccarthy	Actress, Comedian	-.06	**.23****	-.05	.07	.16*	.15*
AmznMovieRevws	Movie Reviews, Humor	.17*	**-.22****	.11	.01	-.11	.02
SweepsAdvantage	Sweepstakes Directory	-.10	**.17***	-.03	.03	.01	.07
CouponsFreebie	Coupons, Giveaways	.04	**.15***	-.04	-.01	.01	.04
SteveCarell	Actor, Comedian	-.06	**.15***	.09	.02	.00	.03
BestBuy_Deals	Technology Deals	-.07	**.14**	.00	.08	.07	.05
RiffTrax	Movie Commentary	-.07	**.12**	.10	-.11	-.03	-.03
batemanjason	Actor	.00	**.12**	.12	-.10	-.11	.03
NIVEAUSA	Skincare Brand	-.05	**.12**	.02	.01	.06	.04
RBReich	Robert Reich, Economist	.09	-.07	**.25*****	.02	-.10	-.02
BBCAMERICA	Television Network	-.02	.04	**.24*****	.05	.00	.04
GilianA	Gillian Anderson, Actress	.10	.13	**.24*****	.06	.05	.07

Table 4 Continued

Twitter Account	Account Description	Neuroticism	Extraversion	Openness	Honesty	Conscientiousness	Agreeableness
ElderScrolls	Video Game	.16*	-.06	**.21****	-.07	-.16*	-.13
bubbawatson	Golfer	-.17*	.15*	**-.21****	-.13	.01	-.11
ABFalecbaldwin	Non-profit Organization	-.13	.08	**.19****	-.01	.05	.10
snopes	Fact Checking Website	.08	-.11	**.19****	-.06	-.17*	-.16*
ThatKevinSmith	Filmmaker	.00	.04	**.19****	-.15*	-.06	-.07
TomiLahren	Political Commentator	-.13	.15*	**-.18***	-.07	.08	-.10
BobsBurgersFOX	Television Show, Humor	.11	.10	**.18***	.03	-.08	.11
CobieSmulders	Actor	-.04	.08	**.18***	.02	.16*	.06
GroovyBruce	Bruce Campbell, Actor	.05	-.02	**.17***	-.06	-.11	-.06
PrettyLights	Music Producer	-.12	.02	**.17***	-.16*	.05	-.02
verge	Technology News	-.10	-.01	**.15***	-.10	-.05	-.01
arnettwill	Actor, Comedian	.11	-.02	**.14***	-.04	-.07	.01
omgthatspunny	Memes, Humor	-.05	.00	**-.13**	.03	.09	.07
jonbonjovious	Mom, Lifestyle Blogger	.05	.10	**-.12**	.09	.05	.01
TheWookieeRoars	Nonprofit Organization	-.02	.03	**.12**	-.01	-.08	.08
TomHall	Twitter Consultant	-.03	.03	**-.10**	-.06	.00	-.04
KellysLuckyYou	Lifestyle Blog	-.05	.02	**-.10**	-.08	-.01	-.05
drewmagary	Journalist	-.01	.02	**.09**	-.06	-.08	-.05
SeeMomClick	Mom, Lifestyle Blogger	.00	.02	**-.04**	-.02	.03	-.02
BarstoolBigCat	Dan Katz, Sports Host	-.24***	.16*	-.13	**-.28*****	-.05	-.02

Table 4 Continued

Twitter Account	Account Description	Neuroticism	Extraversion	Openness	Honesty	Conscientiousness	Agreeableness
Lilpeep	Rapper	.17*	-.09	.17*	**-.23****	-.19**	-.12
solecollector	Sneaker Magazine	-.07	.07	.03	**-.21****	-.03	-.15*
stoolpresidente	Barstool Sports Founder	-.16*	.09	.02	**-.19****	-.02	-.08
FamilyGuyOnFOX	Television Show, Humor	-.04	.13	-.07	**-.18***	-.10	-.10
notch	Video Game Designer	.04	-.13	.08	**-.18***	-.12	-.16*
newcastle	Beer Company	-.17*	.17*	.03	**-.18***	-.01	-.10
joelmchale	Actor, Comedian	.10	.03	.03	**-.16***	-.15*	-.06
normmacdonald	Comedian	.00	.03	.07	**-.15***	-.05	.00
newbelgium	Beer Company	-.07	.09	.07	**-.14***	-.07	-.05
McllroyRory	Golfer	-.09	.11	-.12	**-.14**	-.03	-.12
ders808	Actor, Comedian	.03	.08	.10	**-.14**	-.01	.02
Lovesmytwoboys	Mom, Lifestyle Blogger	-.04	.10	-.08	**.11**	.07	.10
NorthernlionLP	Video Game Streamer	.04	-.02	.12	-.13	**-.23****	-.15*
EdwardNorton	Actor	.15*	-.17*	.18*	-.03	**-.22****	.00
elgatogaming	Audiovisual Technology	.00	-.12	-.01	-.16*	**-.22****	-.16*
ZOWIEbyBenQUSA	Esports Equipment Brand	-.01	-.02	.01	-.08	**-.22****	-.14
thesulk	Writer, Voice Actor	.19**	-.07	.10	-.11	**-.20****	-.05
LogitechG	Gaming Equipment	.04	-.05	.01	-.07	**-.19****	-.12
feedme	Electronic Musician	.10	-.08	.16*	-.07	**-.17***	-.04
SMITEGame	Online Multiplayer Game	-.06	.02	.01	.01	**-.15***	-.08

Table 4 Continued

Twitter Account	Account Description	Neuroticism	Extraversion	Openness	Honesty	Conscientiousness	Agreeableness
michaelianblack	Comedian, Actor	.10	.09	-.08	-.10	**-.15***	-.12
MrCraigRobinson	Actor, Comedian	.06	-.04	.10	-.09	**-.12**	.02
MissingLynxx	Lifestyle Blogger	-.09	.05	-.04	.04	**.12**	.00
teksyndicate	Technology News	-.02	-.11	.07	-.07	**-.12**	-.09
NaNoWriMo	Nonprofit Organization	.03	-.08	.08	.03	**-.11**	.07
ToysRUs	Toy Company	-.09	.03	.05	-.10	**-.10**	-.08
FSOC2011	Money Saving Blog	.04	.00	-.07	-.08	**-.09**	-.06
twokidsandacoupon	Mom, Lifestyle Blogger	-.02	.02	-.05	-.03	**-.06**	-.02
BillCorbett	Writer, Voice Actor	-.02	.04	-.05	-.12	.03	**-.23****
IronsidePC	Gaming Computers	.01	-.03	.03	-.17*	-.20**	**-.22****
ConservamomE	Lifestyle Blogger, Mom	-.14*	.08	-.07	.13	.11	**.21****
lustrelux	Beauty Influencer	.04	.04	.04	.06	.04	**.19****
Monstercat	Electronic Music Label	-.01	-.06	.06	-.10	-.10	**-.16***
Sleepopolis	Mattress Reviews	.07	.05	-.02	-.11	-.10	**-.15***
lifewithheidig	Money Saving Blogger	.02	.05	-.05	.08	.06	**.14***
SpiderManMovie	Movie	-.10	.13	-.04	.09	.04	**.14***
StateDept	Government Organization	.00	.07	.05	-.07	-.04	**-.13**
SanitySuburbia	Mom, Lifestyle Blog	-.05	.03	-.02	.02	.09	**.10**
TheWalkingDead	Television Show	.03	.00	.06	-.07	-.05	**-.10**

Note. Bolded values indicate the highest absolute correlation for each stimuli Twitter account. * correlation is significant at the 0.05 level, ** correlation is significant at the 0.01 level, and *** correlation is significant at the 0.001 level.

Figure 2

Twenty Highest Correlated Twitter Accounts with Neuroticism in Study 1

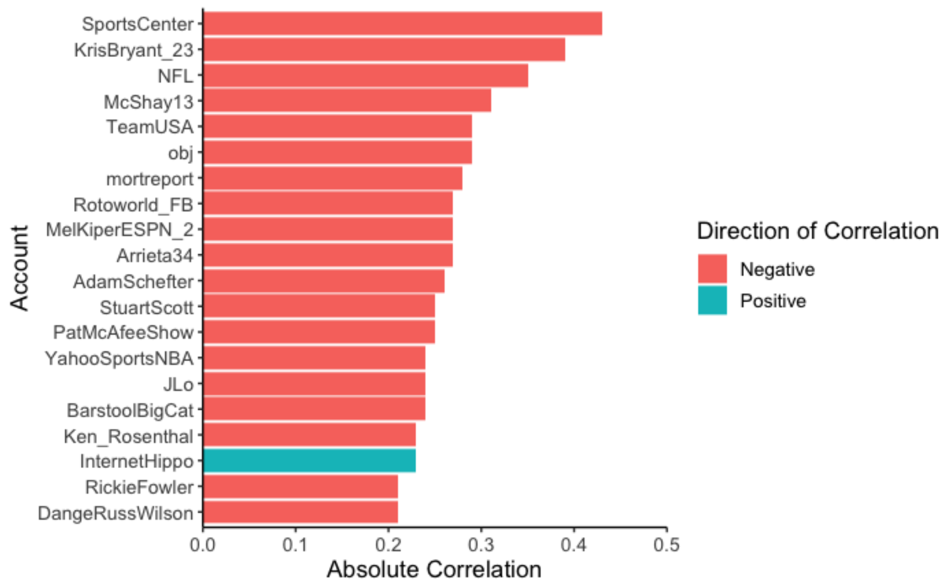


Figure 3

Twenty Highest Correlated Twitter Accounts with Extraversion in Study 1

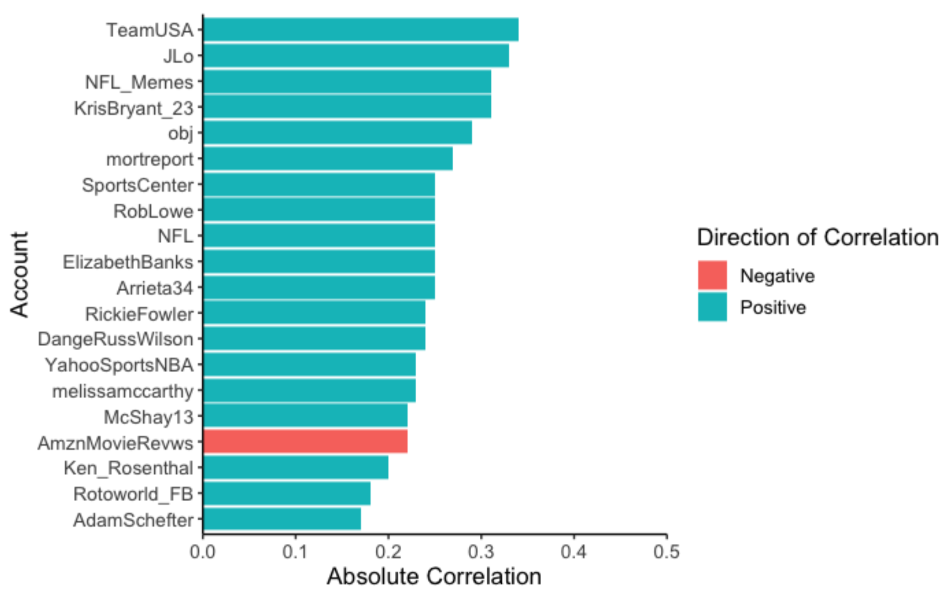


Figure 4

Twenty Highest Correlated Twitter Accounts with Openness in Study 1

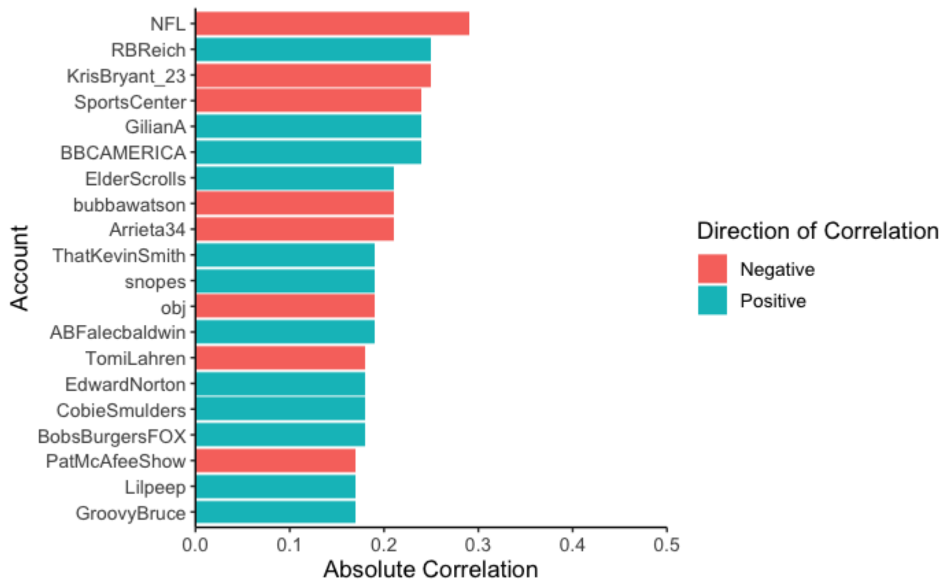


Figure 5

Twenty Highest Correlated Twitter Accounts with Honesty-Propriety in Study 1

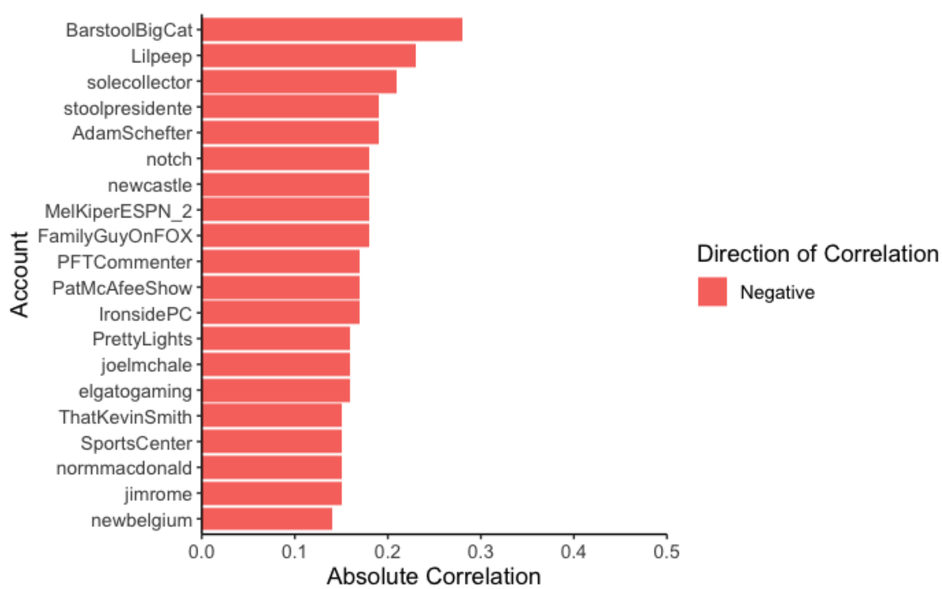


Figure 6

Twenty Highest Correlated Twitter Accounts with Conscientiousness in Study 1

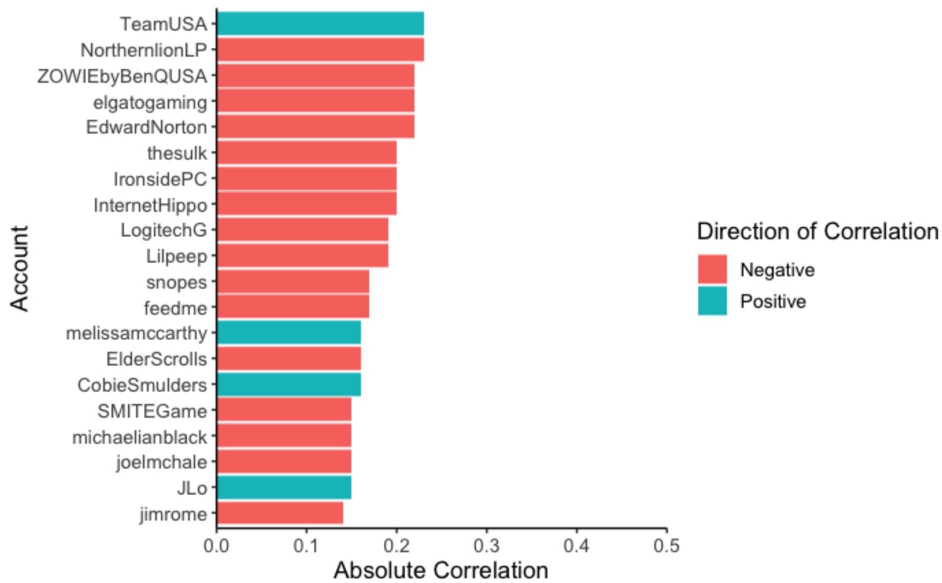


Figure 7

Twenty Highest Correlated Twitter Accounts with Agreeableness in Study 1

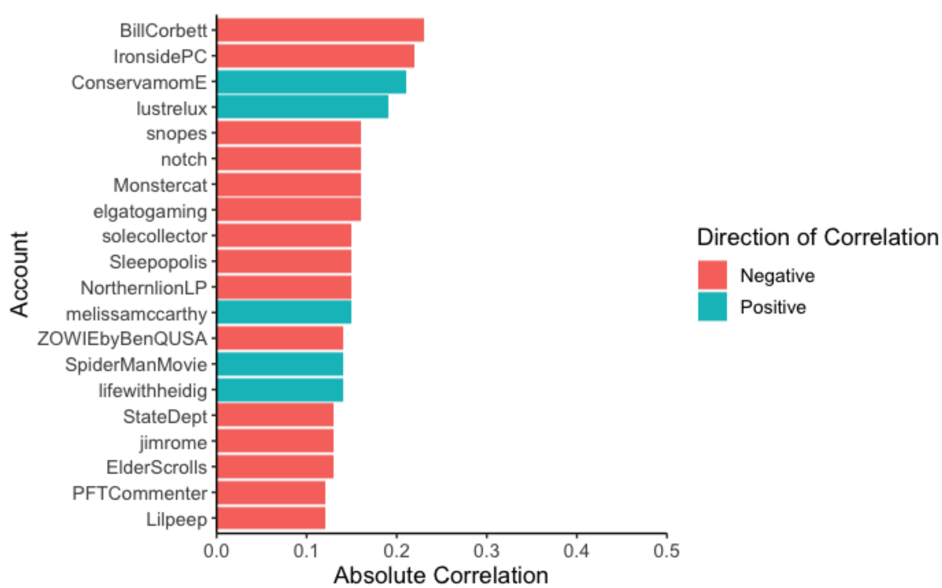


Table 5

Mean of Absolute Correlations Between Twitter Account Interest Ratings and Personality Traits in Study 1

	Mean r
Neuroticism	.12
Extraversion	.11
Openness	.10
Honesty	.09
Conscientiousness	.08
Agreeableness	.08

Table 6

Mean of Absolute Correlations Between Twitter Account Interest Ratings and Demographic Variables in Study 1

	Mean r
Gender	.18
Self-report SES	.08
Age	.05

Table 7

Mean of Absolute Correlations Between Ipsatized Twitter Account Interest Ratings and Personality in Study 1

	Mean r	Correlation with Individual Mean Interest Rating
Neuroticism	.14	-.16
Extraversion	.13	.20
Openness	.12	.02
Honesty	.10	-.14
Conscientiousness	.11	-.06
Agreeableness	.10	-.03

Table 8

Mean of Absolute Correlations Between Ipsatized Twitter Account Interest Ratings and Demographic Variables in Study 1

	Mean r	Correlation with Individual Mean Interest Rating
Gender	.17	-.28
Self-report SES	.09	.01
Age	.05	-.01

Random Forests. To explore the strength of the relationship between personality traits and Twitter stimuli account interest ratings in the aggregate, I used the machine learning technique of random forests to predict personality traits from account interest ratings. The Study 1 sample (N=196) was split into a training and testing (holdout) set using the Tidymodels package in R (Kuhn & Wickham, 2020). The training and holdout samples consisted of 75% ($n_{training}$ = 147) and 25% ($n_{holdout}$ = 49) of the data respectively. Nine models were run, six predicting each of the personality traits individually, one predicting gender (binary male = 1 and female = 2), one predicting age (numeric), and one predicting the self-report of socioeconomic status (on a ten-point scale) (Table 9 & Table 10). There was moderate predictive accuracy across all six personality traits with Neuroticism showing the highest out-of-sample accuracy and Agreeableness showing the lowest. Similarly, in predicting demographic variables from Twitter account interest ratings, gender was predicted with by far the greatest accuracy, age was predicted with moderate accuracy, and socioeconomic status was predicted with the least accuracy. These models demonstrate that account interest ratings can predict personality traits with higher accuracy than age and SES but lower accuracy then gender.

Table 9

Random Forests Performance for Twitter Account Interest Ratings Predicting Personality Traits in Study 1

Trait	Train rmse	Train rsq	Train r	Test rmse	Test rsq	Test r	Lower 95% CI r	Upper 95% CI r
Neuroticism	0.68	.16	.40	0.74	.12	.35	.08	.57
Conscientiousness	0.66	.07	.26	0.64	.12	.35	.08	.57
Extraversion	0.60	.20	.45	0.64	.11	.33	.05	.56
Openness	0.54	.16	.40	0.59	.10	.32	.04	.55
Honesty	0.52	.10	.32	0.59	.06	.24	-.04	.49
Agreeableness	0.55	.07	.26	0.53	.03	.17	-.12	.43

Table 10

Random Forests Performance for Twitter Account Interest Rating Predicting Demographic Variables in Study 1

Trait	Train rmse	Train rsq	Train r	Test rmse	Test rsq	Test r	Lower 95% CI r	Upper 95% CI r
Gender	0.41	.43	.66	0.29	.57	.75	.59	.85
Age	1.70	.06	.24	2.97	.03	.17	-.12	.43
Self-report SES	1.41	.17	.41	1.44	.00	.04	-.24	.32

Aim 2a: Account Ratings Predicted from Profile Features

Multilevel Modeling. Do certain account features make users more or less interested in following those accounts? Multilevel modeling was used to examine the effect of individual Twitter profile features on account interest ratings. The model specification is as follows:

i = subject
j = account

$\text{Account Interest Rating}_{i,j} = b_{0i} + b_{1j} + b_{2i} * \text{Profile Feature}_j + e_{ij}$
$b_{0i} = \gamma_{00} + U_{0i}$
$b_{1j} = \gamma_{10} + U_{1j}$
$b_{2i} = \gamma_{20} + U_{2j}$

I ran a total of 7 models, one for each Twitter profile feature. The models were initially run with a random intercept for subject (U_{0i}), a random intercept for profile (U_{1j}), and a random slope for the effect of feature (U_{2j}). In cases where the model did not converge, I trimmed the random slope for subject (U_{2j}) and re-ran the model. For all models, the account interest rating metric was converted to a proportion of maximum possible score (POMP; Cohen et al., 1999). This transformation converts scores to a percent of the highest score available on the scale, with a theoretical range from 0 to 100, which allows for easier interpretation and communication of model results. To compare the effects of features with different measurement scales, Twitter profile features were z-scored. Features with meaningful units of measurement were converted back to their original scale for individual interpretation. Table 11 summarizes the features that were used as predictors.

Table 11

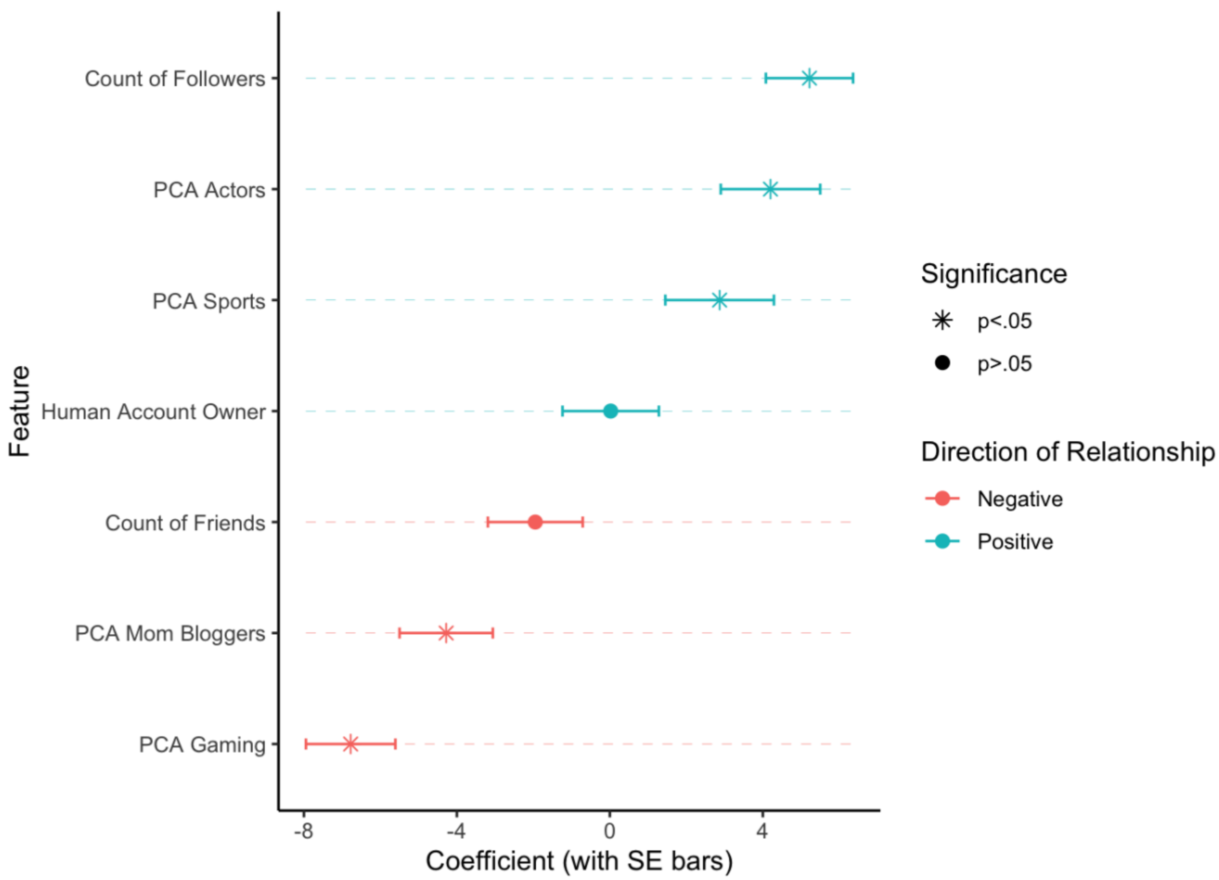
Summary of Twitter Stimuli Profile Features in Study 1

Feature Name	Feature Type	Description	Example
Count of Friends	Account Metadata	Number of accounts that a stimuli Twitter profile follows	JLo has 43,678,741 followers
Count of Followers	Account Metadata	Number of accounts that follow a given stimuli Twitter account	JLo follows 1,693 accounts
Human Account Owner	Perceived Account Characteristic	A binary variable indicating that the Twitter account represents a single person rather than a brand or group of people	JLo is a human, ToysRUs is not a human
PCA Sports	PCA Dimension	Component loadings for the latent category related to sports news and athletes	Accounts that load highly include on this dimension include Sports Center and Rory McIlroy
PCA Gaming	PCA Dimension	Component loadings for the latent category related to video games and associated equipment	Accounts that load highly include on this dimension include Logitech and Elder Scrolls
PCA Actors	PCA Dimension	Component loadings for the latent category related to tv and movie stars	Accounts that load highly include on this dimension include Will Arnett and Joel McHale
PCA Mom Bloggers	PCA Dimension	Component loadings for the latent category related to lifestyle bloggers	Accounts that load highly include on this dimension include PBnWhine and CouponsFreebie

The effect of Twitter profile features on account following interest (and standard error) is displayed in Figure 8. Count of followers and all four of the PCA topics demonstrated significant influence on account interest ratings. Count of followers was the feature that had the largest positive influence on account interest ratings, with a z-scored coefficient of 5.22. In unstandardized units, this means that for every 1 million additional followers a Twitter account has, interest in following increased by 8.6 POMP units. The Gaming PCA topic had the largest negative influence on account ratings. The more strongly an account loaded on the Gaming PCA topic the less participants were interested in following. For every standard deviation increase in gaming content relevance, interest in following decreased by 6.78 POMP units.

Figure 8

Effect of Z-scored Twitter Profile Features on Study 1 Account Interest Ratings



Aim 2b: Personality as a Moderator for the Relationship Between Twitter Profile Interest Ratings and Twitter Profile Features

Moderation Analysis. Do account features have different effects for different participants, as a function of the participant's personality? To examine if participant personality traits moderate the effect of Twitter profile features on Twitter account interest ratings, moderation analysis was incorporated into the multilevel models. The model specification is as follows:

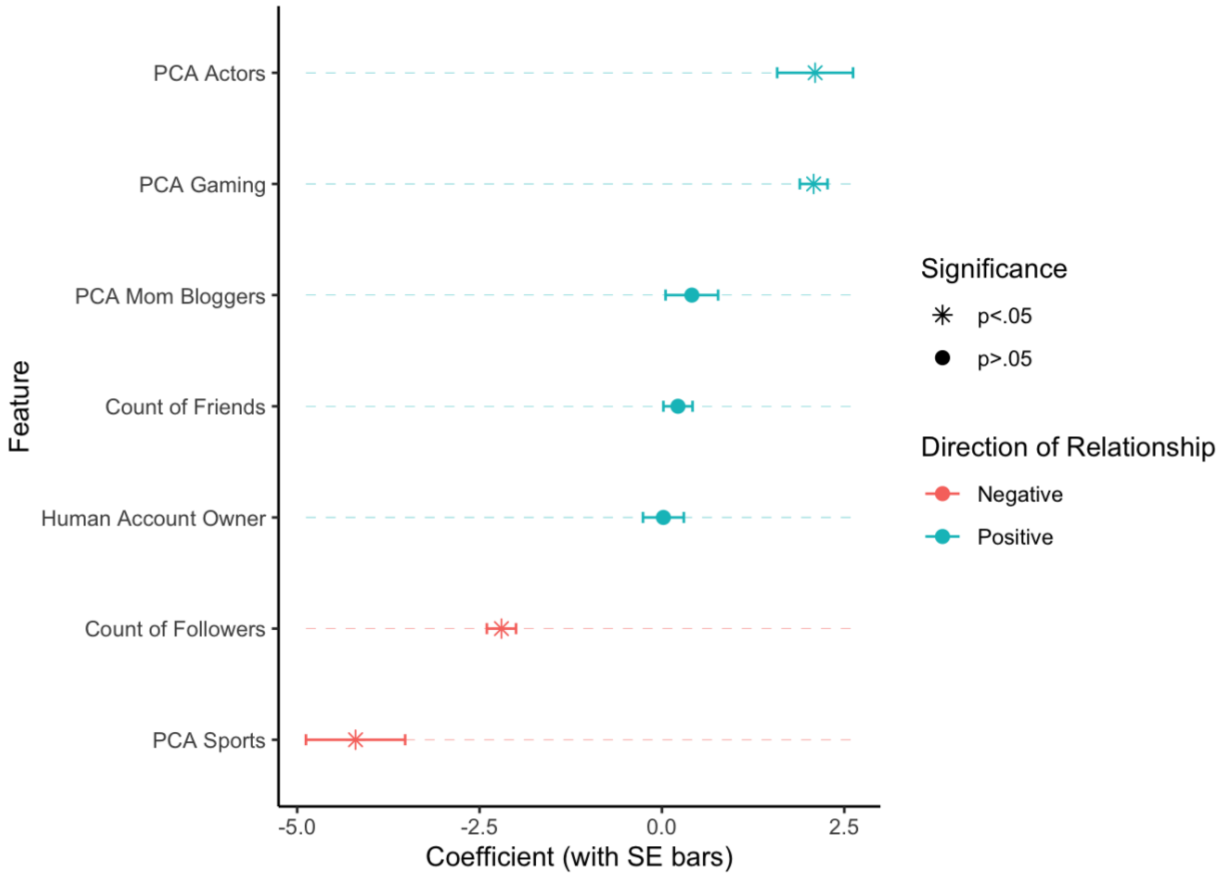
i = subject
j = account

$$\text{Account Interest Rating}_{i,j} = b_{0i} + b_{1j} + b_{2i} * \text{Profile Feature}_j + e_{ij}$$
$$b_{0i} = \gamma_{00} + \gamma_{01} * \text{Trait}_i + U_{0i}$$
$$b_{1j} = \gamma_{10} + U_{1j}$$
$$b_{2i} = \gamma_{20} + \gamma_{21} * \text{Trait}_i + U_{2j}$$

I ran a total of 42 models, one for each combination of 6 personality traits and 7 Twitter profile features. I used the same trimming procedure for models that did not converge as the original multilevel models. The account interest rating metric was converted to POMP scores. Similarly, to compare model results across features with different measurement scales, Twitter profile features and personality traits were z-scored. The full set of 42 interaction coefficients are presented in figures 9, 12, 15, 18, 21, and 23. Positive coefficients indicate that the effect of the feature on interest is relatively more positive for people who score high on the trait (and relatively more negative for people who score low). Negative coefficients indicate the reverse. To aid in interpretation and illustrate what these interaction effects look like when added to the main effects presented in the previous section, I have plotted and elaborated on the interpretation for the most positive and most negative interaction effect for each trait.

Figure 9

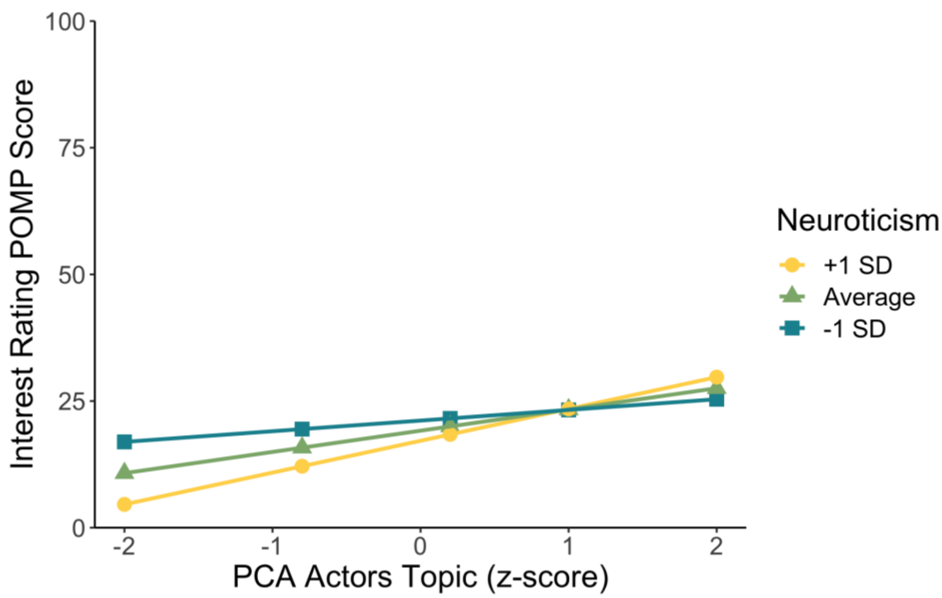
Moderating Effect of Z-scored Neuroticism on the Relationship Between Z-scored Twitter Profile Features and Study 1 Account Interest Scores



Participant Neuroticism significantly moderated the effect of count of followers and the PCA topics of Actors, Gaming, and Sports (Figure 9). The effect of the Actors PCA topic feature was most positively moderated by Neuroticism. To aid in interpretation, I calculated simple effects for people at different levels of neuroticism (Figure 10). The main effect of Actors PCA was positive, indicating that for someone with an average level of Neuroticism, every standard deviation increase in actor content relevance was associated by 4.2 POMP units greater interest in following. For someone 1 standard deviation above the mean in Neuroticism, the simple effect of actor related content was 6.3. For someone 1 standard deviation below the mean in Neuroticism, the simple effect of actor related content was 2.1. In other words, higher levels of Neuroticism were associated with greater sensitivity to actor-related content.

Figure 10

Interaction Plot for Neuroticism and the Effect of PCA Actor Topic



The effect of the Sports PCA topic feature was most negatively moderated by Neuroticism (Figure 11). The main effect of Sports PCA was positive, indicating that at an average level of Neuroticism, every standard deviation increase in sports content relevance resulted in 2.9 POMP units greater interest in following. For someone 1 standard deviation above the mean in Neuroticism, the simple effect of sports related content was -1.3. For someone 1 standard deviation below the mean in Neuroticism, the simple effect of sports related content was 7.1. In other words, lower levels of Neuroticism were associated with greater sensitivity to sports-related content.

Figure 11

Interaction Plot for Neuroticism and the Effect of PCA Sports Topic

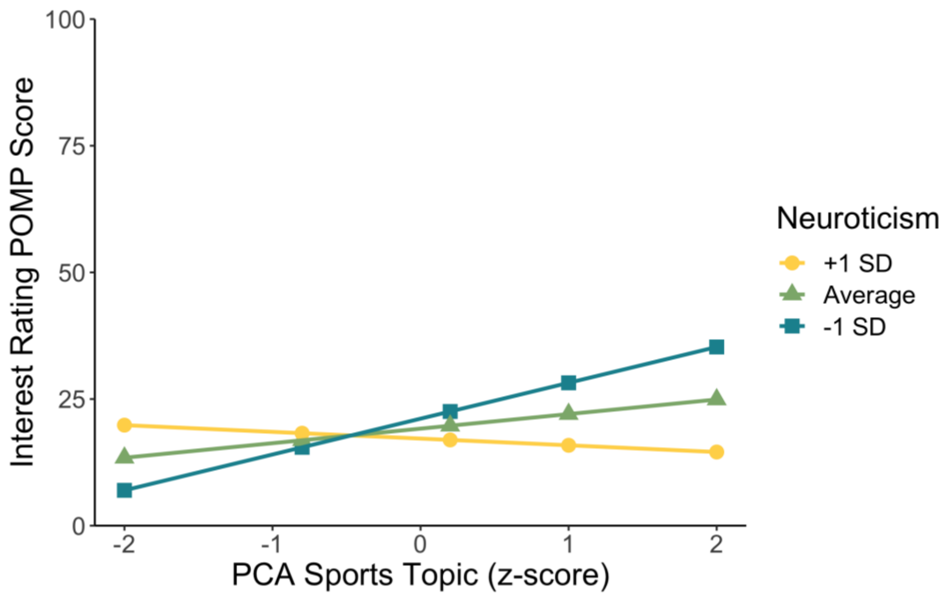
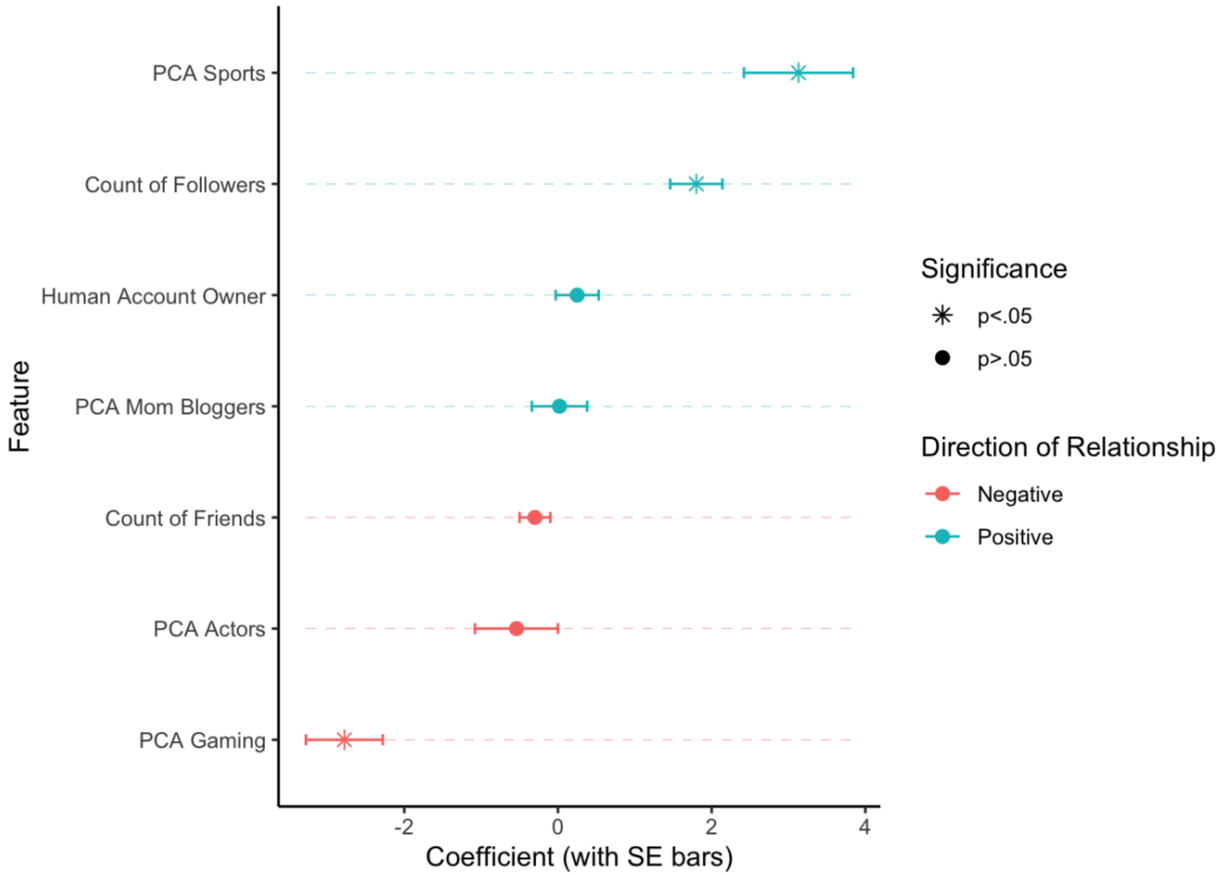


Figure 12

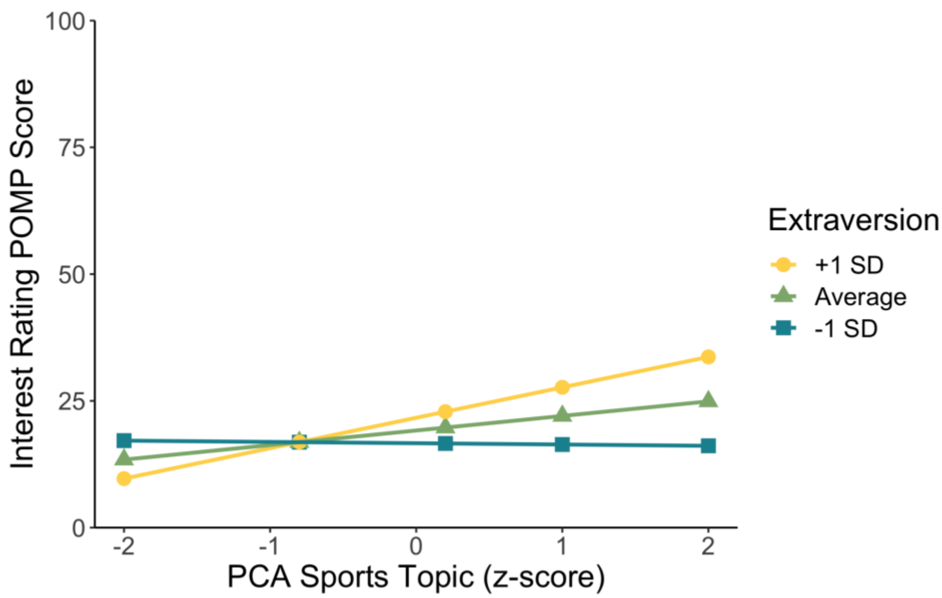
Moderating Effect of Z-scored Extraversion on the Relationship Between Z-scored Twitter Profile Features and Study 1 Account Interest Scores



Participant Extraversion significantly moderated the effect of count of followers and the PCA topics of Sports and Gaming (Figure 12). The effect of the Sports PCA topic feature was most positively moderated by Extraversion (Figure 13). The main effect of Sports PCA was positive, indicating that at an average level of Extraversion, every standard deviation increase in sports content relevance resulted in 2.9 POMP units greater interest in following. For someone 1 standard deviation above the mean in Extraversion, the simple effect of sports related content was 6. For someone 1 standard deviation below the mean in Extraversion, the simple effect of sports related content was -0.2. In other words, higher levels of Extraversion were associated with greater sensitivity to sports-related content.

Figure 13

Interaction Plot for Extraversion and the Effect of PCA Sports Topic



The effect of the Gaming PCA topic feature was most negatively moderated by Extraversion (Figure 14). The main effect of Gaming PCA was negative, indicating that at an average level of Extraversion, every standard deviation increase in sports content relevance resulted in 6.8 POMP units less interest in following. For someone 1 standard deviation above the mean in Extraversion, the simple effect of gaming related content was -9.6. For someone 1 standard deviation below the mean in Extraversion the main effect of gaming related content was -4. In other words, higher levels of Extraversion were associated with greater sensitivity to gaming-related content.

Figure 14

Interaction Plot for Extraversion and the Effect of PCA Gaming Topic

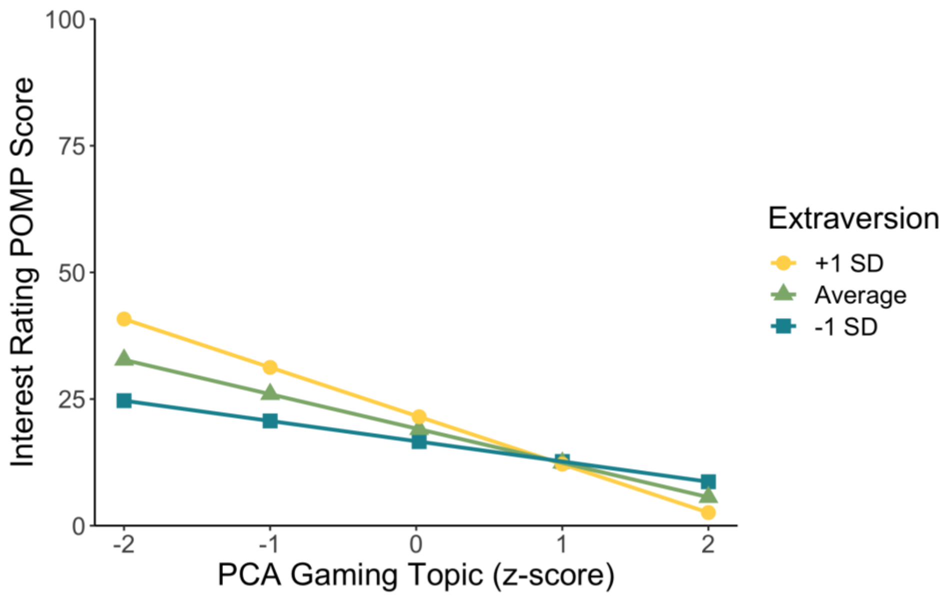
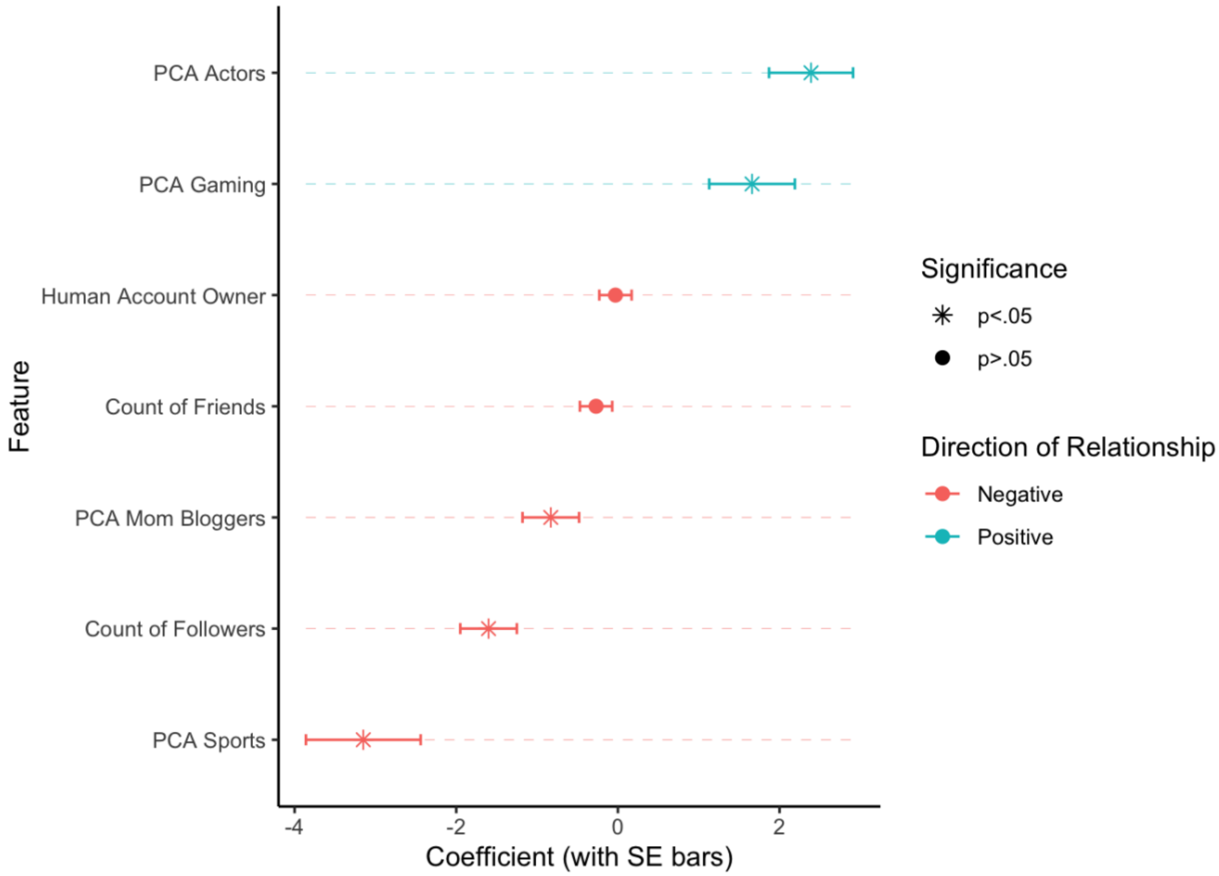


Figure 15

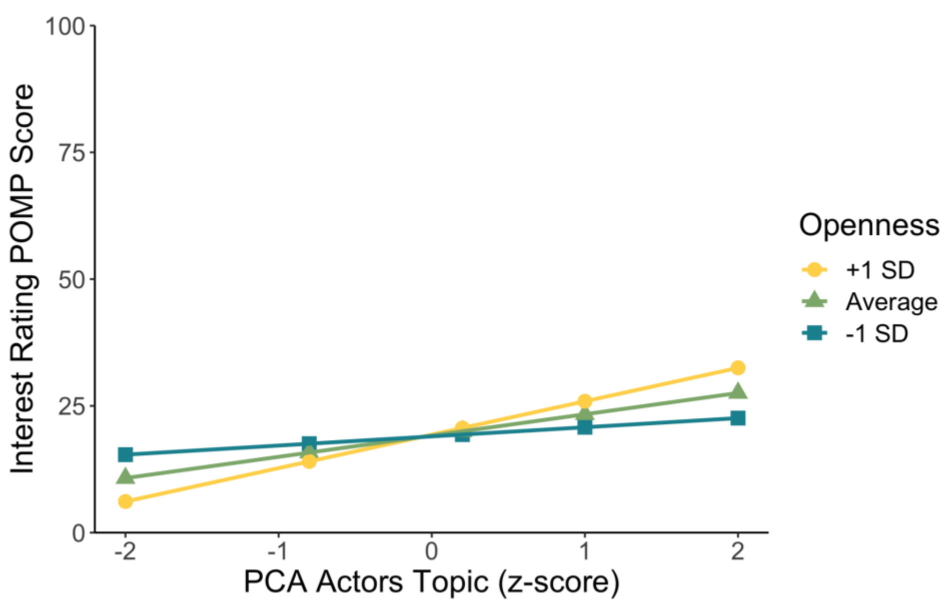
Moderating Effect of Z-scored Openness on the Relationship Between Z-scored Twitter Profile Features and Study 1 Account Interest Scores



Participant Openness significantly moderated the effect of count of followers and all four of the PCA topics on account interest ratings (Figure 15). The effect of the Actors PCA topic feature was most positively moderated by Openness (Figure 16). The main effect of Actors PCA was positive, indicating that at an average level of Openness, every standard deviation increase in actor related content relevance resulted in 4.2 POMP units greater interest in following. For someone 1 standard deviation above the mean in Openness, the simple effect of actor related content was 6.6. For someone 1 standard deviation below the mean in Openness, the simple effect of actor related content was 1.8. In other words, higher levels of Openness were associated with greater sensitivity to actor-related content.

Figure 16

Interaction Plot for Openness and the Effect of PCA Actors Topic



The effect of the Sports PCA topic feature was most negatively moderated by Openness (Figure 17). The main effect of Sports PCA was positive, indicating that at an average level of Openness, every standard deviation increase in sports related content relevance resulted in 2.9 POMP units greater interest in following. For someone 1 standard deviation above the mean in Openness, the simple effect of sports related content was -0.2. For someone 1 standard deviation below the mean in Openness, the simple effect of sports related content was 6. In other words, lower levels of Openness were associated with greater sensitivity to sports-related content.

Figure 17

Interaction Plot for Openness and the Effect of PCA Sports Topic

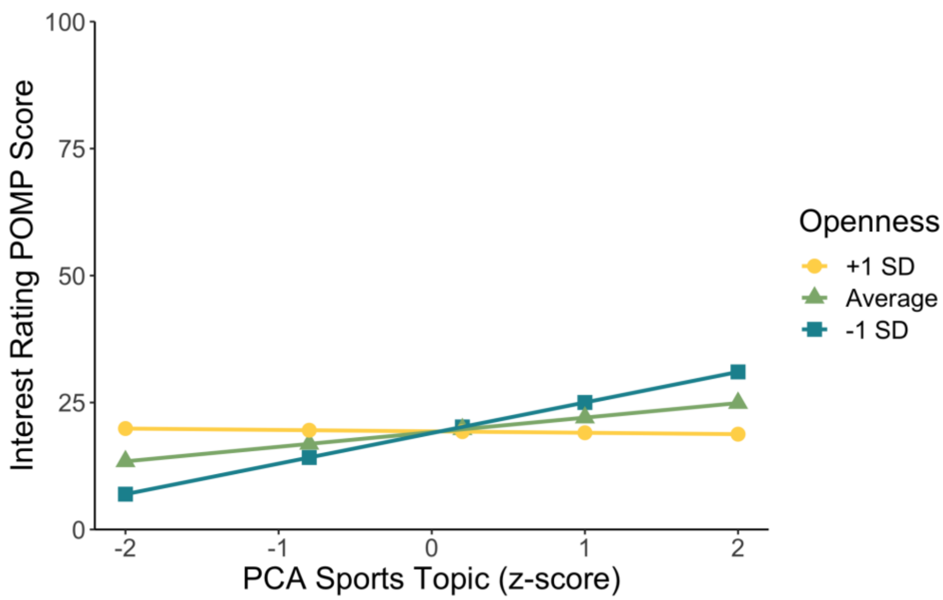
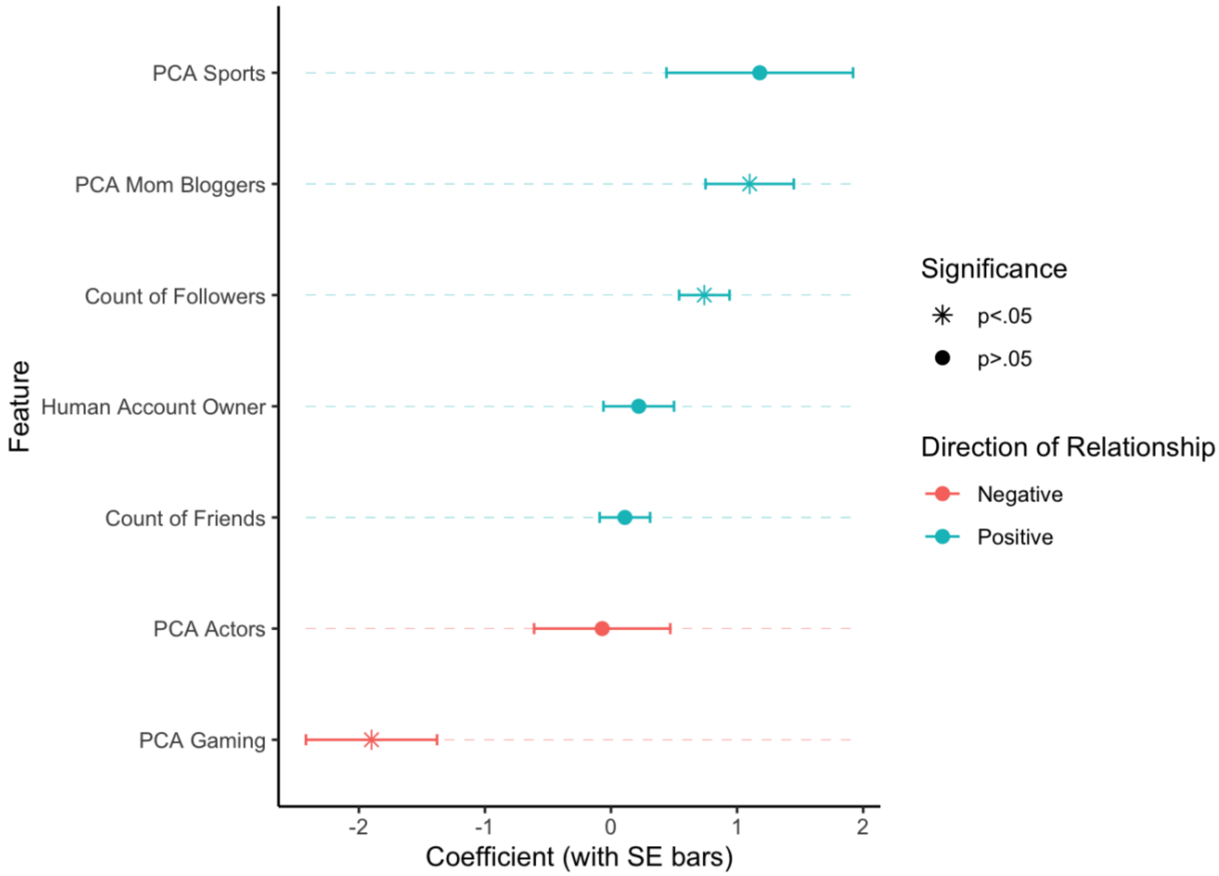


Figure 18

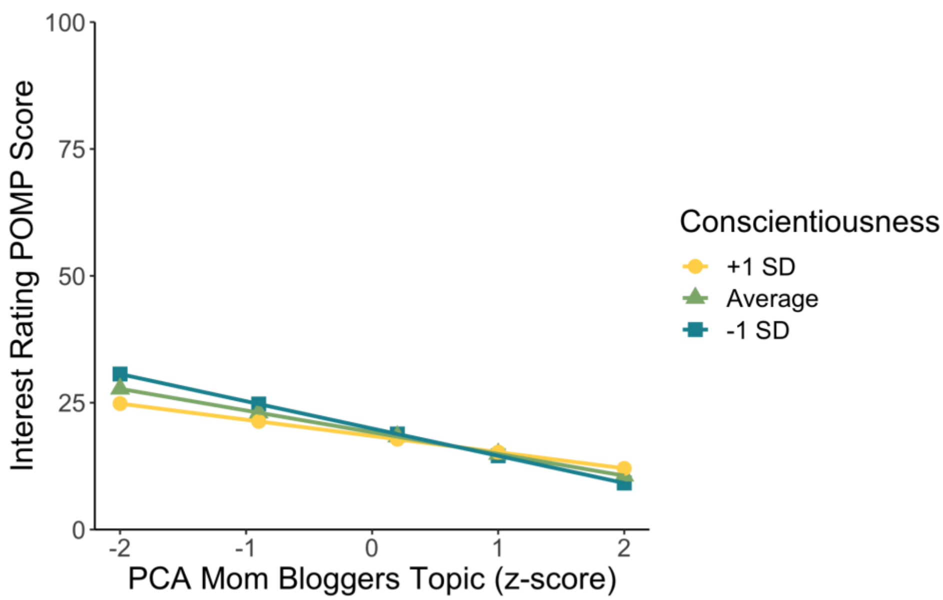
Moderating Effect of Z-scored Conscientiousness on the Relationship Between Z-scored Twitter Profile Features and Study 1 Account Interest Scores



Participant Conscientiousness significantly moderated the effect of count of followers and the PCA topics of Mom Bloggers and Gaming on account interest ratings (Figure 18). The Mom Bloggers PCA topic feature was the most positively significantly moderated by Conscientiousness (Figure 19). The main effect of Mom Bloggers PCA was negative, indicating that at an average level of Conscientiousness, every standard deviation increase in mom blog content relevance resulted in 4.3 POMP units less interest in following. For someone 1 standard deviation above the mean in Conscientiousness, the simple effect of mom blog related content was -3.2. For someone 1 standard deviation below the mean in Conscientiousness, the simple effect of mom blog related content was -5.4. In other words, lower levels of Conscientiousness were associated with greater sensitivity to mom blog-related content.

Figure 19

Interaction Plot for Conscientiousness and the Effect of PCA Mom Bloggers Topic



The effect of the Gaming PCA topic feature was most negatively moderated by Conscientiousness (Figure 20). The main effect of Gaming PCA was negative, indicating that at an average level of Conscientiousness, every standard deviation increase in gaming content relevance resulted in 6.8 POMP units less interest in following. For someone 1 standard deviation above the mean in Conscientiousness, the simple effect of gaming related content was -8.7 and for someone 1 standard deviation below the mean in Conscientiousness, the simple effect of gaming related content was -4.9. In other words, higher levels of Conscientiousness were associated with greater sensitivity to gaming-related content.

Figure 20

Interaction Plot for Conscientiousness and the Effect of PCA Gaming Topic

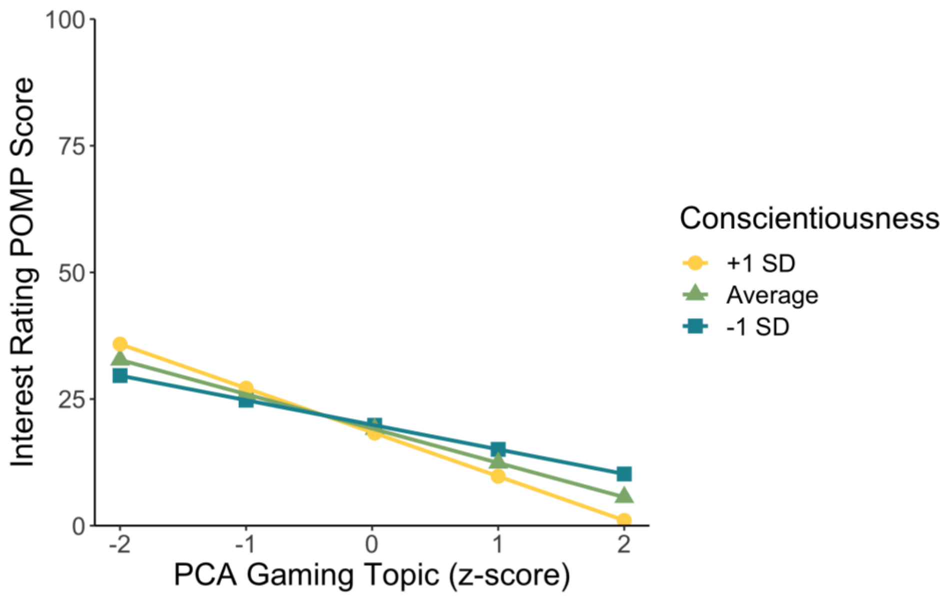
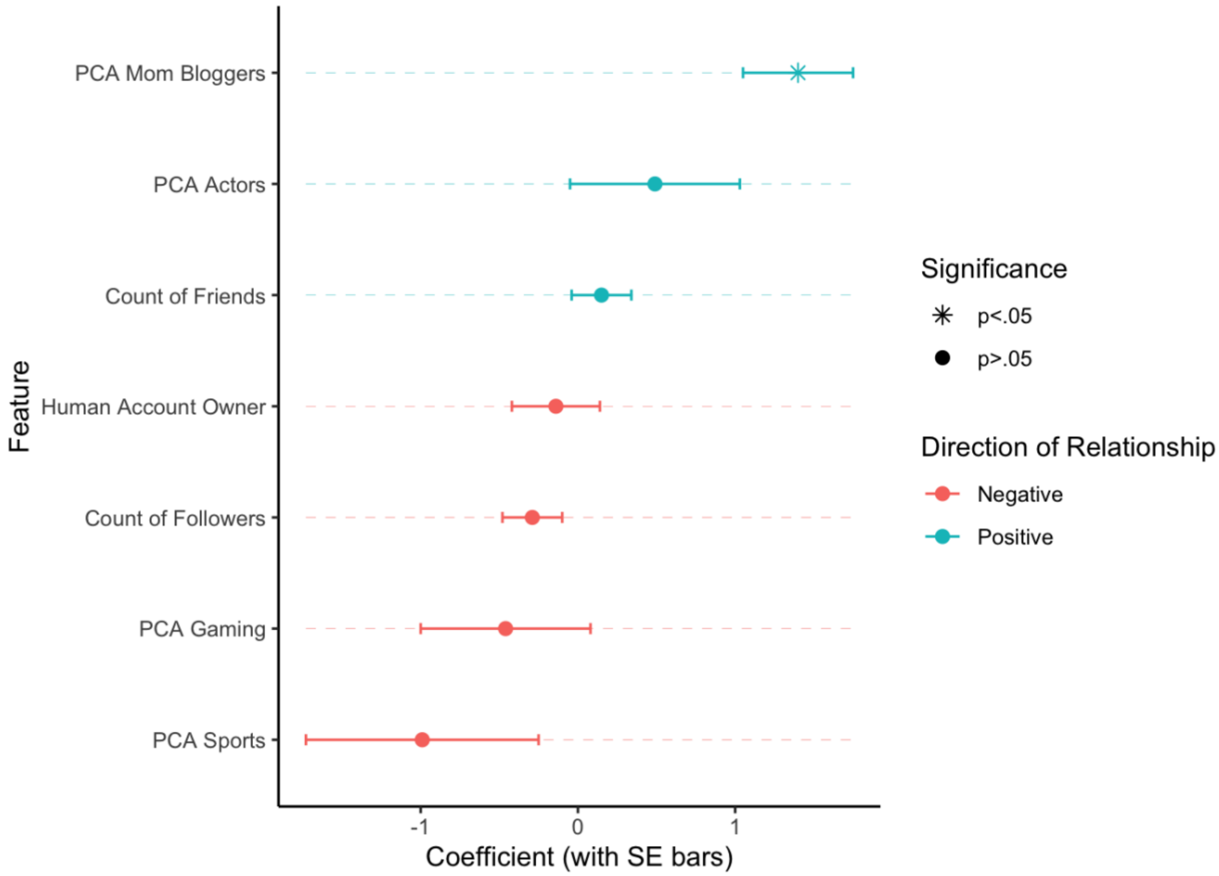


Figure 21

Moderating Effect of Z-scored Honesty-Propriety on the Relationship Between Z-scored Twitter Profile Features and Study 1 Account Interest Scores



Participant Honesty-Propriety significantly moderated the effect of the PCA topic Mom Bloggers (Figure 21). The effect of the Mom Bloggers PCA topic feature was most positively moderated by Honesty-Propriety (Figure 22). The main effect of Mom Bloggers PCA was negative, indicating that at an average level of Honesty-Propriety, every standard deviation increase in mom blog content relevance resulted in 4.3 POMP units less interest in following. For someone 1 standard deviation above the mean in Honesty-Propriety, the simple effect of mom blog related content was -2.9. For someone 1 standard deviation below the mean in Honesty-Propriety, the simple effect of mom blog related content was -5.7. In other words, lower levels of Honesty-Propriety were associated with greater sensitivity to mom blog-related content. Honesty-Propriety did not significantly negatively moderate any of the features tested in Study 1.

Figure 22

Interaction Plot for Honesty-Propriety and the Effect of PCA Mom Bloggers Topic

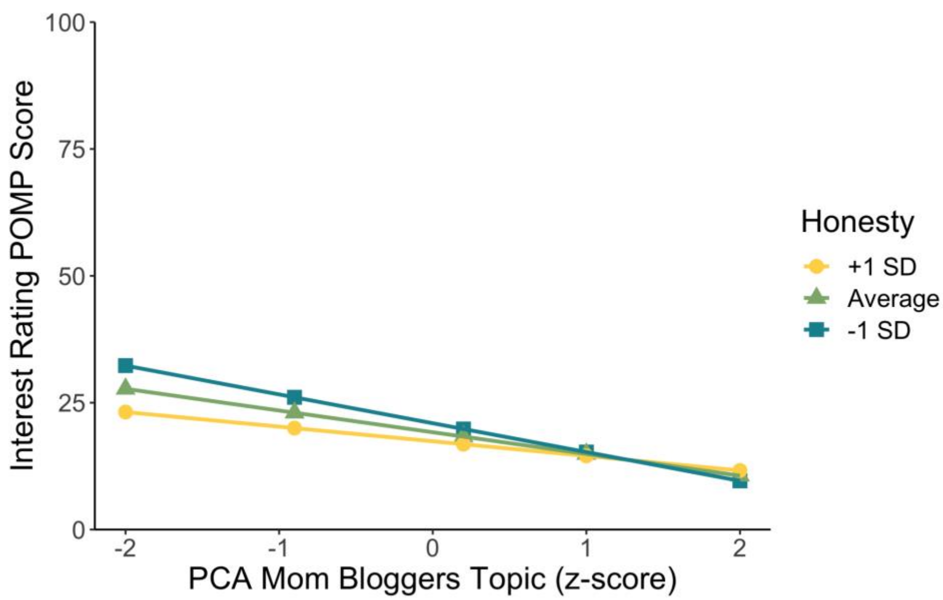
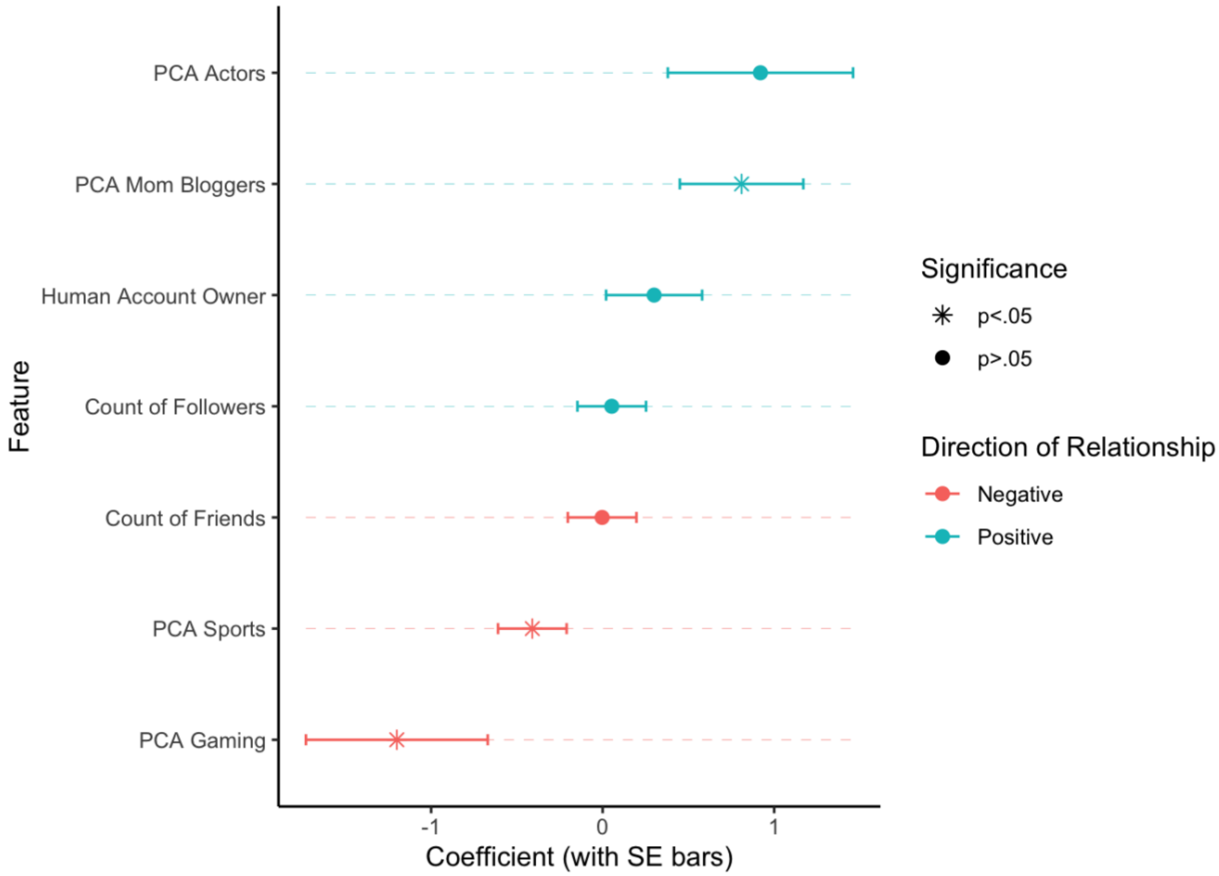


Figure 23

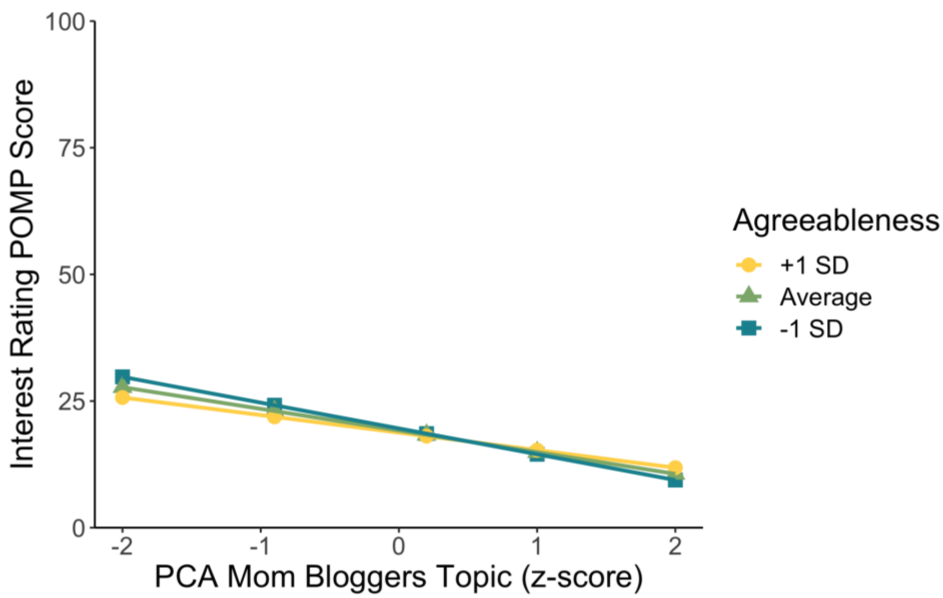
Moderating Effect of Z-scored Agreeableness on the Relationship Between Z-scored Twitter Profile Features and Study 1 Account Interest Scores



Participant Agreeableness significantly moderated the effect of the PCA topics of Mom Bloggers, Sports, and Gaming on account interest ratings (Figure 23). The Mom Bloggers PCA topic feature was the most positively significantly moderated by Agreeableness (Figure 24). The main effect of Mom Bloggers PCA was negative, indicating that at an average level of Agreeableness, every standard deviation increase in actor related content relevance resulted in 4.28 POMP units less interest in following. For someone 1 standard deviation above the mean in Agreeableness, the simple effect of actor related content was -3.47. For someone 1 standard deviation below the mean in Agreeableness, the simple effect of actor related content was -5.09. In other words, lower levels of Agreeableness were associated with greater sensitivity to mom blog-related content.

Figure 24

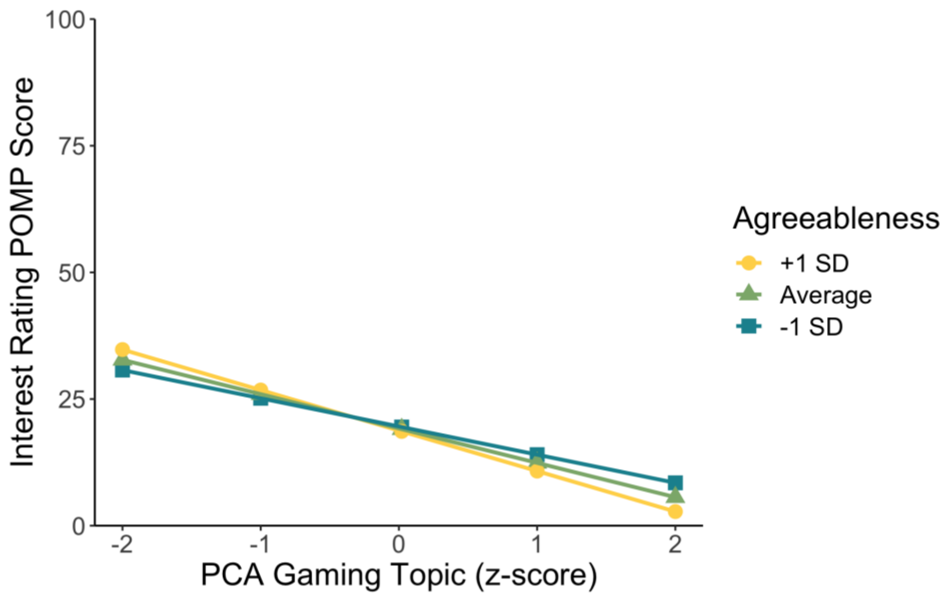
Interaction Plot for Agreeableness and the Effect of PCA Mom Bloggers Topic



The effect of the Gaming PCA topic feature was most negatively moderated by Agreeableness (Figure 25). The main effect of Gaming PCA was negative, indicating that at an average level of Agreeableness, every standard deviation increase in gaming related content relevance resulted in 6.8 POMP units less interest in following. For someone 1 standard deviation above the mean in Agreeableness, the simple effect of gaming related content was -8. For someone 1 standard deviation below the mean in Agreeableness, the simple effect of gaming related content was -5.6. In other words, higher levels of Agreeableness were associated with greater sensitivity to gaming-related content.

Figure 25

Interaction Plot for Agreeableness and the Effect of PCA Gaming Topic



Qualitative Responses

After participants rated their interest in following the Twitter accounts of the stimuli they were presented with, they were asked to reflect on what influenced them to be interested or not interested in following those accounts. Open-ended text responses were read, and themes were extracted. Participants frequently indicated that the primary topic Tweeted by the account was key to determining whether or not they were interested in following the account. Additionally, participants were not likely to follow accounts that did not align with their interests or hobbies. This theme indicates that profile topics may provide useful insight into the relationship between profile features and Twitter following decisions.

> “I gravitated towards accounts that post funny tweets and don't retweet or promote things regularly. I enjoy left-leaning political accounts as well, but not to the same extent. I mostly follow personal accounts rather than businesses or brands.”

> “I don’t like polluting my social media feeds with things I don’t care about.”

Participants also mentioned the authenticity of an account as an influential factor in their following decisions. They indicated a preference for accounts run by the account owner themselves, and noted they were not interested in following accounts that felt like they were run by an agent or corporation.

> “the authenticity of the account; is it a page listed as a blogger but really all their links are ads? Or are the contents genuine and entertaining.”

Another overarching theme observed in participant responses was that the Twitter accounts used as stimuli in Study 1 were not relevant to the interests of the participants. Below are several excerpts from text responses demonstrating participant reactions to the Twitter profiles stimuli used in the study:

> "I don't care about the twitters of Middle aged women moving to Idaho or someone's Mommy Blog. I don't follow those on my real twitter and I didn't follow them in this simulation."

> "I don't care much for nfl players or sports or politics and I'm not a mom and that was like all the accounts except for some actors and youtubers."

> "There are SO many of them that are just repetitive sports accounts, MLM mom marketers, and generic political commentary. There is virtually no variety here, and it just so happens that most of the ones presented are not of the tastes I would be remotely interested in."

Though results still showed significant relationships between account topic and account interest in Study 1, a set of Twitter stimuli more relevant to this sample may uncover even richer relationships between these variables.

Discussion

The findings of Study 1 first and foremost demonstrated the feasibility of the planned analyses. This feasibility was indicated both in mathematical outcomes and interpretability of results. Though I did not have hypotheses with expected outcomes, the findings generated in this study were not notably unusual or particularly unexpected. Additionally, in analyses examining the impact of Twitter profile features on following decisions and the moderating effect of personality traits, almost 50 multilevel models converged with minimal trimming of effects. Overall, the Study 1 results validated the approach and gave greater confidence in the analysis plan to be applied in Study 2 with a better set of account stimuli.

In addition to mathematical feasibility, exploratory results at all steps of analysis were found to be interpretable in Study 1. Correlations and random forests models highlighted differing strengths of personality traits in the relationship between traits and interest in following Twitter accounts. The principal components analysis resulted in a solution with coherent latent categories that captured known popular categories of accounts on Twitter and their relationship to personality traits. Finally, multilevel modeling uncovered interpretable patterns in the effect of individual Twitter profile features on interest in Twitter accounts and how those effects vary based on participant personality traits. Taken together, these analyses quantify and characterize these relationships in multiple ways and will be replicated and expanded upon in Study 2.

Beyond demonstrating feasibility of analysis techniques, the results of Study 1 provided preliminary findings about the nuanced relationship between personality traits and interest in following Twitter accounts. Although these data were originally collected with the purpose of evaluating the effect of mental health variables on Twitter following decisions, data-driven analyses revealed notable and interesting connections across personality traits as well. Overall, the Study 1 findings indicate that aspects of personality are reflected in the accounts that people follow on Twitter, though there is some heterogeneity across domains.

The first aim of Study 1 analyses was to examine if personality traits influence account following decisions. In both the correlations with interest in individual accounts and the random forests models reflecting aggregate interest, Neuroticism and Extraversion showed the strongest relationships with interest in following Twitter accounts. Additionally, Conscientiousness and Openness were predicted with notable accuracy by the random forests models. One factor driving the relationship with Neuroticism may be the use of stimuli Twitter profiles that were chosen for their prior relationship with mental health variables. Previous research has shown a connection between the Neuroticism and the mental health components of anxiety and depression (Duggan et al., 1990; Boyce et al., 1991; Saklofske et al., 1995; Muris et al., 2005). Another notable finding from Study 1 was that the strength of the relationship between personality traits and following decisions was generally in the same realm as demographic variables, giving us a sense of relative effect size for personality traits. These results align with previous research showing that personality and demographic variables have similar predictive power in a number of behavioral and life outcomes.

A number of features of Twitter profiles were examined to further explore factors that influence interest in following accounts. Qualitative responses from participants in Study 1 indicated that the authenticity of an account was a driving factor in their following decisions, specifically citing that they would prefer to follow accounts run by humans rather than brands. However, the analyses revealed almost no effect of human account owner on following interest in this set of stimuli Twitter accounts. This result may indicate that a human vs. non-human measure does not capture the form of authenticity that is relevant to interest in following Twitter accounts. The content of a Twitter account was also mentioned by participants as an influencing factor for following an account, particularly that they were interested in following accounts that put out content related to one of their interests or hobbies. All four PCA topic features tested in Study 1 did show a significant effect on account interest ratings, indicating that content of a Twitter profile is a driving feature in account following decisions.

These analyses revealed not only patterns of following behaviors, but also the effect that personality traits can have on these patterns. For example, sports content had a positive impact on interest in following, but this impact was even more strongly positive for those low in Neuroticism, low in Openness, and high in Extraversion. Similarly, gaming content had an overall negative impact on following but this effect was even more exaggerated for those high in Extraversion, high in Conscientiousness, and high in Agreeableness. These relationships indicate not only that content is a driving factor in following decisions, but also that this relationship changes based on the personality traits of the potential follower. The effect size of individual profile features was small when considered on their own. For example, the number of followers a Twitter profile has the largest positive impact on interest in follow of the features that were tested in Study 1. However, an additional 1 million followers is only associated with an increase in interest of 8.6 POMP units. Depending on the popularity of the account, adding an additional million followers is no small task and even at that, the increase in interest may still not be enough for a user to actually follow the account. These small effect sizes may indicate that there is an upper boundary to what any singular feature can contribute to a following decision. Rather than an individual feature influencing a following decision, several small features may come together to create a profile that is appealing to follow.

While these preliminary results start to uncover how personality is reflected in online environments, there are a number of considerations that will inform Study 2. Qualitative responses indicated that the accounts used in Study 1 did not necessarily match the topics that participants were interested in seeing or the topics that they actually follow on Twitter. While the accounts used as stimuli in this study were drawn from accounts actually followed by participants in Costello et. al (2021), these accounts were collected in 2016-2017 from a non-student sample and may not be relevant to the interests of this population. To simulate a more current and representative Twitter environment, stimuli in Study 2 will be drawn from actual followed accounts of Study 1 participants, a student sample representing a similar population as Study 2 participants. Additionally, collecting this new set of stimuli will allow me to examine relationships with a more representative sample of profiles that have not been selected based on their prior relationship to mental health variables.

Finally, the small number of Twitter profile features tested in Study 1 limit the interpretation of the relative effect size of these features as they represent just a small number of features present in a Twitter profile. The preliminary examination of the effect of these features has provided evidence that it is worth the investment of time to incorporate additional Twitter profile features that are more labor intensive to collect for Study 2 analyses. This includes gathering recent Tweets from stimuli profiles in order to extract various linguistic elements, as well as, utilizing a team of human coders to examine various perceived characteristics of stimuli accounts. This richer set of features will allow me to more fully explore the relationship between Twitter profile features and Twitter account interest in Study 2.

III. STUDY 2: PERSONALITY AND TWITTER FOLLOWED ACCOUNTS AND FEATURES THAT DRIVE THIS RELATIONSHIP

The purpose of Study 2 is to further examine the relationship between personality and Twitter account preferences, replicating analysis techniques that were shown to be feasible in Study 1 but with an updated set of stimuli. The stimuli accounts in Study 2 will be selected for relevance and popularity in the current research population, rather than a prior connection to mental health variables. In the qualitative responses for Study 1, participants noted that the accounts used as stimuli were not representative of accounts they would actually follow. Though the set of stimuli used in Study 1 was a set of Twitter accounts that were followed in real life by previous participants, those accounts were initially collected in 2016 and the participants were Reddit users. Twitter is a quickly evolving platform, and the four-year gap between the collection of stimuli and the data collection of Study 1 plus the difference in populations may have contributed to participant dissatisfaction with the stimuli accounts. I will mitigate this issue in Study 2 by collecting a set of stimuli consisting of popular Twitter accounts that were followed by the participants in Study 1 as of 2021 and showing them to a new sample of participants.

Another aim of Study 2 is to further explore features that drive Twitter following decisions by expanding the breadth of Twitter profile features that are analyzed. Study 1 results indicated that individual features of Twitter profiles significantly drive Twitter following decisions and that personality traits play a role in moderating this relationship. This has provided evidence that it is worth the investment of time to incorporate additional Twitter profile features that are more labor intensive to collect. This includes gathering recent Tweets from stimuli profiles in order to extract various linguistic elements, as well as utilizing a team of human coders to examine various perceived characteristics of stimuli accounts.

When a potential follower views a Twitter profile, they perceive attributes of the account that can inform whether or not they would like to follow that account. To better understand if these subjective account qualities influence Twitter account interest, I will examine two broad categories of perceived attributes. The first is perceived account characteristics, which includes perception of the account's personality traits, as well as positive and negative affect. Measuring the perception of these account characteristics can inform if social processes, such as homophily, play into following decisions. Additionally, I will examine perceptions of who an account appeals to, such as men vs. women or liberals vs. conservatives. These attributes are modeled after the demographic benchmarks used in Study 1 and 2 and will reveal if appeal to a particular group influences following decisions. To capture how these attributes are generally perceived in this set of Twitter stimuli accounts, I will enlist the help of graduate student coders to provide their expert judgment.

Tweets are another prominent element of a Twitter account and examining linguistic features of stimuli account tweets can uncover a variety of characteristics that may factor into a user's following decisions. Dictionary-based analysis approaches are useful for assessing particular linguistic categories of interest. In Study 2, I will look at broad categories of emotion and affect, to examine the influence of sentiment on following decisions. I will also examine a number of social categories such as prosocial behavior and mental health language to better understand the influence of psychological states present in Tweets. Finally, I will incorporate an open-vocabulary analysis to look for additional data-driven linguistic topics. These analyses will uncover if language used in Tweets can inform and drive following decisions.

Taken together, these methodological updates motivate two goals for Study 2: (1) further explore the relationship between personality and Twitter account interest with an updated set of stimuli and (2) expand our understanding of which Twitter profile features influence Twitter account interest and how personality traits moderate that relationship.

Preregistration

Study 2 research questions and method were preregistered on the Open Science Framework at: https://osf.io/yuh9q. Study 2 analysis procedures were preregistered on the Open Science Framework at: https://osf.io/rgxej.

Materials and Methods

Data Collection Procedure

Recruitment for Study 2 took place online within the University of Oregon's Department of Psychology subject-pool website and within the Prolific data collection services website. Participants who read and agreed with the informed consent were redirected to the Qualtrics questionnaire. They completed the questionnaire as described below and provided demographic information. Participants were then shown screenshots of 100 Twitter profiles and asked how interested they were in following each account. Finally, participants responded to open-ended questions about their Twitter account preferences, provided their Twitter handle, and responded to a series of questions about their Twitter usage. After survey completion, participants' Twitter profiles were scraped to collect their Tweets, a list of their followers, and a list of their friends (accounts they are following on Twitter) using the Rtweet package (M. Kearney, 2019).

Participants. Study 2 consisted of two samples of participants. Initially, I planned to only collect one sample of participants from the University of Oregon. However, when the initial data collection ended with less participants than anticipated, I decided to collect another sample of participants using Prolific data collection services, which sampled from the general US population. Survey responses in each sample were screened for anomalous responding and missingness by a blinded analyst.

In the first sample, N = 148 participants were recruited through the University of Oregon human-subjects pool which consisted of students from introductory psychology and linguistics courses. Data collection was open for 8 weeks in Spring 2022 and students could participate at their convenience. Those who completed the survey were compensated with course credit. Upon screening, the analyst determined that 1 participant’s data should be excluded for anomalous responding, resulting in a sample of N = 147 participants for analysis. Participants each rated 100 Twitter profiles for a possible 14,700 profile ratings. Of this, there were only 291 skipped ratings, giving us less than 2% missing data. Participants ranged in age from 18-35 years old with an average age of 19.6 years old.

In the second sample, N = 150 participants were collected through Prolific data collection services in June of 2022. All participants met the prescreening criteria of having a currently active Twitter account that they use at least once a month, speaking English as a first language, and residing in the United States. Those who completed the survey were compensated with $6. The analyst determined that all participants' data should be included, and no exclusions were recommended due anomalous responding or excessive missingness. In this data set, N = 150 participants each rated 100 Twitter profiles for a possible 15,00 profile ratings. Of this, there were only 89 skipped ratings, giving us less than 1% missing data. Participants ranged in age from 18-25 years old with an average age of 21.8 years old.

Demographics for both samples in Study 2 are shown in Tables 12-14. For Study 2 analyses, data from the two samples were merged together for a total of N = 297 participants. In this combined sample, participant age ranged from 18-35 with an average age of 20.7 years old.

Table 12

Participant Gender for Study 2

Gender	UO Sample	Prolific Sample	Combined Sample
Man	40 (27%)	75 (50%)	115 (39%)
Woman	96 (66%)	67 (45%)	163 (55%)
Nonbinary	9 (6%)	8 (5%)	17 (5%)
Another Identity	2 (1%)	0 (0%)	2 (1%)

Table 13

Participant Race for Study 2

Race	UO Sample	Prolific Sample	Combined Sample
American Indian or Alaska Native	5 (3%)	0 (0%)	5 (2%)
Asian	11 (7%)	19 (13%)	30 (10%)
Black or African American	6 (4%)	21 (14%)	27 (9%)
Native Hawaiian or Pacific Islander	3 (2%)	0 (0%)	3 (1%)
White	95 (65%)	98 (65%)	193 (65%)
Other	6 (4%)	3 (2%)	9 (3%)
More than one race	17 (12%)	9 (6%)	26 (9%)
Not reported	4 (3%)	0 (0%)	4 (1%)

Table 14

Participant Ethnicity for Study 2

Ethnicity	UO Sample	Prolific Sample	Combined Sample
Hispanic or Latino	28 (19%)	26 (17%)	54 (18%)
Not Hispanic or Latino	119 (81%)	124 (83%)	243 (82%)

Self-report Measures. Participants completed self-reports of personality traits using a combination of two measures. The Big Five traits (Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Openness) were measured using the Big Five Inventory 2 (BFI 2; Soto & John, 2017), consisting of 60 short statements rated on a scale from one (Disagree strongly) to five (Agree strongly) with a neutral point of three (neither agree nor disagree). Eight items from Questionnaire Big Six measure were used to capture the sixth domain, Honesty-Propriety (Thalmayer et al., 2011). These measures showed expected and adequate internal consistency with alpha coefficients for the BFI-2 scales ranging from .76 for Agreeableness to .92 for Neuroticism and an alpha coefficient for Honesty-Propriety at .62. Though not analyzed as a part of this dissertation participants also completed the following self-report measures: Satisfaction with Life Scale (SWLS; Diener et al., 1985), Scale of Positive and Negative Experiences (SPANE; Diener et al., 2009), PROMIS Depression, Anger, and Anxiety scales (Pilkonis et al., 2011), Trauma Symptom Questionnaire (Brewin et al., 2002).

Twitter Account Stimuli and Ratings. After completing the self-report measures, participants viewed screenshots of 100 Twitter profiles in a randomized order. Each Twitter stimuli screenshot included the banner, profile picture, account name, account bio, account's number of friends and followers, and about 3-6 of the account's most recent Tweets. Participants could scroll to view all components of the screenshot, simulating the experience of viewing a profile page on the Twitter platform. At the end of the screenshot, participants were asked how interested they were in following that account on a 4-point scale (not at all, a little, moderately, very), they could also choose to skip the profile. For the purposes of this study, this measure will be referred to as an *account interest rating*.

To create the set of accounts used for the stimuli in Study 2, I gathered a list of all Twitter accounts followed by the participants in Study 1 (as of March 2021). This resulted in a set of 26,725 friend connections with 17,038 unique accounts represented. The number of participants in Study 1 that follow each account was calculated and the top 100 accounts most frequently followed were selected. The number of Study 1 participants that follow these accounts range from 30 (e.g., Elon Musk) to 8 (e.g., Tom Holland). The selected accounts are generally popular accounts, with follow counts ranging from former president Barack Obama with 130 million followers to internet celebrity Sarah Baska with about half a million followers.

This set of 100 stimuli was used for the UO data collection, which was originally intended to be the only data collection for Study 2. The original set of 100 stimuli included 11 profiles that were specific to a UO student or Oregon population (9 were related to the University, 1 was Oregon senator Jeff Merkley and 1 was Oregon Governor Kate Brown). When I collected the second sample of data using the Prolific participants, those 11 accounts were replaced with the next 11 most followed accounts by Study 1 participants that were not related to UO or the state of Oregon. Study 2 analyses will focus on the 89 Twitter profile stimuli that overlapped between the UO sample and the Prolific sample.

Features of Twitter Profile Stimuli

In order to analyze what features of the Twitter stimuli accounts drove participants' interest, I extracted features of each of the profiles. Features analyzed in Study 2 will include all features analyzed in Study 1 plus additional account metadata features, account characteristics perceived by a team of coders, and linguistic features present in Tweets from the stimuli accounts.

Account Metadata. Features providing information about the Twitter account itself were extracted from the Twitter API (application programming interface) using the RTweet package (M. Kearney, 2019). *Count of Friends* refers to the number of accounts that a given stimuli Twitter profile follows. *Count of Followers* refers to the number of accounts that follow a given stimuli Twitter account. *Average Word Count* indicates the average number of words per Tweet and was calculated with each stimuli account’s 100 most recent Tweets. *Tweet Frequency* indicates the average number of Tweets made by an account per day (calculated using the account’s 100 most recent tweets). *Ratio of Retweets to Tweets* indicates the quantitative relationship between the number of Retweets and total number of Tweets that are not Retweets.

Perceived Account Characteristics. Some profile features of interest had potential for subjectivity based on the perception of the viewer. To quantify these characteristics, a team of six psychology graduate students viewed the stimuli profiles and provided their expert judgements. These judgements were averaged across the six raters for use in Study 2 analyses. To assess agreement among the raters, I calculated intraclass correlations ($ICC_{3,k}$) using a two-way mixed effects model examining consistency of the mean of ratings from all judges (Shrout & Fleiss, 1979; McGraw & Wong, 1996).

Perceived Advertising (*ICC* = .84) was assessed by asking, "How much does this profile feel like advertising?" rated on a scale from not at all (1) to extremely (5). *Perceived Positive Affect* (*ICC* = .79) was assessed by asking, "How positive is the content of this profile?" rated on a scale from not at all (1) to extremely (5). *Perceived Negative Affect* (*ICC* = .86) was assessed by asking, "How negative is the content of this profile?" rated on a scale from not at all (1) to extremely (5).

Perceived personality traits of each account were measured using the Ten-Item Personality Inventory (TIPI; Gosling et al., 2003). This measure consists of 10 sets of adjectives rated on a scale from Disagree strongly (1) to Agree strongly (7) with a neutral point of neither agree nor disagree (4). Responses to these adjectives were scored for: *Perceived Openness* (*ICC* = .85), *Perceived Conscientiousness* (*ICC* = .81), *Perceived Extraversion* (*ICC* = .75), *Perceived Agreeableness* (*ICC* = .86), and *Perceived Neuroticism* (*ICC* = .77). Finally, *Human Account Owner* indicates that the Twitter account represents a single person rather than a brand, organization, or group of people. To determine this metric, I reviewed each stimuli Twitter account and coded it as either a human or non-human account.

Perceived Appeal. Coders were also asked to assess what groups the stimuli profiles would be most appealing to. These judgments were made on a five-point scale with opposing poles representing two different groups (indicated as 1 or 5) and a neutral point indicating the profile was appealing to neither group, both groups, or it was hard to tell (3). *Perceived Gender Appeal* (*ICC* = .89) was assessed by asking how much a profile's content would appeal to primarily men (1) vs. primarily women (5). *Perceived Age Appeal* (*ICC* = .84) was assessed by asking how much a profile's content would appeal to primarily young people (1) vs. primarily old people (5). *Perceived Political Appeal* (*ICC* = .85) was assessed by asking how much a profile's content would appeal to primarily liberal people (1) vs. primarily conservative people (5). *Perceived SES Appeal* (*ICC* = .67) was assessed by asking how much a profile's content would appeal to primarily poor people (1) vs. primarily rich people (5).

Principal Component Analysis Dimensions. To uncover latent categories of Twitter accounts within the set of Study 2 stimuli, I used the dimensionality-reducing technique of principal components analysis (PCA) on account interest ratings, treating each of the 89 accounts as a variable. The scree plot testing a number of components indicated potential solutions between 5 and 8 components. I examined solutions with 5, 6, 7 and 8 components with varimax rotation to enhance interpretability and found the 7-component solution to be the most interpretable (see Appendix B). This solution was additionally tested with an oblimin rotation, but I found that the solution was not notably altered by this rotation. The loadings from these PCA components will be used as features to represent content categories present in this set of stimuli Twitter accounts.

I reviewed the stimuli Twitter accounts with the highest loadings for each of the seven components to determine the associated latent categories of accounts. Component 1 was determined to be related to *Celebrities.* Accounts that loaded highly on this topic tended to be singers such as Harry Styles and actors such as Zendaya and Tom Holland. The musicians tended to skew towards pop music, notably all members of the former band One Direction were included in this component. Component 2 was determined to be related to *Sports.* Accounts that loaded highly on this topic tended to be sports news outlets and athletes. This included organizations such as ESPN and Sports Center as well as players including Steph Curry and Lebron James. There was a particular focus on basketball athletes. Component 3 was determined to be related to *Rap/R&B.* Accounts that loaded highly on this component were primarily popular rap or R&B musicians. This included artists like Kendrick Lamar, Tyler the Creator, and SZA. This component also intersected with adjacent pop stars such as Rihanna and the Weekend. Component 4 was determined to be related to *Mainstream Influencers.* Accounts that loaded highly on this topic included both reality stars that appear on TV shows as well as internet personalities. Notably, all four of the Kardashian/Jenner family members that were included in the stimuli account set loaded highly on this component. Additionally, the two TV hosts included in this stimuli set, Ellen DeGeneres and Jimmy Fallon, loaded highly on this component. Component 5 was determined to be related to *Liberal Politicians.* Particularly, this included liberal politicians like Bernie Sanders and Alexandria Ocasio-Cortez. It also included more official political accounts like the POTUS44 account for Barack Obama. Component 6 was determined to be *Traditional Media Comedy*. Accounts that loaded highly on this component tended to have a focus on humor, such as the Onion. This component included several actors that we associated with humor in TV or movies like Seth Rogen and John Krasinski. Finally, component 7 was determined to be *Social Media Personalities.* This component also had indicated a comedy theme, but only included accounts that were known through online outlets like YouTube. This included accounts like Noel Miller and Caucasian James.

Linguistic Features. I assessed the linguistic content of profiles by extracting the last 100 Tweets from each of the stimuli Twitter accounts. This sampling approach assumes that accounts have a stable linguistic style throughout their Tweets and that this style is conveyed in both the screenshots that participants see, and the sample of Tweets used to assess linguistic features. Additionally, from a practical standpoint, this sample of Tweets allows me to create a set of text for each user large enough to be assessed by linguistic analysis methods. Prior to analysis, Tweets were cleaned by removing URLs, greater-than signs (>), less-than signs (<), ampersands (&), and “RT” (indicates the classic version of retweet). These symbols are generally not handled well in linguistic analyses and do not provide any relevant substantive information about the sample of Tweets.

NRC Positive Sentiment and *NRC Negative Sentiment* indicate on average, how positive or negative a user’s set of tweets is. This was assessed by using the Noncommercial Research (NRC) sentiment lexicon (Mohammad & Kiritchenko, 2015) a sentiment dictionary designed for and validated with tweets. This lexicon consists of over 14,000 words where each word has been assigned a score for positive/negative sentiment (ranging from –6.93 to 7.53). This set of words was then compared to the words in a user’s tweets, and an average positive and negative sentiment for each user was calculated. The same lexicon was also used to assess eight distinct emotion categories in Tweets including: *Disgust, Joy, Anger, Fear, Sadness, Surprise, Anticipation, and Trust*.

Additional linguistic categories were scored for individual profiles using the Linguistic Inquiry Word Count software (LIWC-22; Boyd et al., 2022). This dictionary-based approach includes over 100 different categories aimed at assessing a collection of social and psychological states. To do this, LIWC reads a user’s set of Tweets and compares each word in the text to the list of dictionary words in a given category. Then, a percentage of total words in the text that match that category is calculated, giving us a category score. A set of 13 LIWC categories were chosen as features to include in Study 2 analyses based on representation of categories that were not already covered by other linguistic approaches, relevance to constructs motivating this research, and measures internal consistency (Boyd et al., 2022). These categories include the psychological drives of *Affiliation*, *Achievement*, and *Power*. A set of social behavior categories were selected which include *Prosocial Behavior, Politeness, Interpersonal Conflict, Moralization and Communication.* The content categories of *Mental Health, Substances,* and *Sexual* were included to address potentially relevant topics not covered by other features. Finally, the word type categories of *Swear Words* and *Netspeak* (emojis and abbreviations) were selected to address patterns of language associated with social media posts.

Finally, I used the data-driven approach of Latent Dirichlet Allocation (LDA) to extract topic categories present in Tweets at the profile level. This means there were 89 documents in the LDA analysis (one for each stimuli profile), with each document consisting of the 100 most recent Tweets in a profile. LDA is an unsupervised machine learning algorithm which identifies latent topic information among large document collections (Blei et al., 2003). Similar to other data reduction methods (e.g., factor analysis), researchers must choose the number of latent topics to fit. A combination of perplexity (a quantitative index) and subjective interpretability was used to decide how many topics to fit. Perplexity measures how poorly a probability model predicts a sample. More specifically, the normalized log-likelihood of a held-out test set of data is used to determine how "surprising" the test set is, considering the model. This measure indicated 5-7 topics as possible solutions. For each of these solutions, a set of words most likely to appear in each topic were examined to identify the latent categories. However, upon testing out these solutions, I determined that the topics were difficult to interpret and did not form cohesive categories that would inform analyses. Typically, a larger number of text inputs allow for more cohesive topics to be formed in the model results. The input of text from only 89 stimuli profiles likely limited the emergence of identifiable and distinct topics. Therefore, these model results will not be included as features to be analyzed.

Linguistic Validity. Though Study 2 participants made account interest ratings from a screenshot limited to the most recent 3-6 Tweets from that account, linguistic analyses were based on the account's 100 most recent Tweets in order to provide enough text for the linguistic analyses. To examine how representative these linguistic features are of what subjects are seeing in the screenshots, correlations were calculated between the human coded feature of *Perceived Negative Affect* and the linguistic feature *Negative Text Sentiment* (r = .44), as well as *Perceived Positive Affect* and the linguistic feature *Positive Text Sentiment* (r = .21). We would not expect these measures to be perfectly correlated as text sentiment focuses on the valence of words in tweets and perceived affect includes all visual elements of a Twitter profile. However, seeing a positive relationship between the linguistic and perceived features provides evidence of a stable linguistic style across profile screenshots and tweet samples.

Analytic Procedure

The analytic procedure in Study 2 remained the same as the feasibility-tested procedure in Study 1 (see the analytic procedure section of Study 1 for full details).

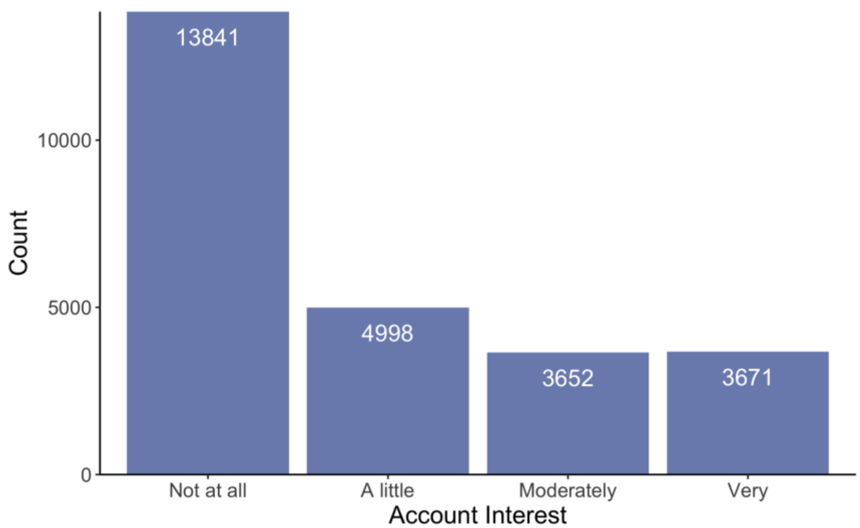
Results

Aim 1: Associations Between Personality Traits and Interest in Following Twitter Accounts

Stimuli Account Ratings. As a preliminary analysis, I examined the distribution of participant interest ratings across all 89 Twitter stimuli profiles. As shown in Figure 26, the distribution is positively skewed, with 53% percent of ratings indicating no interest in following a given account and 47% indicating at least a little interest in following a given account. The mean interest rating across all participants and profiles was 1.89 (*SD* = 1.10) out of a 4-point scale.

Figure 26

Distribution of Study 2 Account Interest Ratings



Note. There were 26,162 total ratings.

Correlations. To examine the strength of the relationship between participant personality traits and Twitter profile account interest ratings, I calculated Pearson product-moment correlations (Table 15). There were a number of significant relationships between personality traits and account interest ratings at the individual account level. Figures 27-32 visualize the twenty highest correlated accounts with each personality trait and the direction of those correlations, identifying the strongest relationships with individual accounts. These figures can also help us compare the strength of these relationships across traits. For example, we can see the highest correlated accounts with Extraversion and Neuroticism generally show stronger correlations than the highest correlated accounts with Openness and Conscientiousness.

Table 15

Correlations Between Study 2 Twitter Account Interest and Personality Traits

Twitter Account	Account Description	Extraversion	Neuroticism	Honesty	Agreeableness	Openness	Conscientiousness
KylieJenner	Celebrity, Makeup	**.37*****	-.06	-.20***	.06	-.13*	.01
rihanna	Singer, Makeup, Fashion	**.36*****	.10	-.08	.21***	.05	.09
Drake	Rapper, Singer	**.33*****	-.15**	-.15*	.06	-.14*	.01
jimmyfallon	Comedian, TV Host	**.33*****	-.13*	-.03	.10	-.07	.15*
KimKardashian	Celebrity, Fashion	**.32*****	-.02	-.17**	.10	-.05	.03
thegreatkhalid	Singer	**.32*****	-.02	-.10	.12*	.06	.11
KendallJenner	Reality Star, Model	**.31*****	.01	-.12*	.03	-.12*	.03
trvisXX	Rapper	**.30*****	-.19**	-.20***	.02	-.08	-.03
khloekardashian	Reality Star, Fashion	**.30*****	-.01	-.08	.05	-.10	.05
TheEllenShow	Talk Show	**.29*****	-.13*	-.03	.07	-.09	.13*
theweeknd	Singer	**.28*****	-.13*	-.12*	.07	-.01	.09
BuzzFeed	Digital Entertainment	**.28*****	-.04	-.10	.12*	-.03	.11
chancetherapper	Rapper	**.27*****	-.09	-.12*	.03	-.03	.05
justinbieber	Singer	**.27*****	-.02	-.04	.10	-.11	.11
DavidDobrik	YouTuber	**.26*****	-.05	-.10	.01	-.25***	.06
Zendaya	Actor	**.26*****	.12*	.00	.24***	.03	.07
sza	Singer	**.26*****	.06	-.13*	.15**	.10	-.03
kendricklamar	Rapper	**.26*****	-.06	-.21***	.03	.02	-.08
asvpxrocky	Rapper	**.26*****	-.11	-.20***	.00	-.06	-.01

Table 15 Continued

Twitter Account	Account Description	Extraversion	Neuroticism	Honesty	Agreeableness	Openness	Conscientiousness
SportsCenter	Sports, News	**.24*** **	-.20***	-.07	.02	-.12*	.08
JoeBiden	Politician	**.24*** **	.01	-.02	.14*	.05	.06
youngthug	Rapper	**.24*** **	-.10	-.23***	-.03	-.08	-.07
zaynmalik	Singer	**.23*** **	.05	.00	.11	.04	.12*
POTUS44	Politician	**.23*** **	.09	.02	.15**	.12*	.02
BarackObama	Politician	**.22*** **	.05	.04	.17**	.12*	.10
Harry_Styles	Singer	**.22*** **	.18**	-.02	.18**	.10	.09
BleacherReport	Sports, News	**.22*** **	-.20***	-.15**	-.06	-.18**	.00
chrissyteigen	Internet Personality	**.21*** **	.00	-.04	.05	-.15*	.04
MileyCyrus	Singer	**.21*** **	.19**	.05	.18**	.09	.08
onedirection	Pop Band	**.21*** **	.16**	-.01	.11	.08	.16**
tylerthecreator	Rapper	**.20*** **	.08	-.19***	-.04	-.03	-.15*
KamalaHarris	Politician	**.20*** **	.09	.01	.09	.03	.00
tanamongeau	YouTuber	**.20*** **	.04	-.11	.03	-.10	-.01
PostMalone	Rapper	**.19*** **	.04	.01	.08	.04	.12*
MichelleObama	Former First Lady	**.19*** **	.12*	.06	.17**	.11	.03
LiamPayne	Singer	**.18** **	.07	-.09	.11	-.03	.11
KidCudi	Rapper	**.18** **	.01	-.16**	-.05	.01	-.07
quenblackwell	Social Media Personality	**.17** **	.14*	-.01	.14*	.07	.01
lilyachty	Rapper	**.17** **	.00	-.15**	-.02	.01	-.03

Table 15 Continued

Twitter Account	Account Description	Extraversion	Neuroticism	Honesty	Agreeableness	Openness	Conscientiousness
Dame_Lillard	Basketball Player	**.16****	-.15**	-.12*	.03	-.08	-.02
jamescharles	Beauty Influencer	**.16****	-.03	-.04	-.02	-.08	.02
codyko	YouTuber	**.15***	.07	-.08	.07	-.02	-.05
jaden	Rapper	**.15****	.09	-.13*	.04	.02	.06
johnkrasinski	Actor	**.15***	-.02	.03	.09	.04	.05
Lin_Manuel	Composer, Actor	**.14***	.07	-.05	.01	.12*	.01
Nick_Colletti	Comedian	**.13***	.07	.01	.00	.00	-.01
colesprouse	Actor	**.12***	.04	.02	.11	-.06	.08
Sethrogen	Comedian	**.12***	.01	-.08	.03	-.01	-.06
dylanobrien	Actor	**.12***	.11	-.03	.11	.06	.03
shanedawson	YouTube	**.09**	-.02	-.08	.02	-.09	.07
elonmusk	Business, Tech	.21***	**-.27*****	-.13*	-.01	-.08	.07
troyesivan	Singer	.06	**.26*****	-.12*	-.09	.16**	-.02
theestallion	Rapper	.08	**.24*****	-.01	.08	.08	.03
StephenCurry30	Basketball Player	.21***	**-.23*****	-.01	.10	-.08	.06
espn	Sports, News	.22***	**-.22*****	-.04	.06	-.13*	.09
bretmanrock	Beauty Influencer	.10	**.22*****	-.03	.03	-.07	.01
wojespn	Sports Columnist	.13*	**-.22*****	-.05	.00	-.08	.04
KingJames	LeBron James, Basketball	.20***	**-.21*****	-.06	.10	-.09	.06
BernieSanders	Politician	-.03	**.20*****	.02	.05	.14*	-.05

Table 15 Continued

Twitter Account	Account Description	Extraversion	Neuroticism	Honesty	Agreeableness	Openness	Conscientiousness
stephenasmith	Sports Commentator	.13*	**-.18****	-.04	.02	-.03	.01
WorldWideWob	NBA Twitter Personality	.06	**-.17****	-.06	-.04	-.11	-.01
SarahBaska	YouTuber	.11	**.17****	-.06	.01	-.01	.01
snitchery	Makeup, Cosplay	-.15*	**.16****	-.02	-.12*	.10	-.15**
dog_rates	Cute Content, Humor	-.03	**.14***	.09	.11	.01	.01
thenoelmiller	Comedian, YouTuber	.05	**.14***	-.04	.06	.00	-.08
AOC	Politician	.02	**.13***	.01	.09	.13*	.00
dylansprouse	Actor	.10	**.13***	.04	.08	.00	.02
caseykfrey	Social Media Personality	.09	**.11**	-.10	.02	.05	-.09
MrBeast	YouTuber	.06	**-.11**	.00	-.01	-.05	.05
CaucasianJames	Twitter Personality	.06	**.09**	-.02	.04	.06	.02
archillect	AI Created Content	.03	**-.08**	-.08	-.02	.00	.04
kanyewest	Rapper	.26***	-.17**	**-.28*****	-.01	-.04	-.11
LILUZIVERT	Rapper	.11	-.01	**-.14***	-.04	-.11	-.1
DemetriusHarmon	Social Media Personality	.07	.00	**-.10**	.00	.04	-.01
TomHolland1996	Actor	.13*	-.02	.04	**.17****	.03	.11
WORLDSTAR	Entertainment News	.16**	-.05	-.16**	**-.17****	-.07	-.02
NiallOfficial	Singer	.13*	.08	.01	**.15***	.09	.15**
RobertDowneyJr	Actor	.12*	-.04	.07	**.13***	.09	.08
ericandre	Comedian, Actor	.03	-.10	-.09	**-.12***	-.01	-.11

Table 15 Continued

Twitter Account	Account Description	Extraversion	Neuroticism	Honesty	Agreeableness	Openness	Conscientiousness
SenSanders	Politician	-.02	.19**	.06	.08	**.21*** **	-.05
TheOnion	Humor News, Satire	-.14*	.09	.03	-.01	**.15** **	-.14*
Luke5SOS	Singer	.10	.13*	.00	.08	**.14***	.13*
JordanPeele	Comedian, Filmmaker	.03	.02	-.01	.05	**.14***	-.10
YourAnonCentral	Hacking Collective	.00	-.03	-.01	.05	**.06**	-.01
netflix	Streaming Platform	.14*	.01	.02	.12*	.07	**.18** **
BrotherNature	Social Media Personality	.03	.03	-.04	.06	.01	**.07**
richbrian	Rapper	.02	-.01	-.02	.01	.02	**.02**

Note. Bolded values indicate the highest absolute correlation for each stimuli Twitter account. * correlation is significant at the 0.05 level, ** correlation is significant at the 0.01 level, and *** correlation is significant at the 0.001 level.

Figure 27

Twenty Highest Correlated Twitter Accounts with Extraversion in Study 2

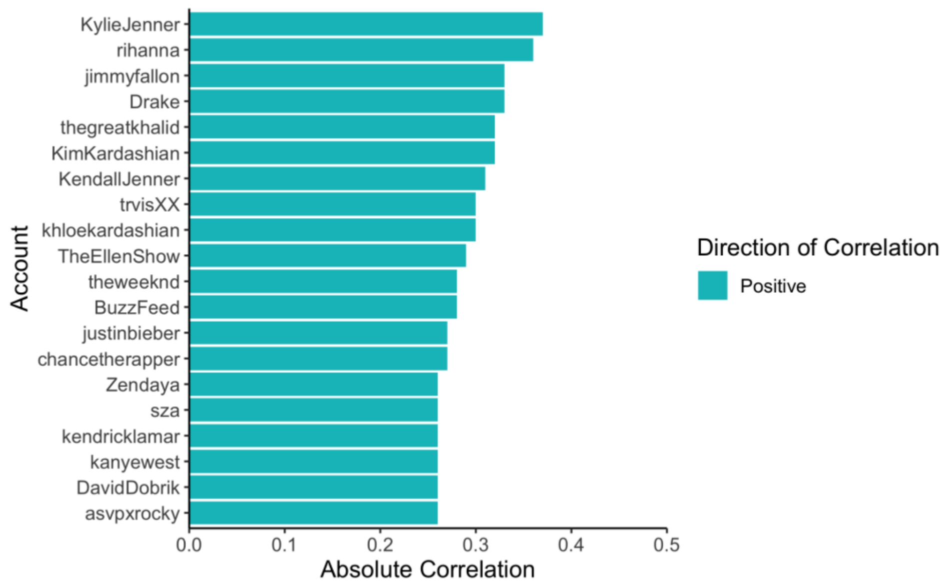


Figure 28

Twenty Highest Correlated Twitter Accounts with Neuroticism in Study 2

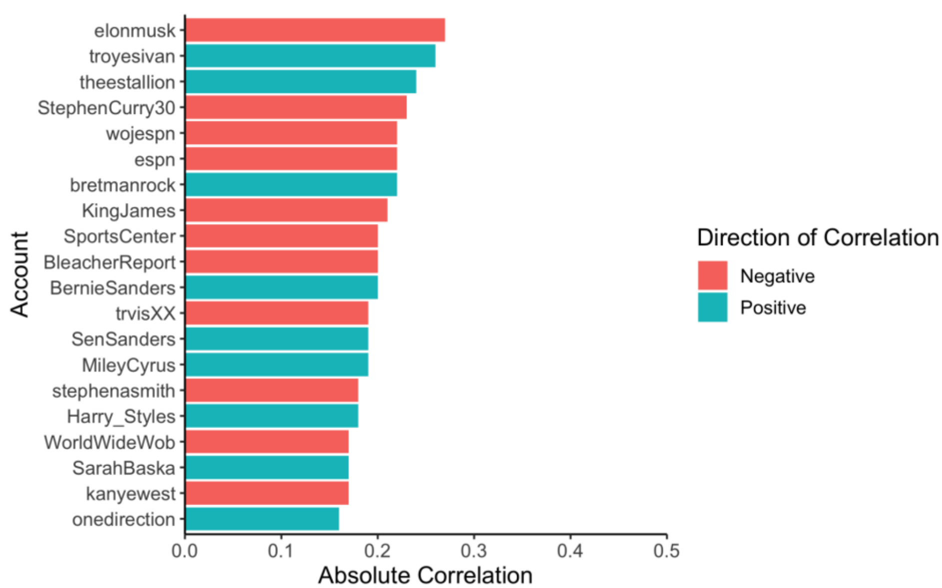


Figure 29

Twenty Highest Correlated Twitter Accounts with Honesty-Propriety in Study 2

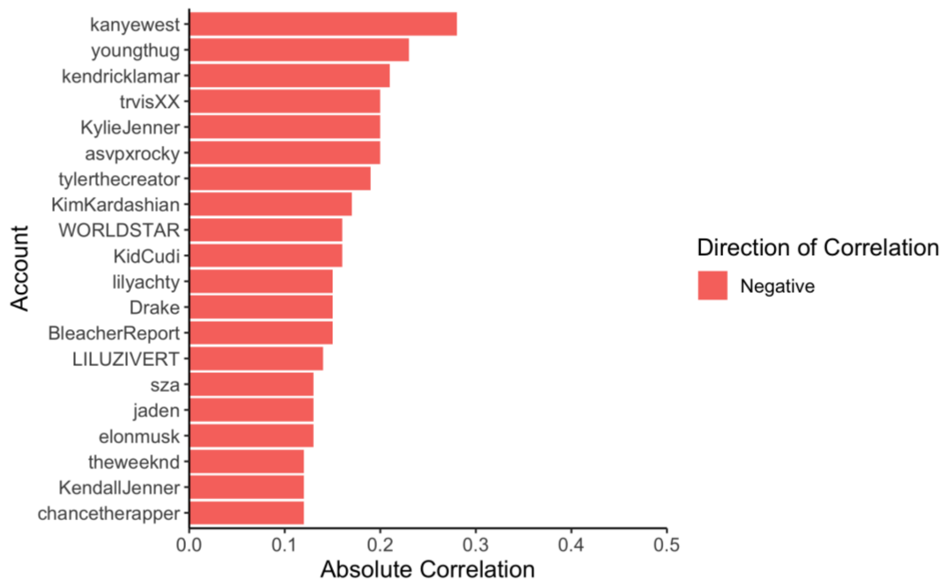


Figure 30

Twenty Highest Correlated Twitter Accounts with Agreeableness in Study 2

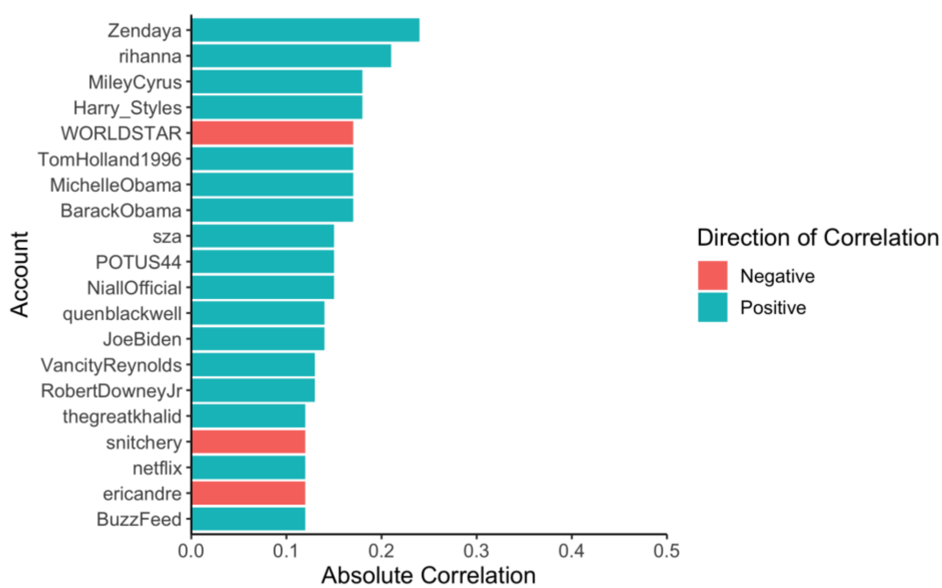


Figure 31

Twenty Highest Correlated Twitter Accounts with Openness in Study 2

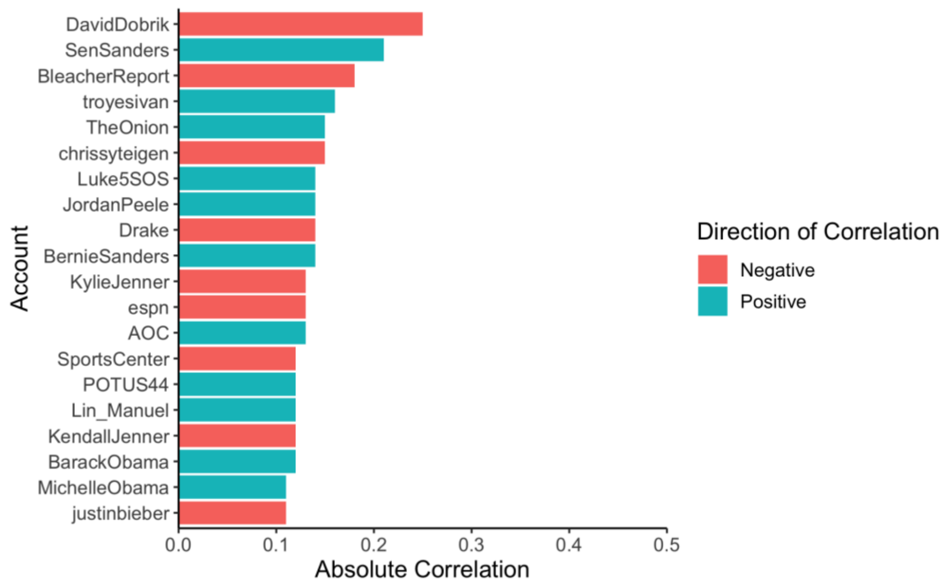
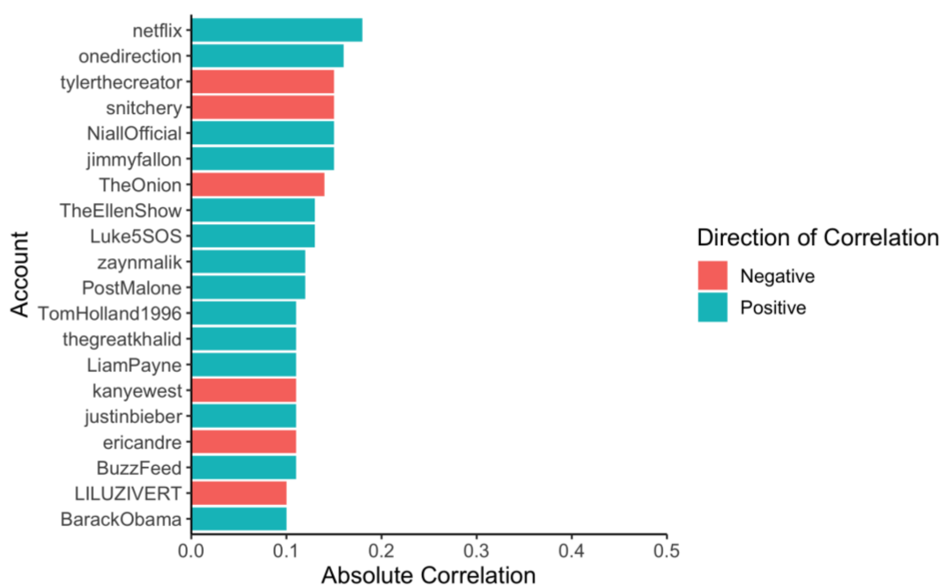


Figure 32

Twenty Highest Correlated Twitter Accounts with Conscientiousness in Study 2



To further summarize this data, a mean of absolute correlations was calculated for each trait across the 89 stimuli profiles (Table 16). This metric provides further information about the typical level of interest in an account by personality trait. The mean of absolute correlations also showed the strongest relationship with Extraversion (r = .17) and the weakest with Conscientiousness (r = .06). The same correlation summary was calculated for the demographic variables of gender (binary male = 1 and female = 2, non-binary was removed from these analyses due a low number of responses), age (numeric, range = 18-39 years), socioeconomic status (a numeric self-report on a ten-point scale), and political orientation (7-point scale from extremely liberal to extremely conservative) (Table 17). Gender had the highest mean absolute correlation (*r = .18*). Political orientation (*r = .15*), self-reported SES (*r = .13*), and age (*r = .12*) showed stronger relationships with account interest than all personality traits, except Extraversion. Study 2 data generally showed that personality traits have similar but slightly weaker relationships with account interest than demographic variables. Extraversion was an exception, demonstrating a relationship with account interest that was on par with gender.

Table 16

Mean of Absolute Correlations Between Twitter Account Interest Ratings and Personality Traits in Study 2

	Mean r
Extraversion	.17
Neuroticism	.10
Honesty	.07
Agreeableness	.07
Openness	.07
Conscientiousness	.06

Table 17

Mean of Absolute Correlations Between Twitter Account Interest Ratings and Demographic Variables in Study 2

	Mean r
Gender	.18
Political Orientation	.15
Self-report SES	.13
Age	.12

To examine how similar results were between the two populations sampled, I repeated the correlation calculations between individual account interest and personality traits (the same as seen in Table 15) with subgroups of UO and Prolific participants. I then correlated the results from each subsample with each other by trait (Table 18). Associations were positive for all traits indicating similar results between the two subsamples. Additionally, traits that correlated more strongly with accounts in the combined sample (Table 16) showed more consistency across subsamples.

Table 18

Correlations Between UO and Prolific Subsample Results

	r
Neuroticism	.68
Extraversion	.46
Openness	.42
Agreeableness	.41
Honesty	.40
Conscientiousness	.33

To explore whether these associations occur because some personality traits are associated with generally wanting to follow more accounts, as opposed to interest in the specific account content, I standardized account interest ratings within participants (ipsatized) and recalculated correlations with this data (Table 19 & 20). The mean of absolute correlations decreased for Extraversion, age, and socioeconomic status between regular and ipsatized data. To look at this relationship more specifically, I calculated correlations between an individual's mean account interest rating and their personality traits/demographic variables. While many of these correlations were close to zero, Extraversion showed a notable positive correlation (*r* = .35) indicating that those higher in Extraversion generally showed more interest in following accounts. Age also showed a positive relationship (*r* = .26) indicating that older participants generally showed more interest in following accounts. Notable negative relationships were demonstrated with self-report of SES (*r* = -.25) indicating that those high in SES generally showed less interest in following accounts.

Table 19

Mean of Absolute Correlations Between Ipsatized Twitter Account Interest Ratings and Personality in Study 2

	Combined Sample Mean r	Correlation with Individual Mean Interest Rating
Extraversion	.10	.35
Neuroticism	.11	.02
Honesty	.07	-.12
Agreeableness	.07	.12
Openness	.08	.00
Conscientiousness	.07	.05

Table 20

Mean of Absolute Correlations Between Ipsatized Twitter Account Interest Ratings and Demographic Variables in Study 2

	Combined Sample Mean r	Correlation with Individual Mean Interest Rating
Gender	.21	.08
Political Orientation	.16	.00
Self-report SES	.08	-.25
Age	.08	.26

Random Forests. To explore the strength of the relationship between personality traits and Twitter stimuli account interest ratings in the aggregate, I used the machine learning technique of random forests to predict personality traits from account interest ratings. The study 2 sample (N = 297) was split into a training and holdout sample consisting of 75% ($n_{training}$ = 222) and 25% ($n_{holdout}$ = 75) of the data respectively. Ten models were run, six predicting each of the personality traits individually, one predicting gender (binary male = 1 and female = 2), one predicting age (numeric), one predicting the self-report of socioeconomic status (on a ten-point scale), and one predicting political orientation (7-point scale from extremely liberal to extremely conservative) (Table 21 & Table 22). The highest predictive accuracy was seen for Political Orientation, Gender, and Extraversion. At least moderate predictive accuracy was seen across all other traits and demographic variables, though Openness notably demonstrated the least predictive accuracy in the test data. While these models demonstrate heterogeneity in predictive accuracy across personality traits and demographic variables, account interest ratings can generally predict personality traits with accuracy similar to some demographic variables.

Table 21

Random Forests Performance for Twitter Account Interest Ratings Predicting Personality Traits in Study 2

Trait	Train rmse	Train rsq	Train r	Test rmse	Test rsq	Test r	Lower 95% CI r	Upper 95% CI r
Extraversion	0.74	.16	.40	0.73	.25	.50	.31	.65
Agreeableness	0.57	.04	.20	0.45	.18	.42	.21	.59
Neuroticism	0.85	.11	.33	0.82	.13	.36	.14	.54
Honesty	0.55	.03	.17	0.55	.13	.36	.14	.54
Conscientiousness	0.69	.03	.17	0.70	.09	.30	.08	.49
Openness	0.62	.09	.30	0.61	.03	.17	-.06	.38

Table 22

Random Forests Performance for Twitter Account Interest Rating Predicting Demographic Variables in Study 2

Trait	Train rmse	Train rsq	Train r	Test rmse	Test rsq	Test r	Lower 95% CI r	Upper 95% CI r
Gender	0.35	.55	.74	0.52	.37	.61	.44	.74
Age	2.20	.19	.44	1.98	.13	.36	.14	.54
Self-report SES	1.61	.18	.42	1.58	.10	.32	.10	.51
Political Orientation	1.20	.49	.70	1.01	.61	.78	.67	.86

Aim 2a: Account Ratings Predicted from Profile Features

Multilevel Modeling. Do certain account features make users more or less interested in following those accounts? Multilevel modeling was used to examine the effect of individual Twitter profile features on account interest ratings. The model specification is as follows:

i = subject
j = account

$$\text{Account Interest Rating}_{i,j} = b_{0i} + b_{1j} + b_{2i} * \text{Profile Feature}_j + e_{ij}$$
$$b_{0i} = \gamma_{00} + U_{0i}$$
$$b_{1j} = \gamma_{10} + U_{1j}$$
$$b_{2i} = \gamma_{20} + U_{2j}$$

I ran a total of 48 models, one for each Twitter profile feature. The models were initially run with a random intercept for subject (U_{0i}), a random intercept for profile (U_{1j}), and a random slope for the effect of feature (U_{2j}). In cases where the model did not converge, I trimmed the random slope for subject (U_{2j}) and re-ran the model. For all models, the account interest rating metric was converted to a proportion of maximum possible score (POMP; Cohen et al., 1999). This transformation converts scores to a percent of the highest score available on the scale, with a theoretical range from 0 to 100, which allows for easier interpretation and communication of model results. To compare the effects of features with different measurement scales, Twitter profile features were z-scored. Features with meaningful units of measurement were converted back to their original scale for individual interpretation. Table 23 summarizes the features that were used as predictors.

Table 23

Summary of Twitter Stimuli Profile Features in Study 2

Feature Name	Feature Type	Description	Example
Count of Friends	Account Metadata	Number of accounts that a stimuli Twitter profile follows	Barack Obama has 130,685,919 followers
Count of Followers	Account Metadata	Number of accounts that follow a given stimuli Twitter account	Barack Obama follows 586,761 accounts
Average Word Count	Account Metadata	Average number of words per Tweet. Calculated using the account's last 100 Tweets	On average, Chrissy Teigen uses 18.3 words per Tweet.
Tweet Frequency	Account Metadata	Average number of Tweets an account creates per day. Calculated using the account's last 100 Tweets	On average Khloe Kardashian Tweets 6.82 times per day.
Ratio of Retweets to Tweets	Account Metadata	Number of retweets divided by number of non-retweets. Calculated using the account's last 100 Tweets	Kendrick Lamar has .59 retweets for every non-retweet
Perceived Age Appeal	Perceived Appeal	How much does this profile's content appeal to primarily young people (1) vs. primarily old people (5)? Numeric variable rated by human coders	Lil Uzi Vert appeals primarily to young people while Ellen DeGeneres appeals primarily to older people
Perceived Gender Appeal	Perceived Appeal	How much does this profile's content appeal to primarily men (1) vs. primarily women (5)? Numeric variable rated by human coders	ESPN appeals primarily to men while Kylie Jenner appeals primarily to women
Perceived SES Appeal	Perceived Appeal	How much does this profile's content appeal to primarily poor people (1) vs. primarily rich people (5)? Numeric variable rated by human coders	Bernie Sanders appeals primarily to poor people while Kim Kardashian appeals primarily to rich people
Perceived Political Appeal	Perceived Appeal	How much does this profile's content appeal to primarily liberal people (1) vs. primarily conservative people (5)? Numeric variable rated by human coders	Alexandria Ocasio-Cortez appeals primarily to liberal people while Elon Musk appeals primarily to rich people
Perceived Advertising	Perceived Account Characteristic	How much does this profile feel like advertising? Numeric variable rated by human coders on a scale from not at all (1) to extremely (5)	Kanye West's account does not feel like advertising, but Kendal Jenner's does

Table 23 Continued

Feature Name	Feature Type	Description	Example
Perceived Extraversion	Perceived Account Characteristic	Extraversion measured using the Ten-Item Personality Inventory Numeric variable rated by human coders on a scale from Disagree strongly (1) to Agree strongly (7)	Kendrick Lamar is perceived to be low in Extraversion while Miley Cyrus is perceived to be high in Extraversion
Perceived Agreeableness	Perceived Account Characteristic	Agreeableness measured using the Ten-Item Personality Inventory Numeric variable rated by human coders on a scale from Disagree strongly (1) to Agree strongly (7)	Kanye West is perceived to be low in Agreeableness while Barack Obama is perceived to be high in Agreeableness
Perceived Conscientiousness	Perceived Account Characteristic	Conscientiousness measured using the Ten-Item Personality Inventory Numeric variable rated by human coders on a scale from Disagree strongly (1) to Agree strongly (7)	Seth Rogen is perceived to be low in Conscientiousness while Michelle Obama is perceived to be high in Conscientiousness
Perceived Neuroticism	Perceived Account Characteristic	Neuroticism measured using the Ten-Item Personality Inventory Numeric variable rated by human coders on a scale from Disagree strongly (1) to Agree strongly (7)	Barack Obama is perceived to be low in Neuroticism while Kanye West is perceived to be high in Neuroticism
Perceived Openness	Perceived Account Characteristic	Openness measured using the Ten-Item Personality Inventory Numeric variable rated by human coders on a scale from Disagree strongly (1) to Agree strongly (7)	Sports Center is perceived to be low in Openness while Miley Cyrus is perceived to be high in Openness
Perceived Positive Affect	Perceived Account Characteristic	How positive is the content of this profile? Numeric variable rated by human coders on a scale from not at all (1) to extremely (5)	Travis Scott’s profile is perceived to be low in positive affect while Michelle Obama’s profile is perceived to be high in positive affect
Perceived Negative Affect	Perceived Account Characteristic	How negative is the content of this profile? Numeric variable rated by human coders on a scale from not at all (1) to extremely (5)	Harry Styles’s profile is perceived to be low in negative affect while Bernie Sanders’s profiles is perceived to be high in negative affect
Human Account Owner	Perceived Account Characteristic	Binary variable indicating that the Twitter account represents a single person rather than a brand or group of people	Harry Styles is a human, ESPN is not a human

Table 23 Continued

Feature Name	Feature Type	Description	Example
PCA Celebrities	PCA Dimension	Component loadings for the latent category related to young pop musicians and actors	Accounts that load highly include on this dimension include Harry Styles and Zendaya
PCA Sports	PCA Dimension	Component loadings for the latent category related to sports news and athletes	Accounts that load highly include on this dimension include Sports Center and LeBron James
PCA Rap/R&B	PCA Dimension	Component loadings for the latent category related to popular rap or R&B musicians	Accounts that load highly include on this dimension include Kendrick Lamar and Drake
PCA Mainstream Influencers	PCA Dimension	Component loadings for the latent category related to TV reality stars and internet personalities	Accounts that load highly include on this dimension include Kim Kardashian and Chrissy Teigen
PCA Liberal Politicians	PCA Dimension	Component loadings for the latent category related to left-leaning political accounts	Accounts that load highly include on this dimension include Kamala Harris and Bernie Sanders
PCA Traditional Media Comedy	PCA Dimension	Component loadings for the latent category related to comedic actors and online humor accounts	Accounts that load highly include on this dimension include Seth Rogen and The Onion
PCA Social Media Personalities	PCA Dimension	Component loadings for the latent category related to humorous online personalities	Accounts that load highly include on this dimension include Noel Miller and Caucasian James
NRC Anger	Linguistic	Average anger score for an account’s last 100 Tweets Calculated using the NRC dictionary	Words that score as high in sadness include agitated, infuriated, fury
NRC Anticipation	Linguistic	Average anticipation score for an account’s last 100 Tweets Calculated using the NRC dictionary	Words that score as high in sadness include aroused, expectation, unveil
NRC Disgust	Linguistic	Average disgust score for an account’s last 100 Tweets Calculated using the NRC dictionary	Words that score as high in sadness include repulsed, sickened, revolting

Table 23 Continued

Feature Name	Feature Type	Description	Example
NRC Fear	Linguistic	Average fear score for an account's last 100 Tweets Calculated using the NRC dictionary	Words that score as high in sadness include terrify, phobic, anxious
NRS Joy	Linguistic	Average joy score for an account's last 100 Tweets Calculated using the NRC dictionary	Words that score as high in sadness include happy, elation, uplifting
NRC Sadness	Linguistic	Average sadness score for an account's last 100 Tweets Calculated using the NRC dictionary	Words that score as high in sadness include dreary, depressed, bereavement
NRC Surprise	Linguistic	Average surprise score for an account's last 100 Tweets Calculated using the NRC dictionary	Words that score as high in surprise include amused, mindblown, surprised
NRC Trust	Linguistic	Average trust score for an account's last 100 Tweets Calculated using the NRC dictionary	Words that score as high in trust include valued, trusted, vetted
LIWC Affiliation	Linguistic	Percent of words in an account's last 100 Tweets that match the affiliation category Calculated using the LIWC dictionary	Includes words like we, our, us, help
LIWC Achievement	Linguistic	Percent of words in an account's last 100 Tweets that match the achievement category Calculated using the LIWC dictionary	Includes words like work, better, best, working
LIWC Power	Linguistic	Percent of words in an account's last 100 Tweets that match the power category Calculated using the LIWC dictionary	Includes words like own, order, allow, power
LIWC Prosocial	Linguistic	Percent of words in an account's last 100 Tweets that match the prosocial category Calculated using the LIWC dictionary	Includes words like care, help, thank
LIWC Politeness	Linguistic	Percent of words in an account's last 100 Tweets that match the politeness category Calculated using the LIWC dictionary	Includes words like please, thanks, good morning

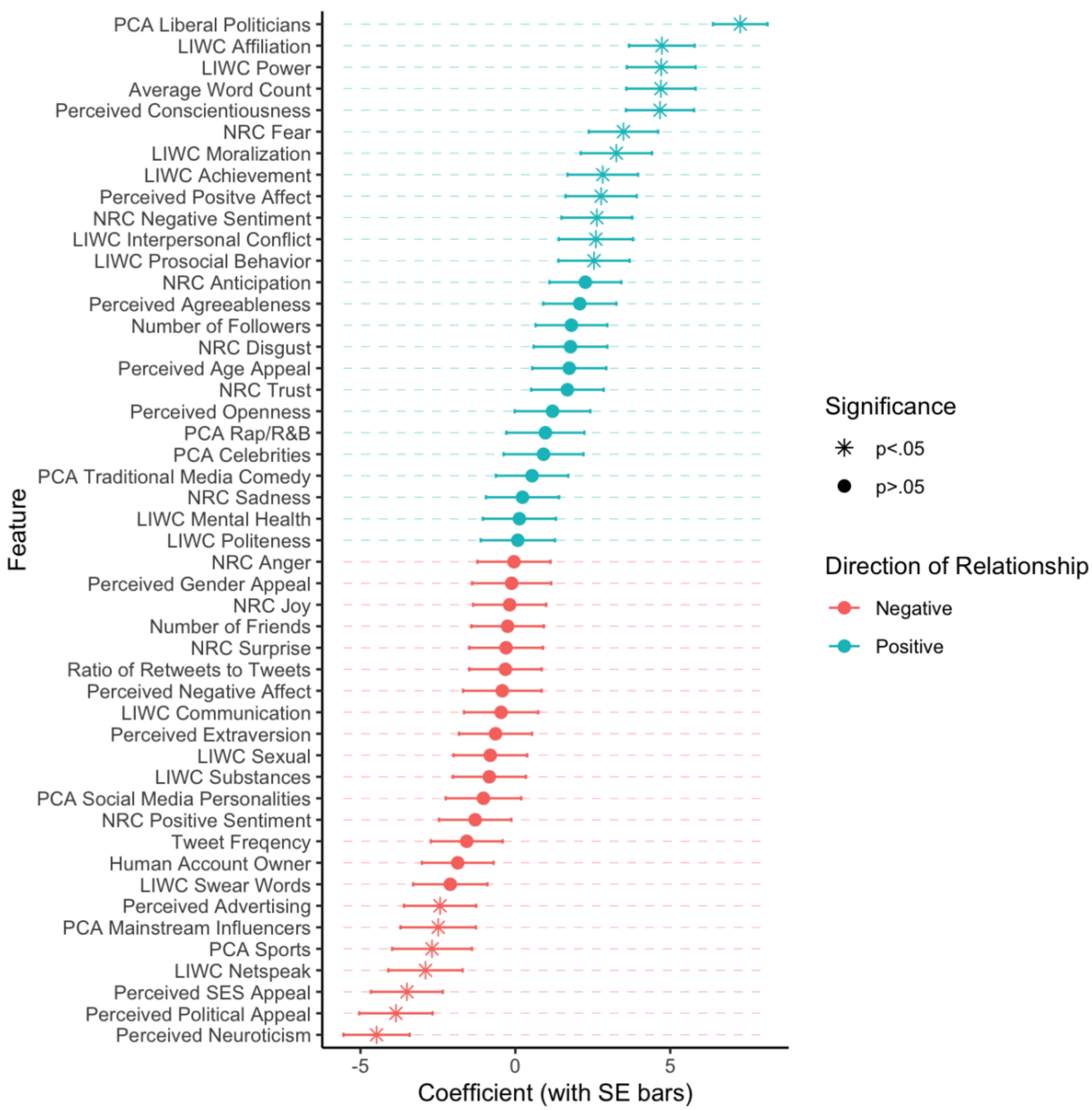
Table 23 Continued

Feature Name	Feature Type	Description	Example
LIWC Interpersonal Conflict	Linguistic	Percent of words in an account's last 100 Tweets that match the interpersonal conflict category Calculated using the LIWC dictionary	Includes words like fight, kill, killed, attack
LIWC Moralization	Linguistic	Percent of words in an account's last 100 Tweets that match the moralization category Calculated using the LIWC dictionary	Includes words like wrong, honor, deserve, judge
LIWC Communication	Linguistic	Percent of words in an account's last 100 Tweets that match the communication category Calculated using the LIWC dictionary	Includes words like say, tell, thank
LIWC Mental Health	Linguistic	Percent of words in an account's last 100 Tweets that match the mental health category Calculated using the LIWC dictionary	Includes words like mental health, depressed, trauma
LIWC Swear Words	Linguistic	Percent of words in an account's last 100 Tweets that match the swear words category Calculated using the LIWC dictionary	Includes words like shit, fuck, damn
LIWC Substances	Linguistic	Percent of words in an account's last 100 Tweets that match the substances category Calculated using the LIWC dictionary	Includes words like beer, wine, drunk, cigar
LIWC Sexual	Linguistic	Percent of words in an account's last 100 Tweets that match the sexual category Calculated using the LIWC dictionary	Includes words like sex, gay, pregnant
LIWC Netspeak	Linguistic	Percent of words in an account's last 100 Tweets that match the netspeak category Calculated using the LIWC dictionary	Includes words like :), u, lol, haha

The effect of Twitter profile features on account following interest (and standard error) is displayed in Figure 33. The PCA Liberal Politicians topic had the largest positive influence on account ratings. The more strongly an account loaded on the Liberal Politicians PCA topics the more participants were interested in following. For every standard deviation increase in liberal political content relevance, interest in following increased by 7.26 POMP units. Perceived Neuroticism had the largest negative influence on account ratings. Participants were less interested in following accounts viewed as higher in Neuroticism. For every standard deviation increase in perceived Neuroticism, interest in following decreased by 4.48 POMP units.

Figure 33

Effect of Z-scored Twitter Profile Features on Study 2 Account Interest Ratings



Aim 2b: Personality as a Moderator for the Relationship Between Twitter Profile Interest Ratings and Twitter Profile Features

Moderation Analysis. Do account features have different effects for different participants, as a function of the participant's personality? To examine if participant personality traits moderate the effect of Twitter profile features on Twitter account interest ratings, moderation analysis was incorporated into the multilevel models. The model specification is as follows:

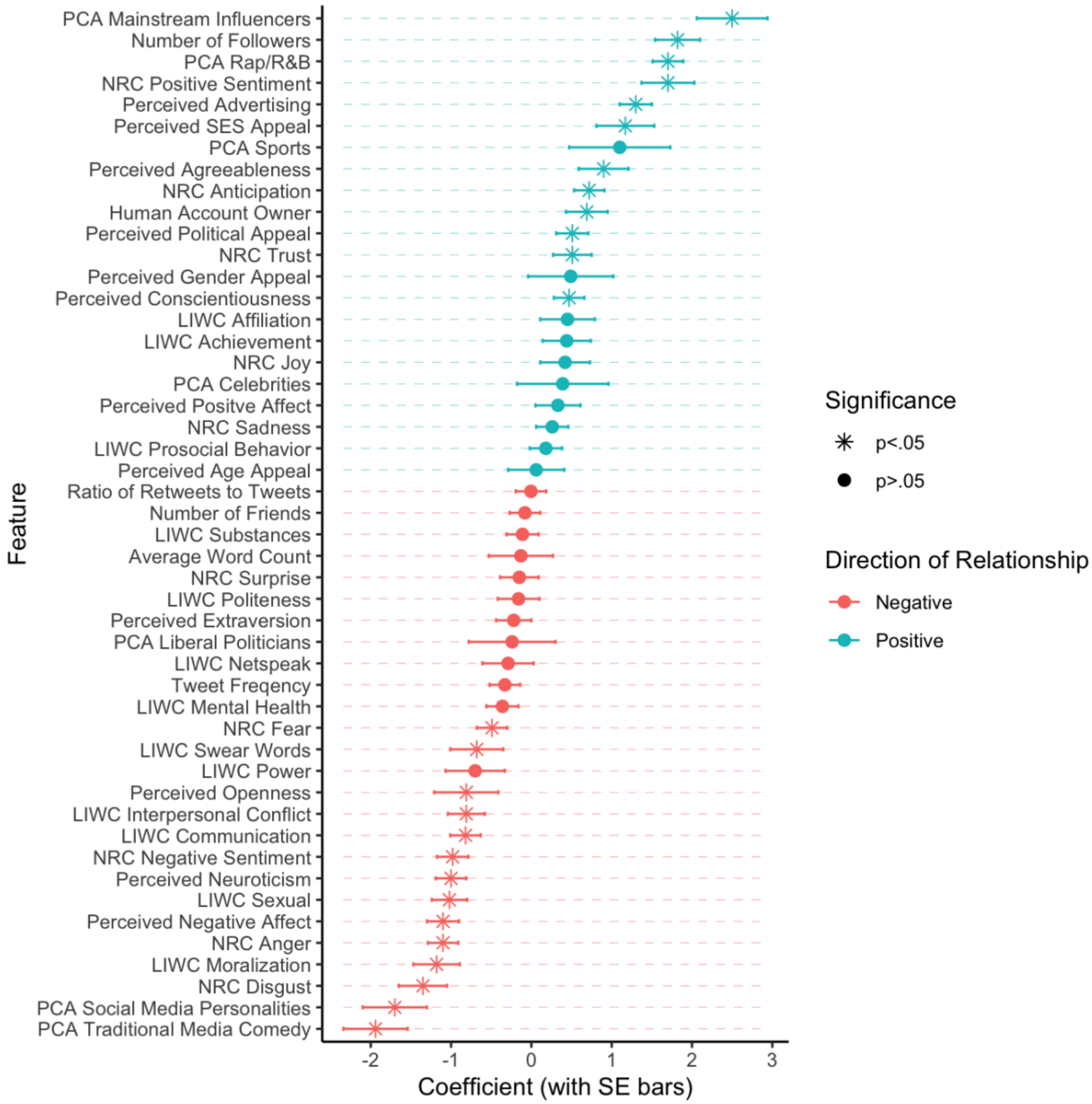
i = subject
j = account

$$\text{Account Interest Rating}_{i,j} = b_{0i} + b_{1j} + b_{2i} * \text{Profile Feature}_j + e_{ij}$$
$$b_{0i} = \gamma_{00} + \gamma_{01} * \text{Trait}_i + U_{0i}$$
$$b_{1j} = \gamma_{10} + U_{1j}$$
$$b_{2i} = \gamma_{20} + \gamma_{21} * \text{Trait}_i + U_{2j}$$

I ran a total of 288 models, one for each combination of 6 personality traits and 48 Twitter profile features. I used the same trimming procedure for models that did not converge as the original multilevel models. The account interest rating metric was converted to POMP scores. Similarly, to compare model results across features with different measurement scales, Twitter profile features and personality traits were z-scored. Features with meaningful units of measurement were converted back to their original scale for individual interpretation. The full set of 288 interaction coefficients are presented in figures 34, 37, 40, 43, 46, and 49. Positive coefficients indicate that the effect of the feature on interest is relatively more positive for people who score high on the trait (and relatively more negative for people who score low). Negative coefficients indicate the reverse. To aid in interpretation and illustrate what these interaction effects look like when added to the main effects presented in the previous section, I have plotted and elaborated on the interpretation for the most positive and most negative interaction effect for each trait.

Figure 34

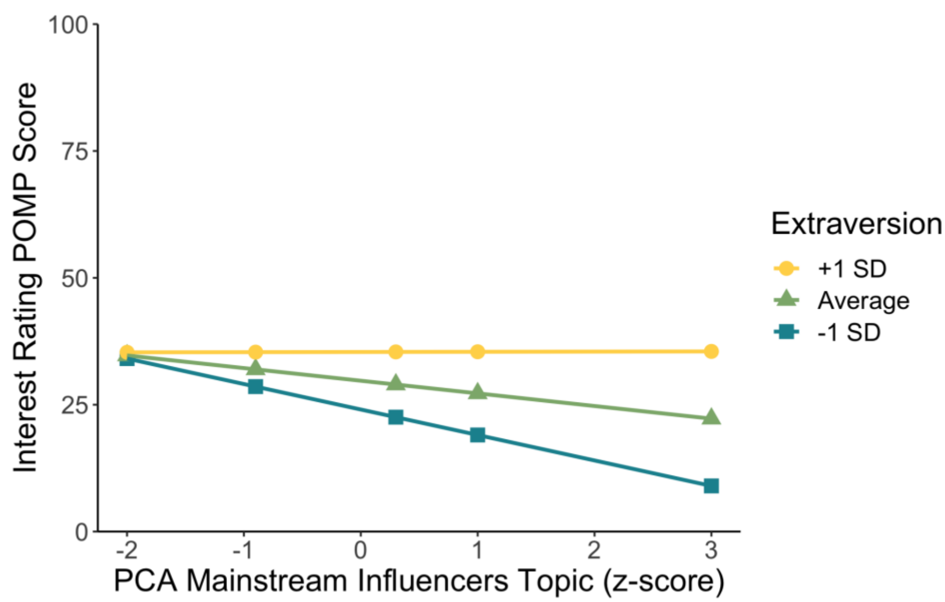
Moderating Effect of Z-scored Extraversion on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



The effect of PCA Mainstream Influencers topic feature was most positively moderated by Extraversion (Figure 35). The main effect of PCA Mainstream Influencers was negative, for every standard deviation increase in reality star content relevance, interest in following decreased by 2.49 POMP units less interest in following. For someone 1 standard deviation above the mean in Extraversion, the simple effect of reality star content was 0.01. For someone 1 standard deviation below the mean in Extraversion, the simple effect of reality star content was -5.99. In other words, lower levels of Extraversion were associated with greater sensitivity to reality star-related content.

Figure 35

Interaction Plot for Extraversion and the Effect of PCA Mainstream Influencers



The effect of PCA Traditional Media Comedy topic feature was most negatively moderated by Extraversion (Figure 36). The main effect of PCA Traditional Media Comedy was positive, for every standard deviation increase in traditional media comedy content relevance, interest in following increased by 0.54 POMP units greater interest in following. For someone 1 standard deviation above the mean in Extraversion, the simple effect of traditional media comedy content was -1.40. For someone 1 standard deviation below the mean in Extraversion, the simple effect of traditional media comedy content was 2.48. In other words, lower levels of Extraversion were associated with greater sensitivity to reality traditional media comedy-related content.

Figure 36

Interaction Plot for Extraversion and the Effect of PCA Traditional Media Comedy

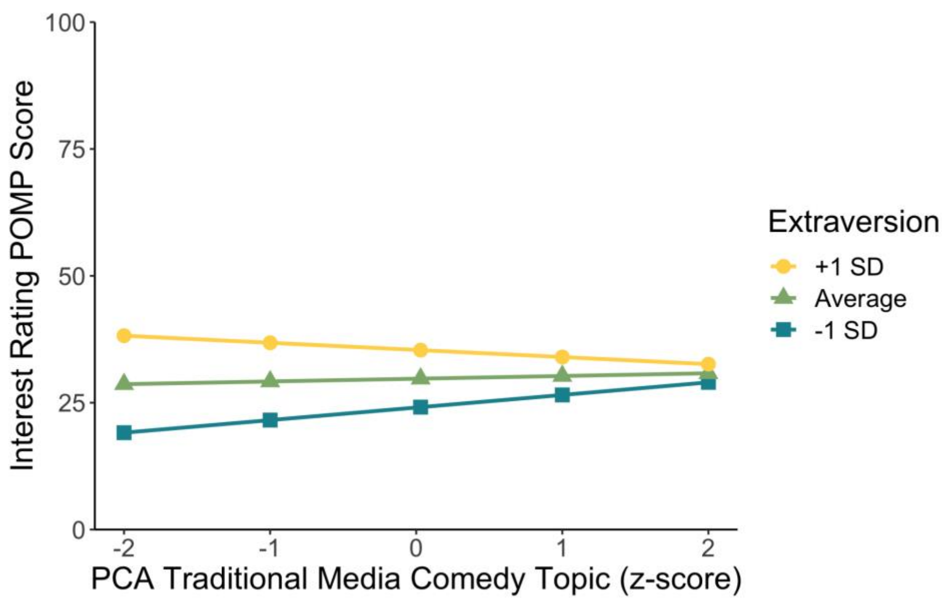
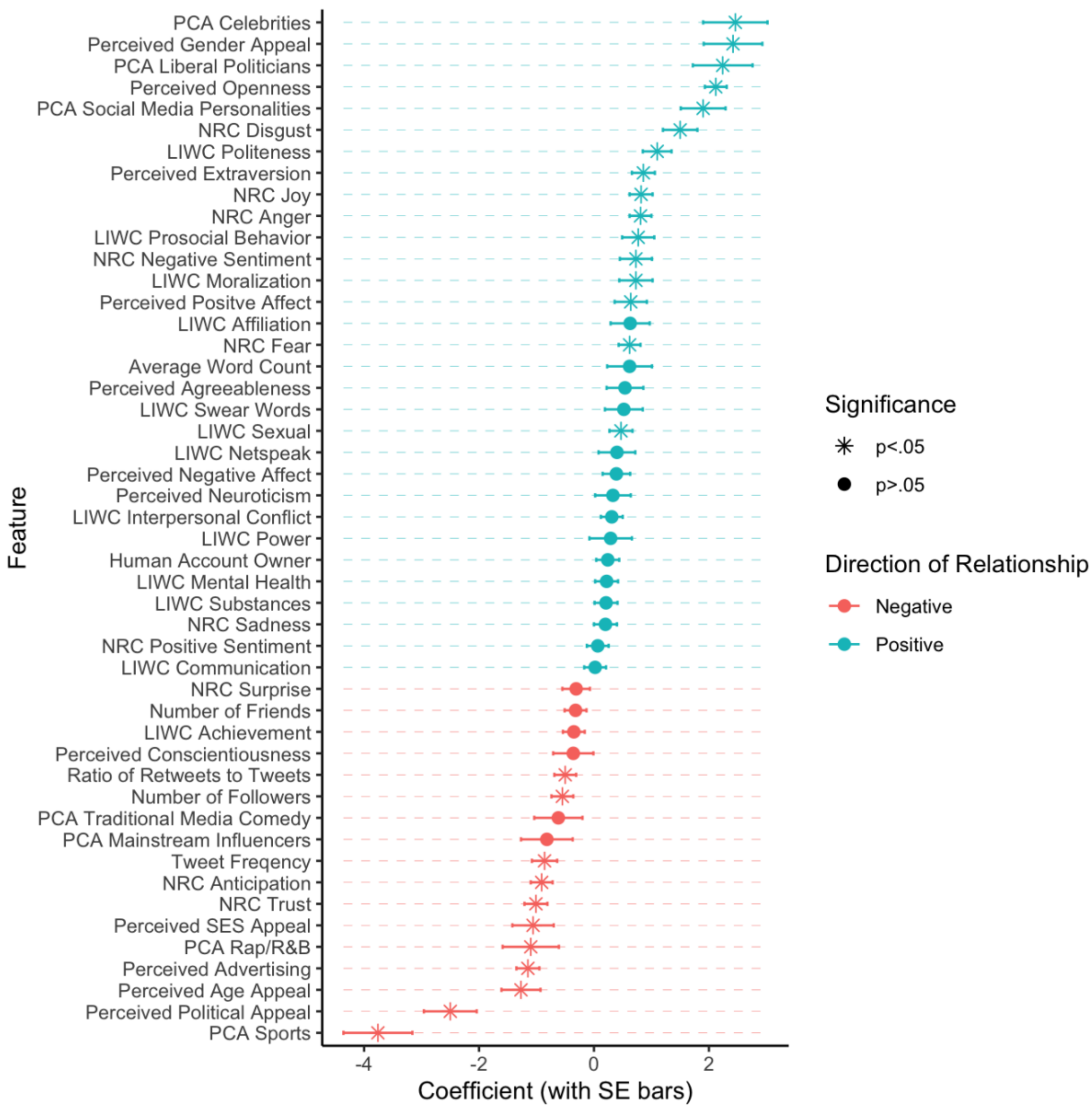


Figure 37

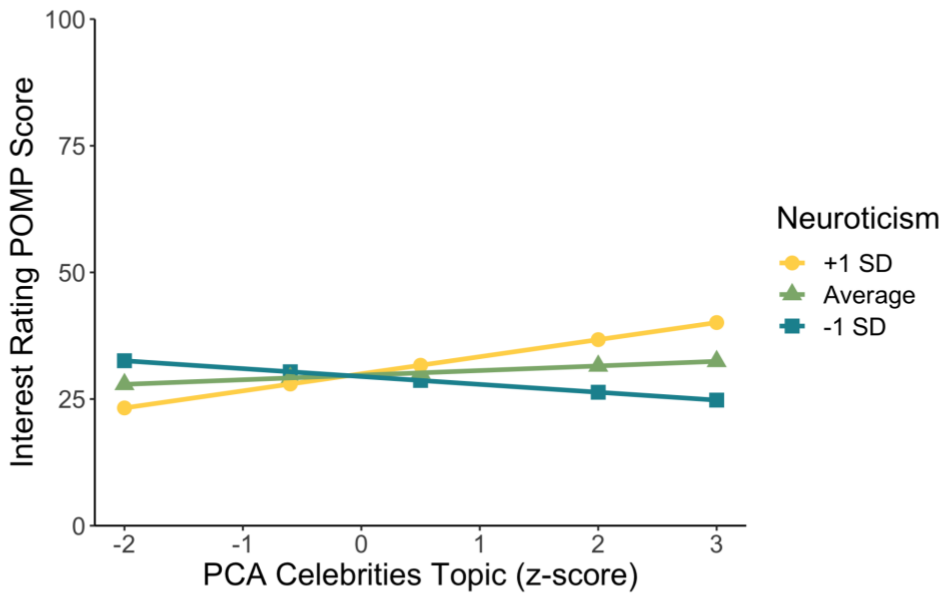
Moderating Effect of Z-scored Neuroticism on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



The effect of PCA Celebrities topic feature was most positively moderated by Neuroticism (Figure 38). The main effect of PCA Celebrities was positive, indicating that at an average level of Neuroticism, every standard deviation increase in celebrity related content relevance resulted in 0.91 POMP units greater interest in following. For someone 1 standard deviation above the mean in Neuroticism, the simple effect of celebrity content was 3.37. For someone 1 standard deviation below the mean in Neuroticism the main effect of celebrity content was -1.55. In other words, higher levels of Neuroticism were associated with greater sensitivity to celebrity-related content.

Figure 38

Interaction Plot for Neuroticism and the Effect of PCA Celebrities



The effect of PCA Sports topic feature was most negatively moderated by Neuroticism (Figure 39). The main effect of PCA Sports was negative, indicating that at an average level of Neuroticism, every standard deviation increase in sports related content relevance resulted in 2.69 POMP units less interest in following. For someone 1 standard deviation above the mean in Neuroticism, the simple effect of sports content was -6.72. For someone 1 standard deviation below the mean in Neuroticism, the simple effect of sports content was 1.07. In other words, higher levels of Neuroticism were associated with greater sensitivity to sports-related content.

Figure 39

Interaction Plot for Neuroticism and the Effect of PCA Sports

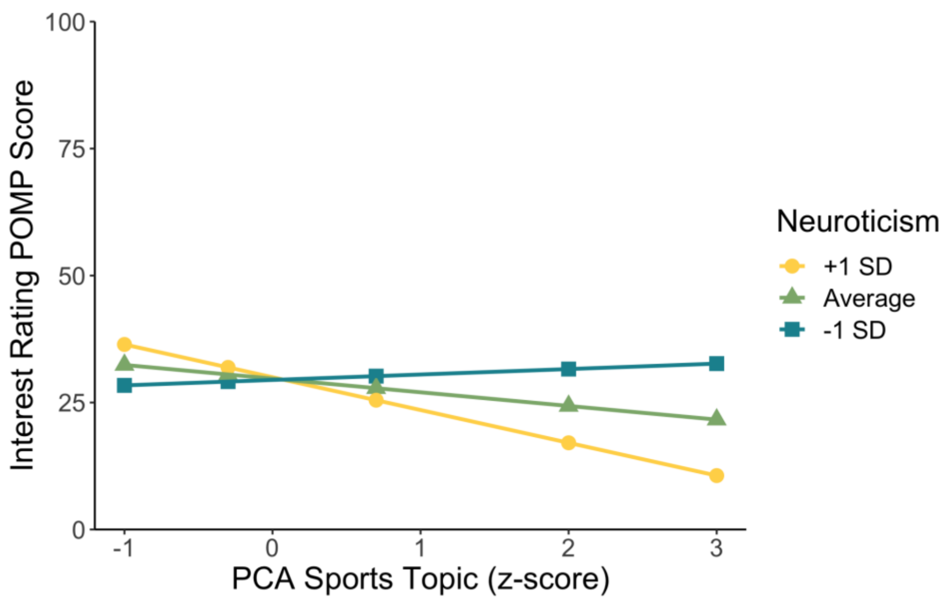
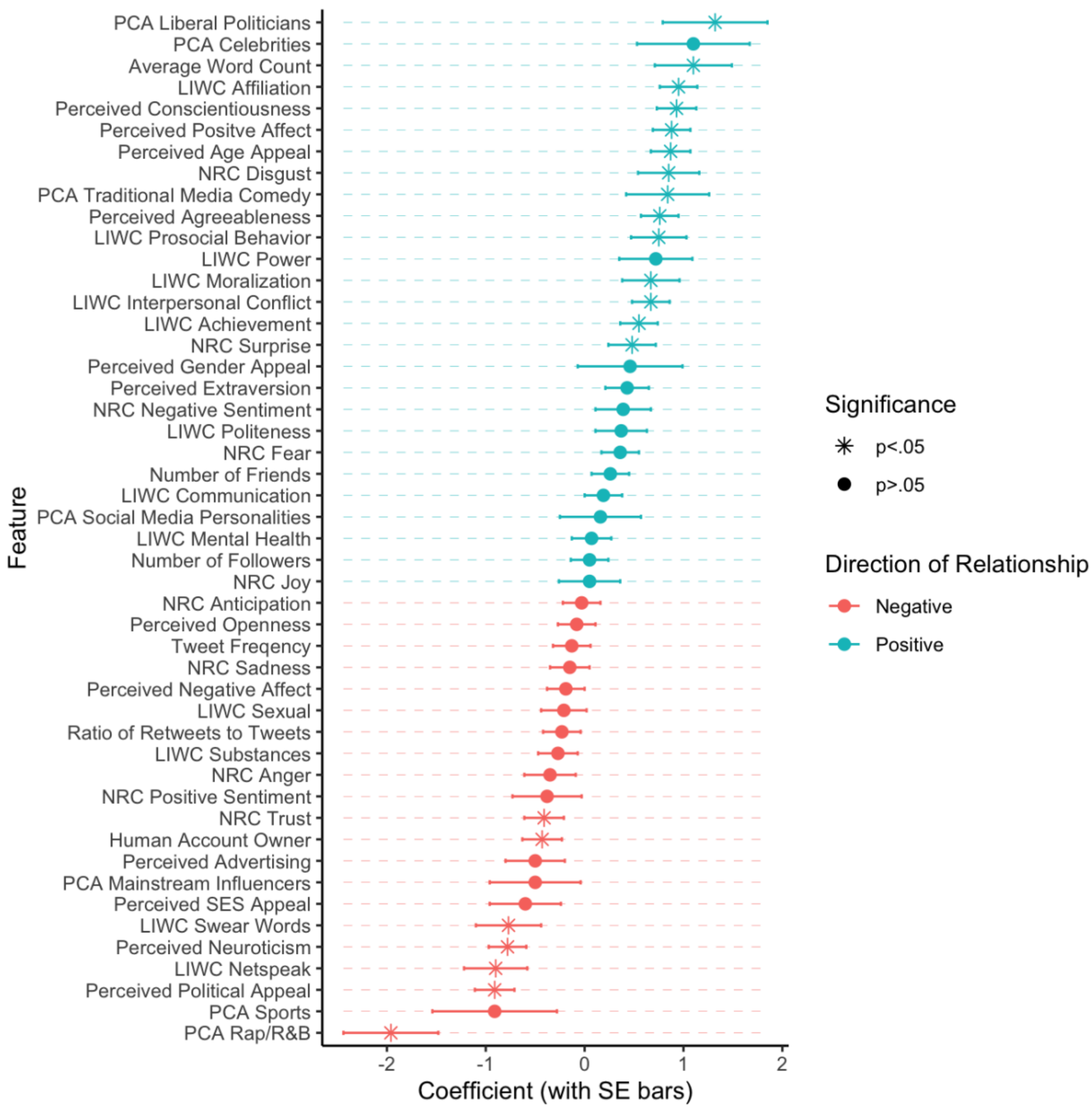


Figure 40

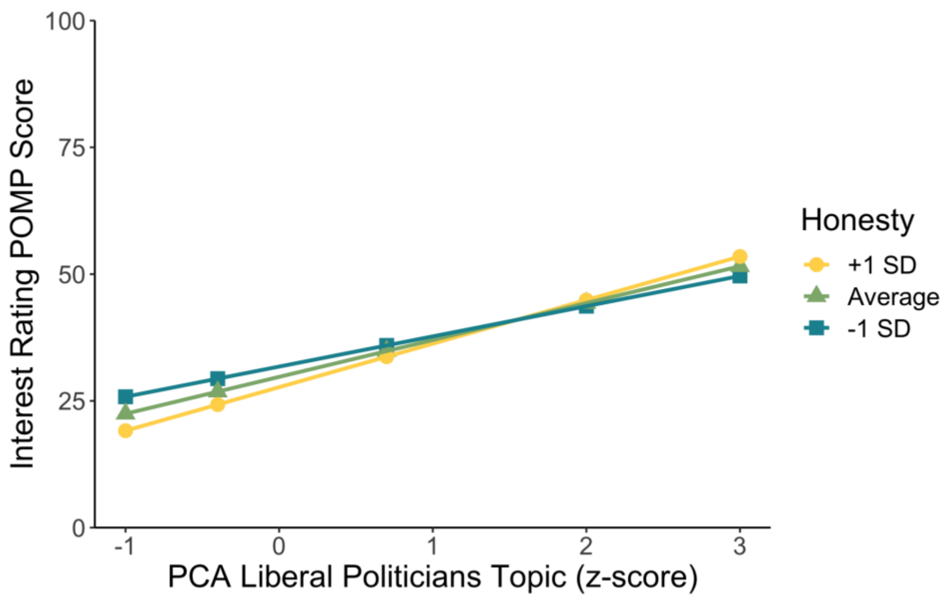
Moderating Effect of Z-scored Honesty-Propriety on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



The effect of PCA Liberal Politicians topic feature was most positively moderated by Honesty-Propriety (Figure 41) The main effect of PCA Liberal Politicians was positive, indicating that at an average level of Honesty-Propriety, every standard deviation increase in liberal political related content relevance resulted in 7.26 POMP units greater interest in following. For someone 1 standard deviation above the mean in Honesty-Propriety, the simple effect of liberal political content was 8.58. For someone 1 standard deviation below the mean in Honesty-Propriety, the simple effect of liberal political content was 5.94. In other words, higher levels of Honesty-Propriety were associated with greater sensitivity to liberal political-related content.

Figure 41

Interaction Plot for Honesty-Propriety and the Effect of PCA Liberal Politicians



The effect of PCA Rap/R&B topic feature was most negatively moderated by Honesty-Propriety (Figure 42). The main effect of PCA Rap/R&B was positive, indicating that at an average level of Honesty-Propriety, every standard deviation increase in rap/R&B related content relevance resulted in 0.97 POMP units greater interest in following. For someone 1 standard deviation above the mean in Honesty-Propriety, the simple effect of rap/R&B content was -0.99. For someone 1 standard deviation below the mean in Honesty-Propriety, the simple effect of rap/R&B content was 2.93. In other words, lower levels of Honesty-Propriety were associated with greater sensitivity to rap/R&B-related content.

Figure 42

Interaction Plot for Honesty-Propriety and the Effect of PCA Rappers

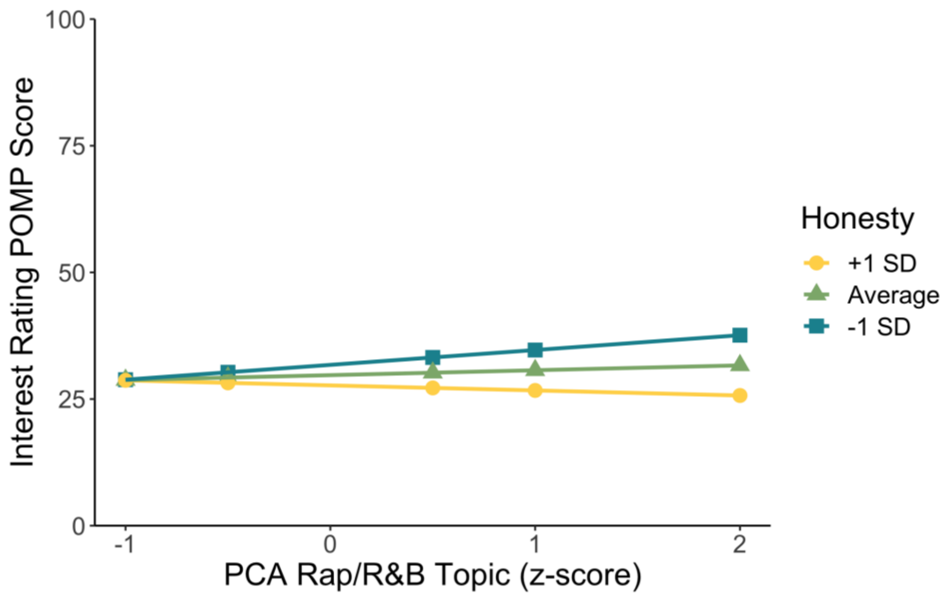
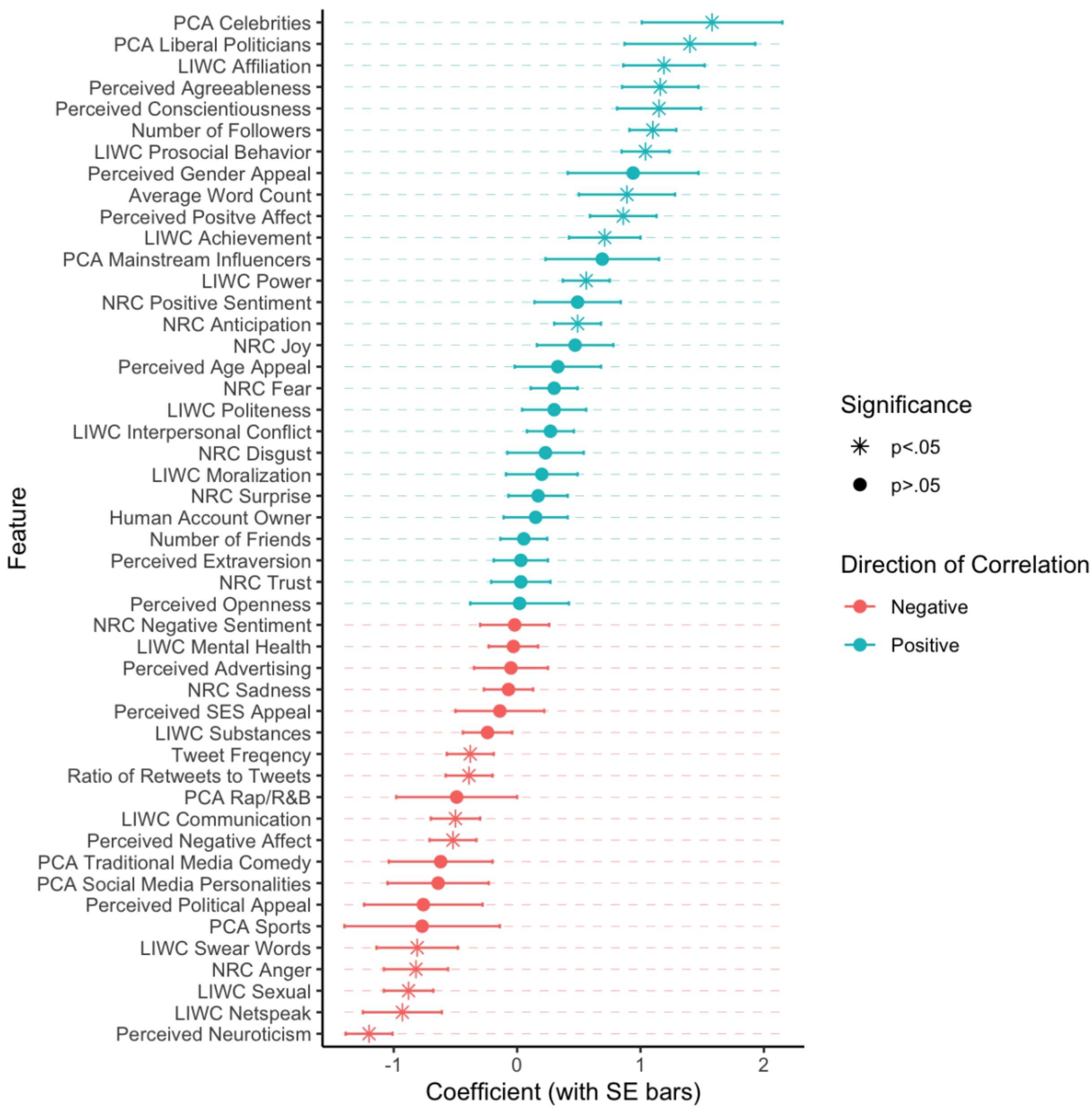


Figure 43

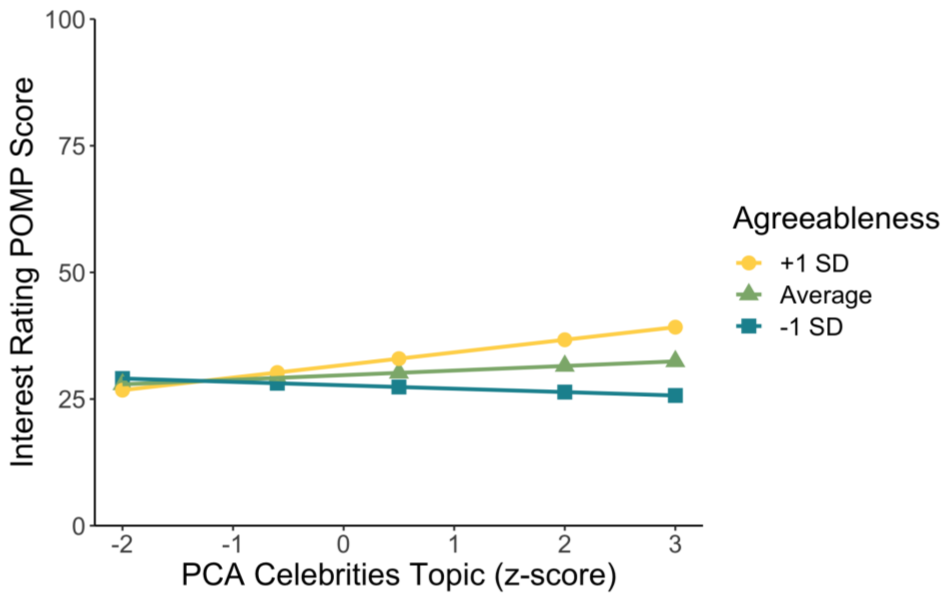
Moderating Effect of Z-scored Agreeableness on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



The effect of PCA Celebrities topic feature was most positively moderated by Agreeableness (Figure 44). The main effect of PCA Celebrities was positive, indicating that at an average level of Agreeableness, every standard deviation increase in celebrity related content relevance resulted in 0.91 POMP units greater interest in following. For someone 1 standard deviation above the mean in Agreeableness, the simple effect of celebrity content was 2.49. For someone 1 standard deviation below the mean in Agreeableness, the simple effect of celebrity content was -0.67. In other words, higher levels of Agreeableness were associated with greater sensitivity to celebrity-related content.

Figure 44

Interaction Plot for Agreeableness and the Effect of PCA Celebrities topic



The effect of perceived Neuroticism was most negatively moderated by Agreeableness (Figure 45). The main effect of perceived Neuroticism was negative, indicating that at an average level of Agreeableness, every standard deviation increase in perceived Neuroticism resulted in -4.48 POMP units less interest in following. For someone 1 standard deviation above the mean in Agreeableness, the simple effect of perceived Neuroticism was -5.68. For someone 1 standard deviation below the mean in Agreeableness, the simple effect of perceived Neuroticism was -3.28. In other words, higher levels of Agreeableness were associated with greater sensitivity to perceived Neuroticism in Twitter profiles.

Figure 45

Interaction Plot for Agreeableness and the Effect of Perceived Neuroticism

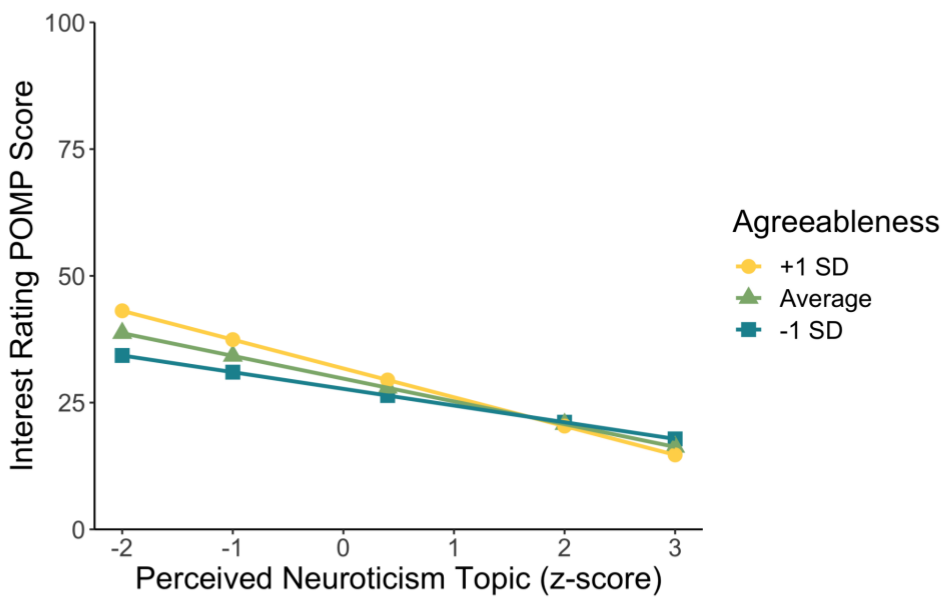
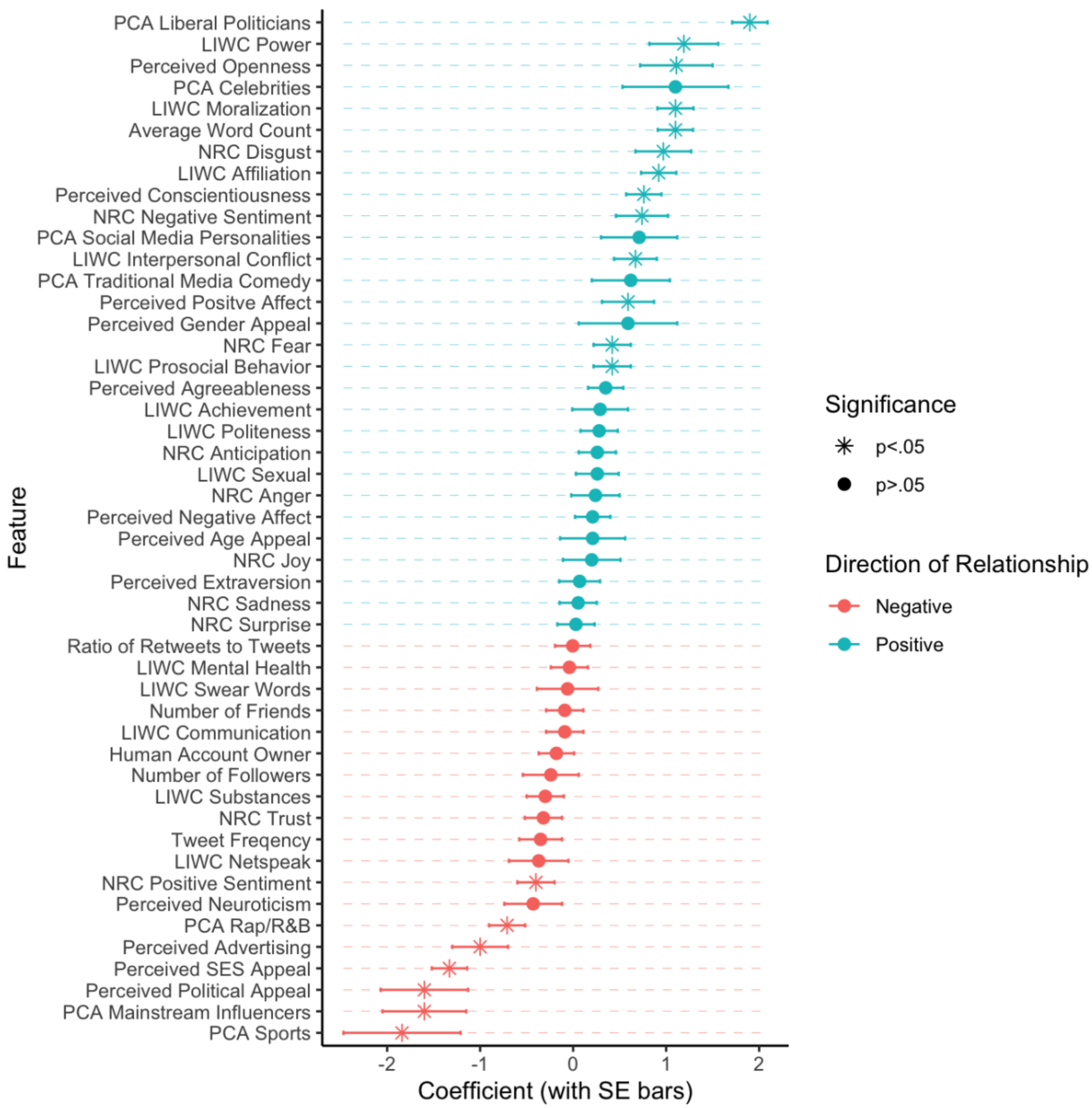


Figure 46

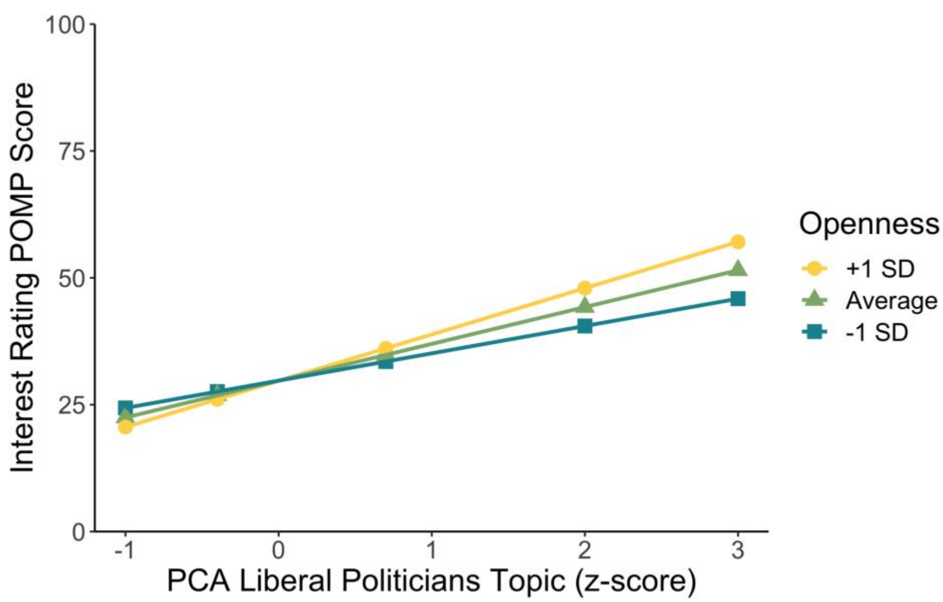
Moderating Effect of Z-scored Openness on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



The effect of PCA Liberal Politicians topic feature was most positively moderated by Openness (Figure 47). The main effect of PCA Liberal Politicians was positive, indicating that at an average level of Openness, every standard deviation increase in liberal political content relevance resulted in 7.26 POMP units greater interest in following. For someone 1 standard deviation above the mean in Openness, the simple effect of liberal political content was 9.16. For someone 1 standard deviation below the mean in Openness, the simple effect of political related content was 5.36. In other words, higher levels of Openness were associated with greater sensitivity to liberal political-related content.

Figure 47

Interaction Plot for Openness and the Effect of PCA Liberal Politicians



The effect of PCA Sports topic feature was most negatively moderated by Openness (Figure 48). The main effect of PCA Sports was negative, indicating that at an average level of Openness, every standard deviation increase in sports related content relevance resulted in 2.69 POMP units less interest in following. For someone 1 standard deviation above the mean in Openness, the simple effect of sports content was -4.53. For someone 1 standard deviation below the mean in Openness, the simple effect of sports content was -0.85. In other words, higher levels of Openness were associated with greater sensitivity to sports-related content.

Figure 48

Interaction Plot for Openness and the Effect of PCA Sports Topic

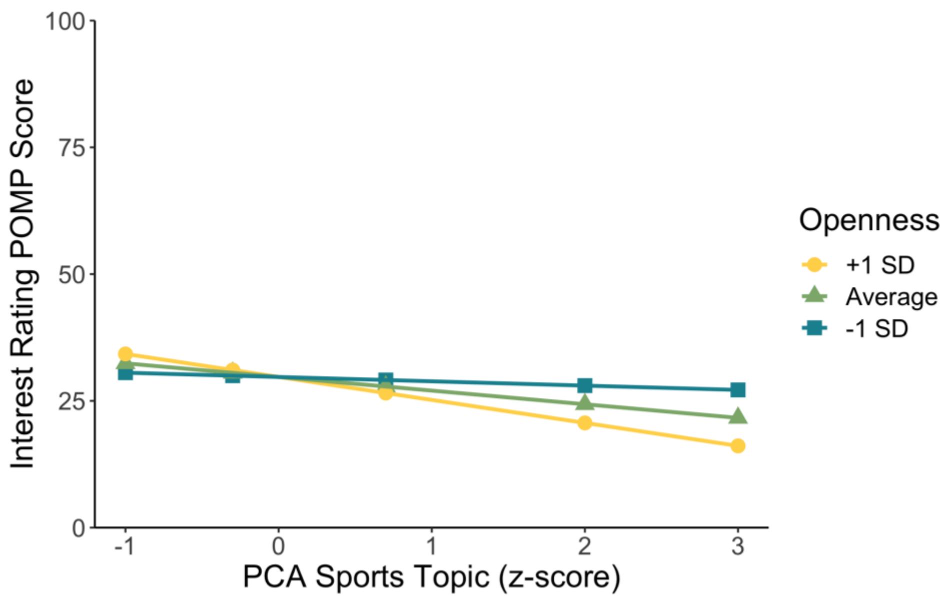
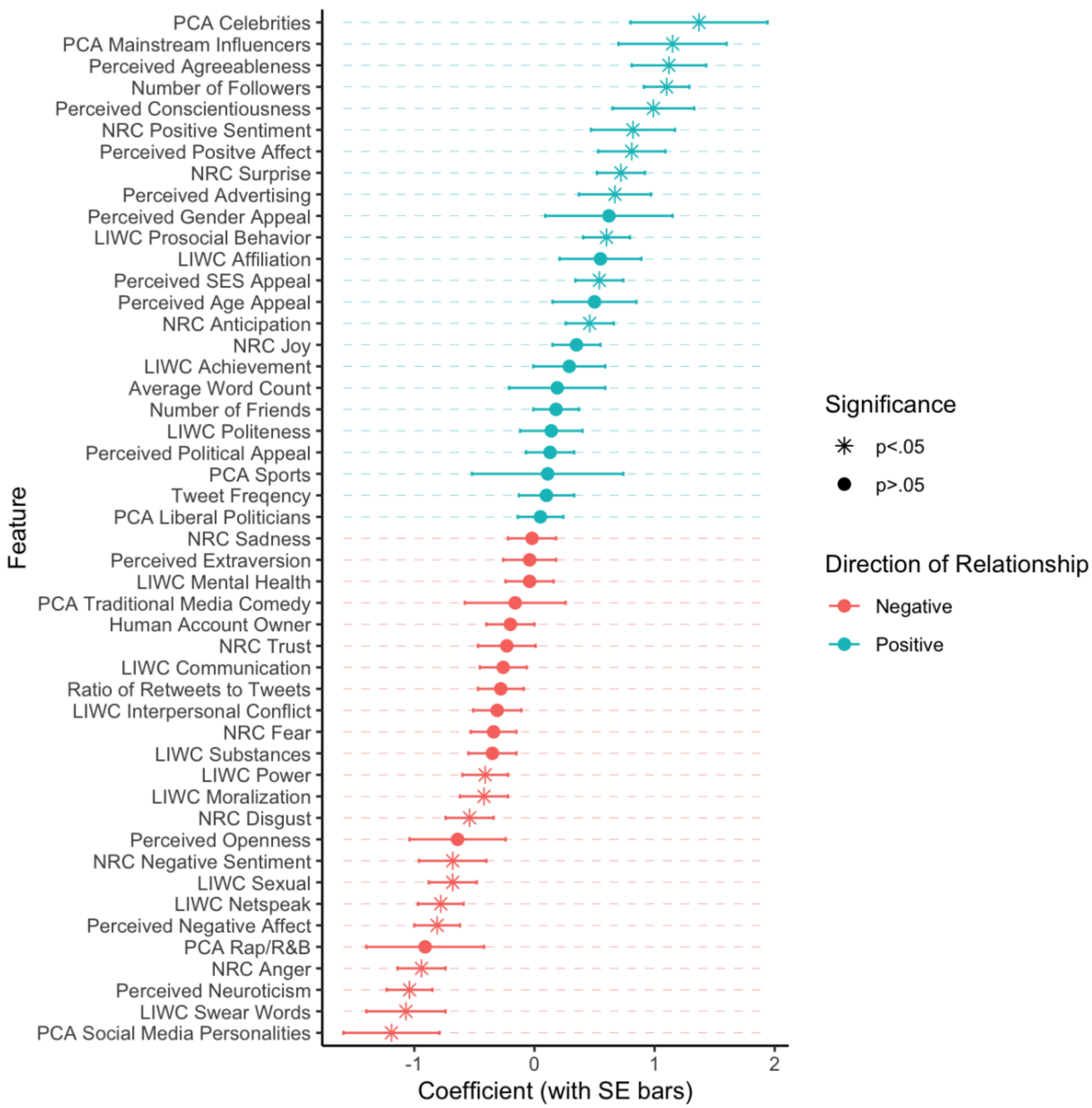


Figure 49

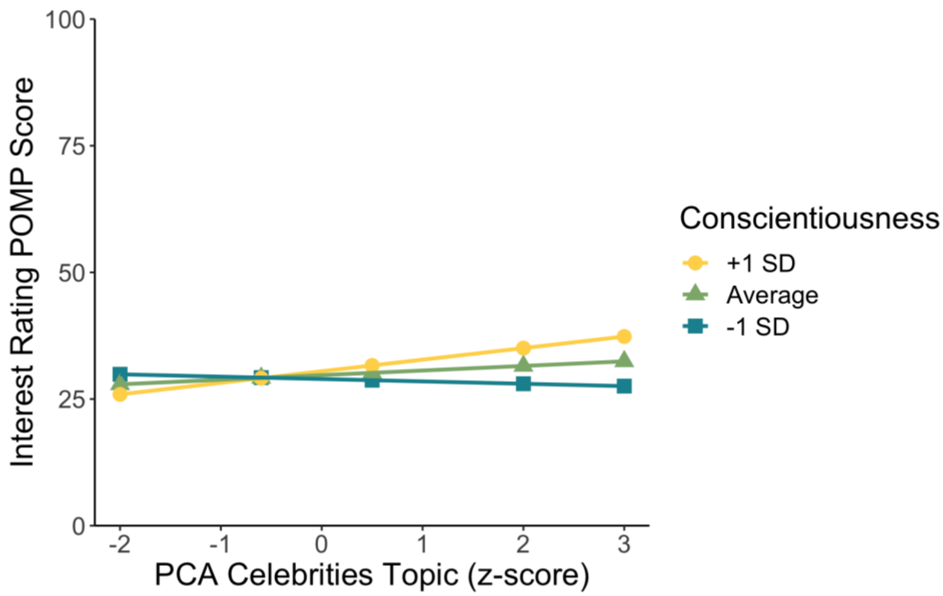
Moderating Effect of Z-scored Conscientiousness on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



The effect of PCA Celebrities topic feature was most positively moderated by Conscientiousness (Figure 50). The main effect of PCA Celebrities was positive, indicating that at an average level of Conscientiousness, every standard deviation increase in celebrity related content relevance resulted in 0.91 POMP units greater interest in following. For someone 1 standard deviation above the mean in Conscientiousness, the simple effect of celebrity content was 2.28. For someone 1 standard deviation below the mean in Conscientiousness, the simple effect of celebrity content was -0.46. In other words, higher levels of Conscientiousness were associated with greater sensitivity to celebrity-related content.

Figure 50

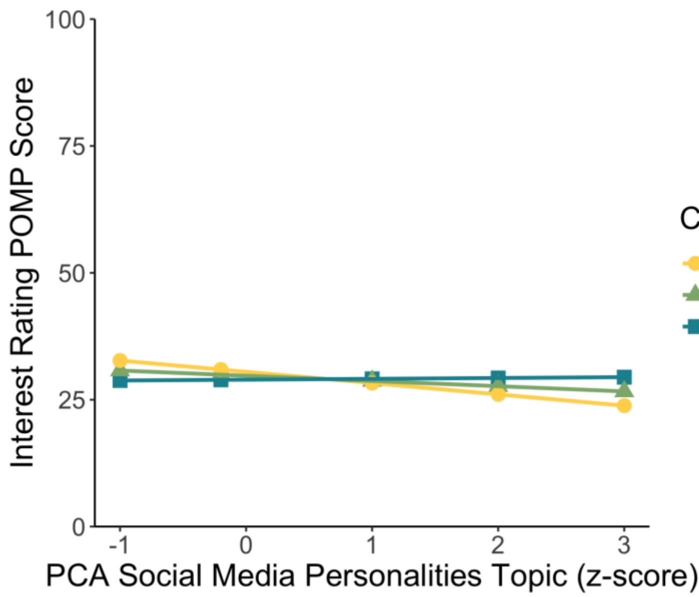
Interaction Plot for Conscientiousness and the Effect of PCA Celebrities Topic



The effect of PCA Social Media Personalities topic feature was most negatively moderated by Conscientiousness (Figure 51). The main effect of PCA Social Media Personalities was negative, indicating that at an average level of Conscientiousness, every standard deviation increase in social media personality related content relevance resulted in 1.03 POMP units less interest in following. For someone 1 standard deviation above the mean in Conscientiousness, the simple effect of social media personality content was -2.22. For someone 1 standard deviation below the mean in Conscientiousness, the simple effect of social media personality content was 0.16. In other words, higher levels of Conscientiousness were associated with greater sensitivity to social media personality-related content.

Figure 51

Interaction Plot for Conscientiousness and the Effect of PCA Social Media Personalities Topic



Discussion

The results of Study 2 add to the findings in Study 1 to provide a fuller and richer perspective of what influences following decisions on Twitter and how personality traits moderate that relationship. An updated sample of Twitter stimuli allowed me to reexamine what traits and features influence interest ratings with accounts that were more relevant to the population of participants. Additionally, expanding the set of Twitter features evaluated in Study 2 allowed for a broader exploration of the elements of a Twitter profile that impact interest in following Twitter accounts. Overall, the results of Study 2 demonstrated notable relationships that further quantify and characterize associations between personality traits and Twitter following decisions.

The first aim of Study 2 was to examine if personality traits influence account following decisions. Similar to Study 1, in both the correlations with interest in individual accounts and the random forests models reflecting aggregate interest, Extraversion and Neuroticism showed the strongest relationships with interest in following Twitter accounts relative to other traits. Those high in Extraversion showed higher interest in following accounts in general, driving some of the strength in this relationship. However, even when accounting for global interest in following accounts a notable relationship was found between Extraversion and interest in following accounts. Comparisons between UO and Prolific subsamples revealed similarity in associations between populations, particularly for traits that correlated most strongly with accounts in the combined sample.

The majority of personality traits demonstrated changes in the strength of relationship between trait and account interest between Study 1 and Study 2. However, in examining this variability more closely, we can see that these changes generally do not vary systematically. For example, the out-of-sample accuracy for Agreeableness increases notably between Study 1 and Study 2 in the random forests analysis. In contrast, the accuracy in the training set for the same trait, decreases between Study 1 and Study 2. A similar pattern occurred for Extraversion, Conscientiousness, and Honesty-Propriety. While training data results generally do not weigh heavily on interpretation, the opposing direction of these relationships indicates noise in our results that make substantive interpretation difficult. This noise is likely a result of the sample sizes of both Study 1 and Study 2. Though there is not a required minimum number of participants, samples of under 300 are on the low end of typical random forest models.

Openness stood out as the only trait whose relationship with account interest systematically decreased between Study 1 and Study 2. One possible explanation for this change is that Openness is related to seeking out more niche or novel content in Twitter profiles. In Study 2, I drew from the most popular accounts followed by this population, which shifted the stimuli to represent more mainstream content for this group. This change in stimuli may have limited individual differences in Openness being expressed in account interest ratings. In opposition, Neuroticism demonstrated a consistent relationship across Study 1 and Study 2. In Study 1, there was concern that the use of stimuli Twitter profiles that were chosen for their prior relationship with mental health variables drove the relationship with Neuroticism. However, the strength of relationship between Neuroticism and account interest remained similar between the two studies, indicating this association is not strongly affected by stimulus selection.

In Study 2, the influence of demographic variables on account following decisions was also tested to get a sense of relative effect size. In general, demographic variables showed slightly stronger relationships with account interest than personality traits. The notable exception was Extraversion, which demonstrated a relationship similar to the demographic variables tested.

Age and SES demonstrated systematic increases in the strength of their relationship with account interest ratings between Study 1 and Study 2. It may be the case that Study 2 stimuli are better for detecting demographic differences in account interest or that the addition of Prolific participants in Study 2 expanded the range of age and SES represented, allowing for a greater effect to be seen.

An extended set of Twitter profiles features were examined to further explore factors that influence interest in following accounts. While all four of the PCA categories in Study 1 had a significant effect on account interest ratings, only three of the seven categories in Study 2 showed a significant effect. However, the strongest positive effect of any feature came from the PCA category of liberal politicians, indicating that the more liberal political content an account had the more interest there was in following that account. Though, this finding may be specific to this population of mostly young University students with generally liberal leaning political views. A number of linguistic features showed a significant influence on account following decisions. Particularly, the LIWC categories of affiliation and power had a strong positive influence on account interest and the category of netspeak (abbreviations in language commonly associated with language used on the internet) had a strong negative influence on account interest. Rater-perceived Neuroticism had the strongest negative impact on following interest, indicating that accounts that were generally viewed as high in Neuroticism were less appealing to potential followers.

These analyses revealed not only patterns of following behaviors, but also the effect that personality traits can have on these patterns. In general, the effects of PCA categories were most strongly moderated by personality traits. For example, the liberal political content had a positive impact on interest in following, but this impact was even more strongly positive for those high in

Neuroticism, Honesty-Propriety, Agreeableness, and Openness. The effect of linguistic features was also frequently moderated by personality traits. For example, the use of netspeak language (emojis and abbreviations) had a negative impact on interest in following, but this impact was even more strongly negative for those high in Honesty-Propriety, Conscientiousness, and Agreeableness. Though perceived characteristics were not often significantly moderated by personality traits, there were a couple notable exceptions. The positive effect of perceived Conscientiousness was even more positive for participants high in Conscientiousness. Similarly, the slightly positive effect of perceived Agreeableness was even more positive for participants high in Agreeableness and the slightly positive effect of perceived Openness was even more positive for participants high in Openness. While I am not testing homophily directly in this study, these results indicate a positive association when similarity in particular personality traits occurs.

Similar to Study 1 results, the effect size of individual profile features were small when considered on their own. This indicates that there may be an upper boundary to what any singular feature can contribute to a following decision. Rather than an individual feature determining a following decision, multiple small effects may come together to create a profile that is appealing to follow. Overall, the results of Study 2 build off of Study 1 results to expand our understanding of the nuanced relationship between personality traits and Twitter account following decisions, though there are a number of limitations to keep in mind. Both Study 1 and Study 2 test a very narrow population of primarily young, educated adults. While 30% of US Twitter users are between the age of 18 and 29 (Wojcik & Hughes, 2019), this group's decisions about following Twitter accounts may not be representative of the larger population. Additionally, these studies utilize hypothetical following decisions in a survey as a proxy for the actual following decisions that a user may make in their own personal Twitter environment. Taken together, these issues limit the generalizability of conclusions related to the relationship between personality and Twitter followed accounts.

Do the results of Study 2 extend to real world following decisions? To further explore the generalizability of these results, Study 3 will examine the relationship between personality traits and profile features in real-world followed accounts. This will allow me to more fully understand if the results found in Study 2 are limited to the controlled environment of a laboratory study or if they extend beyond. Additionally, Study 3 will sample from an expanded US population, which will indicate if these results are population specific.

IV. STUDY 3: GENERALIZING THE INFLUENCE OF PERSONALITY TRAITS TO REAL-WORLD FOLLOWED ACCOUNTS

The purpose of Study 3 is to test the generalizability of findings from Study 2, by examining the relationship between personality traits and Twitter profile features in real-world followed accounts (Aim 3). In this dissertation, I am proposing that when people see a Twitter account, their interest in following that account is at least partly driven by the features of that account. Additionally, this effect may vary based on user personality traits. For example, the features that drive interest for high-Neuroticism people might be different than the ones for low-Neuroticism people. In Study 2, I tested this by showing people different accounts in an experimental context and extracting the features of those accounts to examine how they affected interest in following the accounts. Then, I then used interaction terms to determine if the effects of features on interest correlated with personality traits.

To test the generalizability of these results in Study 3, I will use a naturalistic study design where people have already been exposed to accounts with a variety of features and have made decisions about whether or not to follow those accounts. If the same process is happening for real-world following decisions as I tested in Study 2, then people are choosing to follow different accounts as a function of how their personalities lined up with the features of the accounts they encountered. If this is the case, the results of Study 3 will indicate that people with different personalities have systematic differences in the features of their followed accounts.

An additional concern with the generalizability of Study 2 results is the limited age range of participants, which focused on a population primarily in their early 20s. Study 3 will sample from a non-student focused source, expanding the age range of participants. This updated sample will indicate if the results from Study 2 can be applied beyond this narrow population. Taken together, Study 3 analyses will further broaden our understanding of the associations between personality traits and Twitter following decisions and the generalizability of this relationship. Better understanding this generalizability can also inform future research by indicating if hypothetical following decisions can be used in laboratory studies to accurately study the process of Twitter following decisions.

Materials and Methods

Data Collection Procedure

Study 3 participants were recruited from the “r/beermoney” Reddit community. This community is an online forum where people are able to discuss opportunities to make small amounts of money. All who were interested in participating took a prescreening survey asking if they resided in the United States, spoke fluent English, had at least 25 followers on Twitter, followed at least 25 Twitter accounts, and had posted at least 25 Tweets. To verify eligibility, those who indicated that they met all the requirements had their Twitter account individually checked by a researcher. Eligible participants were sent a link to a Qualtrics survey where they gave their informed consent to participate in the study (or for participants under 18, the participants gave assent and their parents gave informed consent). Consenting participants completed several self-report measures, provided demographic information, answered questions about their Twitter usage, and provided their Twitter handle. After survey completion, participants' Twitter profiles were scraped to collect their Tweets, a list of their followers, and a list of their friends (accounts they are following on Twitter) using the Rtweet package (M. Kearney, 2019).

Data collection for Study 3 occurred in three waves. The first wave of data was collected in 2018 and consisted of N = 284 participants. The second wave of data was collected in 2020 and consisted of N = 358 participants. The final wave of data was collected in 2021 and consisted of N = 60 participants. The combined sample of all three waves of data collection totaled N = 702 participants. Of the 702 participants, friend lists were successfully retrieved from 681 participants. This discrepancy generally arises because participants either deleted, locked, or changed their account name between the time when they completed the survey and when their Twitter data was downloaded.

Data Preparation. A random sample of high degree accounts followed by participants was collected. In this study, a high degree account was defined as an account that had at least 100,000 followers and 100 Tweets. Accounts were also required to use English as their sole or primary language. For each participant, I randomly selected twenty of their followed accounts that met these criteria. This sample of accounts will represent the following decisions of each participant and will be referred to in this study as *followed accounts*. Participants were included in the study if they followed at least 20 high degree accounts, resulting in a sample of N = 452 participants with 9,040 participant and followed account combinations, and 4,236 unique followed accounts.

Participants. Demographics for Study 3 participants are shown in Tables 24-26 and Figure 52. Participants ranged in age from 14-61 years old with an average age of 27.9 years old. Participants received $10 for their participation in the form of either an Amazon gift card or a physical check.

Table 24

Participant Sex for Study 3

Sex	n
Male	238 (53%)
Female	205 (45%)
Other	8 (2%)
Not Reported	1 (1%)

Table 25

Participant Race for Study 3

Race	n
American Indian or Alaska Native	3 (1%)
Asian	40 (9%)
Black or African American	40 (9%)
White	334 (74%)
More than one race	16 (3%)
Other	18 (3%)
Not Reported	1 (1%)

Table 26

Participant Ethnicity for Study 3

Ethnicity	n
Hispanic or Latino	58 (12%)
Not Hispanic or Latino	393 (87%)
Not Reported	1 (1%)

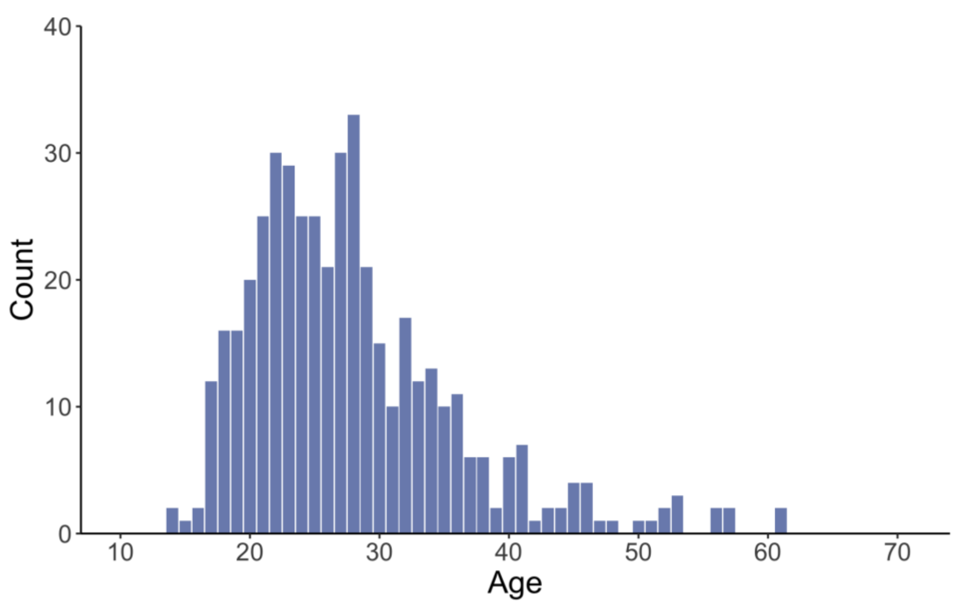
Table 27

Participant Education for Study 3

Highest Education Completed	n
Less than High School	11 (2%)
High School/GED	56 (12%)
Some College	129 (29%)
2-year College Degree	36 (8%)
4-year College Degree	173 (38%)
Master’s degree	36 (8%)
Doctoral Degree	6 (1%)
Professional Degree (e.g., JD, MD)	4 (1%)
Not Reported	1 (1%)

Figure 52

Distribution of Ages for Participants in Study 3



Measures

Self-report Measures. Participants completed self-reports of personality traits using a combination of two measures. The Big Five traits (Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Openness) were measured using the Big Five Inventory 2 (BFI-2; Soto & John, 2017), consisting of 60 short statements rated on a scale from 1 (Disagree Strongly) to 5 (Agree Strongly) with a neutral point of 3 (Neither Agree nor Disagree). Eight items from Questionnaire Big Six measure were used to capture the sixth domain, Honesty-Propriety (Thalmayer et al., 2011). These scales showed adequate internal consistency, with alpha coefficients for the BFI-2 scales ranging from .80 for Agreeableness to .91 for Neuroticism and an alpha coefficient for Honesty-Propriety at .65. Though not analyzed as a part of this dissertation, participants also completed the following self-report measures: Scale for Positive and Negative Experience (SPANE; Diener et al., 2009), Agentic and Communal Values Scale (ACV; Trapnell & Paulhus, 2012), and Interpersonal Support Evaluation List (ISEL; Cohen et al., 1985).

Twitter Profile Features. A select set of Twitter profile features from Study 2 were chosen to be evaluated in Study 3. Due to the large number of followed accounts being included in these analyses, this study will focus on Twitter metadata profile features and linguistic features of Tweets, which can be evaluated at scale (see Table 23 for full description of features). To evaluate linguistic features of Tweets, a set of 100 Tweets was randomly sampled from each of the followed accounts. To give the most accurate representation possible of what a participant may have viewed on an account, the Tweets sampled from an account occurred no later than the year the participant's data was collected. For example, if a participant completed the survey in

2021 and Barack Obama was one of their followed accounts sampled, the Tweets extracted from Barack Obama's accounts would be from no later than 2021. If another participant who followed Barack Obama's account completed the survey in 2018, then another set of Tweets would be pulled from Barack Obama's account from no later than 2018.

Analytic Procedure

Aim 3. The aim of Study 3 is to test the generalizability of the relationship between personality and Twitter profile features with actual followed Twitter accounts. To prepare the data, each followed account will be evaluated for selected Twitter profile features. Then, I will take the mean of each evaluated feature across the set of 20 sampled accounts for each participant, to give us a feature average for the followed accounts of each participant.

I will use Pearson product-moment correlations to assess the relationship between features of participants' followed accounts and personality traits. Each of the 28 features from a participant's sample of followed accounts will be correlated with each of the six participant personality traits, to give us 168 profile feature and personality trait correlation combinations. These correlations will be compared to Study 2 results to assess generalizability of results between the two studies. For each feature and trait combination, the Study 2 interaction coefficient from the multilevel model will be correlated with the corresponding correlation from Study 3. This will result in one correlation per trait, with a higher correlation value indicating a stronger relationship between the effect of that trait in Study 2 and Study 3.

Results

Aim 3: Associations Between Personality Traits and Profile Features in Followed Accounts

Correlations. To examine the strength of the relationship between participant personality traits and Twitter profile features, I calculated Pearson product-moment correlations (Table 30). Honesty-Propriety demonstrated the highest number of notable relationships with profile features, particularly with the linguistic features assessed. Linguistic content related to prosocial behavior and politeness demonstrated the strongest positive relationships while swear words and netspeak demonstrated the most negative relationships. Neuroticism additionally showed several notable associations, including positive relationships with mental health words, swear words, and disgust. Conscientiousness was positively associated with affiliation and achievement words and negatively associated with swear words and anger. Agreeableness was most positively associated with positive sentiment and also showed a notable negative relationship with negative sentiment. Openness showed a positive relationship with positive sentiment, sexual words, and anger. Finally, Extraversion showed very weak relationships with this set of Twitter profile features, with the only notable positive relationship being with substance words (e.g. beer, drunk, cigar).

As noted earlier, in Study 2, the interaction terms reflected how traits were correlated with sensitivity to features. In Study 3, the correlations showed how traits were associated with the actual presence of features. To compare these results across studies, I correlated the interaction terms from Study 2 with the corresponding correlation values in Table 28. These results demonstrated that all traits except Openness showed similarity across studies, with correlations ranging from .30 to .60 (Table 29). Relationships with Neuroticism and Conscientiousness were most similar between Study 2 and Study 3. The influence of Openness showed almost no similarity between Study 2 and Study 3.

Table 28

Correlations Between Twitter Profile Features and Personality Traits in Study 3

	Honesty	Neuroticism	Conscientiousness	Agreeableness	Openness	Extraversion
Count of Followers	**-.11**	-.04	.04	.00	.00	.08
Count of Friends	.05	-.08	.06	.02	.07	.01
Tweet Frequency	.08	.00	.03	-.01	.00	-.05
Ratio of Retweets to Tweets	.00	-.03	.02	-.04	-.03	-.02
Average Word Count	**.10**	-.07	.04	-.01	.06	-.01
LIWC Affiliation	**.12**	-.05	**.10**	.07	.05	-.01
LIWC Achievement	.06	**-.14**	**.10**	.04	-.06	-.01
LIWC Power	.00	-.02	.06	-.06	.00	.04
LIWC Prosocial	**.15**	-.05	.03	.04	-.01	-.01
LIWC Politeness	**.11**	.05	-.08	-.02	-.03	-.06
LIWC Conflict	-.06	.04	-.01	**-.12**	.04	.02
LIWC Moralization	.00	**.11**	-.03	-.08	.01	-.07
LIWC Communication	**.10**	.01	-.04	-.04	.03	.00
LIWC Mental Health	-.03	**.13**	-.02	.01	-.02	-.04
LIWC Swear Words	**-.11**	**.15**	-.07	-.01	.04	.01
LIWC Substances	-.02	.03	-.05	.00	.06	**.10**
LIWC Sexual	-.05	**.15**	**-.10**	.01	**.14**	.00
LIWC Netspeak	**-.12**	**.11**	-.06	.00	.00	-.01

Table 28 Continued

	Honesty	Neuroticism	Conscientiousness	Agreeableness	Openness	Extraversion
NRC Pos Sentiment	.07	.05	.01	**.12**	**.15**	.05
NRC Neg Sentiment	-.02	.07	.00	**-.10**	.00	.02
NRC Anger	-.07	**.10**	**-.14**	-.05	**.13**	-.06
NRC Anticipation	.02	-.07	.04	.01	.02	-.05
NRC Disgust	-.07	**.19**	-.09	-.07	**.12**	.00
NRC Fear	.05	.00	.07	-.02	.06	.03
NRC Joy	**.10**	.07	-.03	.05	.06	-.02
NRC Sadness	.01	.09	-.02	-.07	-.02	.01
NRC Surprise	-.05	.03	-.05	-.05	.06	.03
NRC Trust	.05	**-.10**	.07	.05	.04	-.02

Note. Correlations greater than or equal to .1 or less than or equal to -.1 are bolded to indicate notable relationships.

Table 29

Correlations Between Study 2 Interaction Coefficients and Study 3 Correlations

	r
Neuroticism	.60
Conscientiousness	.51
Honesty	.42
Agreeableness	.31
Extraversion	.30
Openness	.01

Discussion

Study 3 explored the generalizability of the relationship between personality and Twitter profile features. Using a sample of actual followed Twitter accounts, I examined if the effects between personality traits and profile features found in Study 2 could be extended beyond a set of hypothetical following decisions in an expanded population. Overall, I found substantial positive relationships between the results of Study 2 and Study 3 for all traits except Openness, indicating a notable amount of generalizability of these findings.

In comparing the relationship between personality traits and Twitter profile features for hypothetical and real-world following decisions, Neuroticism demonstrated the strongest association between Study 2 and Study 3 results. This finding indicates that the moderating effect of Neuroticism on features in Study 2 is most generalizable to real-world following decisions in this expanded population. From this we can infer that Neuroticism affects the hypothetical decision-making process similarly to the actual commitment of following an account in the real-world. One possible explanation is that Neuroticism also influences how people find Twitter accounts to follow in the real-world. For example, a person high in Neuroticism may be more likely to click into a Twitter profile and examine historical Tweets before deciding if they want to follow the account. As a result, the laboratory method of showing Twitter profiles to participants may better reflect individual differences in Neuroticism than other traits.

In contrast, Openness stood out as having almost no relationship between Study 2 and Study 3 results. Several features that had a positive relationship with Openness in Study 2 showed a negative relationship in Study 3 and vice versa. This finding indicates that Openness may affect the hypothetical decision-making process differently than the actual commitment of following an account in the real-world. Between Study 1 and Study 2, the effect of Openness was most systematically influenced by the set of stimuli used. Relating this to back the broader possibility that Openness may influence how much a person seeks out niche or novel content, using a set of chosen Twitter stimuli in a laboratory study may not accurately represent how individual differences in Openness influence the evaluation of Twitter accounts in the real world. Another possibility is that the relationship between Openness and profile features is more closely tied to age than other traits. Openness in general tends to increase after the age of 20 (Soto et al., 2011), which may indicate that Study 2 findings related to Openness were specific to that restricted age group and do not generalize to a population with an expanded age range.

Correlations between personality traits and the select set of profile features evaluated in the sample of followed accounts revealed a number of intuitive relationships. Honesty-Propriety demonstrated positive relationships with linguistic content related to prosocial behavior and politeness but demonstrated negative relationships to swear words and netspeak. In considering Honesty-Propriety as a trait broadly related to conformity of social decorum, these relationships indicate that participants high in this trait curate their social media feeds in accordance with these social values. Neuroticism showed positive relationships with mental health words, swear words, and disgust. This not only further supports previous connections seen between Neuroticism and the mental health variables (Duggan et al., 1990; Boyce et al., 1991; Saklofske et al., 1995; Muris et al., 2005), but also further indicates associations with features related to negative mood states. Finally, Agreeableness was most positively associated with positive sentiment and also showed a notable negative relationship with negative sentiment, indicating an association with affect of text. Previous research has indicated that people high in Agreeableness tend to go out of their way to look at pleasant rather than unpleasant things (Bresin & Robinson, 2015) and the same process may be occurring in Twitter following behaviors with Agreeableness impacting following accounts based on valence of text.

The findings reproduced in Study 3 endured a robust test of generalizability in a real-world context. Overall, these results demonstrate strong evidence that the effects between personality traits and profile features found in Study 2 can be extended to real-world following decisions for all traits except Openness. With the ultimate goal of understanding real-world following decisions on Twitter, these results indicate that hypothetical following decisions used in laboratory methods are a viable means to study the relationship between personality traits and features of Twitter accounts.

V. GENERAL DISCUSSION

Social media and its growing popularity have brought about a new domain of social interaction. In understanding these behaviors, it is important to consider the transactional relationship between the person and the situation that shapes our online experience. To what extent is personality reflected in the environments we create for ourselves online? In this series of studies, I took an exploratory approach to examining if personality traits can be used to better understand the account following decisions we make on Twitter. Though this research is a first step in exploring the influence of personality on the vast number behaviors that occur on social media, these findings establish a foundational understanding of this relationship and inform future research.

Personality psychologists are interested in exploring how people select and modify their environment and what drives those behaviors. The results of these studies provided an opportunity to begin examining the directionality of the connection between personality and social media behaviors. The laboratory method of collecting personality metrics prior to asking about hypothetical account interest distilled the impact of personality on following decisions. This ruled out effects in the opposite direction (i.e., that a followed account can affect personality in this data). This is not to say that this relationship doesn't exist outside of these experiments. Like many behaviors, real-world following decisions are likely a two-way relationship with both the person and the environment impacting each other. However, being able to pull apart these effects has demonstrated that personality traits have their own independent effect on how we curate our social environment.

Overall, results demonstrated that personality does influence the decisions we make about which Twitter accounts to follow and in turn, how our social media environment is curated. The strength and stability of this relationship showed some heterogeneity across traits, though is generally comparable to the effect of some commonly used demographic variables. Additionally, these results demonstrated some similarities to findings from previous analogous studies. Azucar's (2018) meta-analytic assessment of the relationship between personality traits and a wide variety of online behaviors also found the strongest association with Extraversion and the weakest with Agreeableness. Additionally, Costello et al.'s (2022) naturalistic examination of the relationship between personality traits and Twitter accounts followed by participants found the weakest associations with Conscientiousness and Agreeableness. Taken together, these results support a converging body of evidence demonstrating that some traits are more strongly expressed in our online behaviors than others.

Underlying Psychological Processes Linking Traits to Account Features

Beyond understanding the strength of these relationships, I examined characteristics and groups of Twitter accounts in relation to user personality traits to identify patterns in behavior. Findings indicated that personality traits of users align with characteristics of Twitter accounts and moderate the effect of different Twitter profile features on following decisions. What principles and theories can be used to explain this alignment of characteristics and features? Drawing on broader knowledge of personality traits, we can take an interpretive approach to speculating about psychological processes that drive these relationships which can inform the development of theories to be tested in future research.

Extraversion demonstrated the strongest relationship with account following interest. Further analysis indicated that some of the strength in the relationship with Extraversion was driven by extraverts wanting to follow more accounts in general. However, notable relationships with specific accounts were found even when accounting for this propensity. Those high in Extraversion were more positively influenced by features that could be interpreted as "popularity," (such as number of followers or content related to mainstream influencers) than for those low in Extraversion. Extraverts themselves tend to be popular (Feiler & Kleinbaum, 2015) and enjoy socializing with lots of people, which may be reflected in having interests that are shared by other people. Additionally, features such as number of followers can be seen as symbols of social status, which extraverts often pursue (DesJardins et al., 2015). Though previous studies have shown a connection between Extraversion and higher word usage (Gill & Oberlander, 2019), this relationship was not demonstrated in interest in followed accounts. While extraverts may use more words, they aren't necessarily attracted to Twitter accounts that also use more words. On the opposing end, Extraversion showed a strong negative relationship with features related to negative affect. While negative emotionality is typically connected to Neuroticism, those features may also represent the opposite of mainstream popularity. That is to say, introverts may be less put off by accounts that have generally unappealing traits such as negative affect than their extraverted counterparts.

Neuroticism demonstrated consistently strong relationships with account interest and Twitter profile features across all three studies. This finding is particularly notable because there was concern that the use of stimuli Twitter profiles that were chosen for their prior relationship with mental health variables drove the relationship with Neuroticism in Study 1. However, the strength of relationship between Neuroticism and account interest remained similar between Study 1 and Study 2, indicating this association is not strongly affected by stimulus selection. Those high in Neuroticism were most strongly influenced by features that related to general perceptions about which gender, age, and political beliefs a Twitter profile appealed to.

Specifically, there was a positive influence of accounts that appeal to women, young people, and liberals, which align with gender differences and age trends in Neuroticism (Schmitt et. al, 2008; Soto et. al, 2011) and match the attributes of the majority of participants (see Appendix C for gender-centered analyses). In addition, the influence of several content categories examined aligns with the perceived gender appeal of accounts. Those high in Neuroticism were more positively influenced by accounts with celebrity content, which notably included Harry Styles and other former One Direction members who are known to have a largely female fanbase. On the opposing end, those low in Neuroticism were more positively influenced by accounts with rap/R&B and sports content, which are topics more commonly associated with men. Neuroticism is also associated with social anxiety (Naragon-Gainey & Watson, 2011), indicating a heightened awareness of judgements of others. Those high in Neuroticism may view their list of followed accounts as a public behavior subject to judgment from others. If so, general perceptions of who an account appeals to would be particularly salient to their following decisions.

Agreeableness demonstrated notably strong relationships with a number of linguistic categories. Those low in Agreeableness were less negatively influenced by linguistic features that could be interpreted as socially improper (swear words, sexual language, anger) than their counterparts high in Agreeableness. A related study from Schwartz et. al (2013) found that the use of anger-related words in Facebook statuses was predictive of being low in Agreeableness, indicating similar connections between personality and linguistic features that people produce as well as follow. Conversely, those high in Agreeableness were more positively influenced by perceived Agreeableness and affiliation words (e.g., we, our, us). Agreeableness is associated with cooperation and prosocial behavior, and people tend to form relationships with others who have similar levels of agreeableness (Tracey et al., 2001). Interest in accounts that indicate similar communal views and an avoidance of accounts that do not adhere to prosocial norms may reflect individual differences in Agreeableness.

Honesty-Propriety demonstrated some similar relationships with features as Agreeableness, with swear words and netspeak (emojis and abbreviations) having a higher negative influence for those high in Honesty-Propriety. As a trait broadly related to social decorum, those high in Honesty-Propriety may view these language cues as a sign of nonconformity with proper social values. Similarly, those high in Honesty-Propriety were positively influenced by accounts that used words associated with prosocial behavior and were perceived as appealing to an older age. These features may have indicated attributes of maturity and adherence to social values that are seen as desirable by those high in Honesty-Propriety.

Those high in Openness were more positively influenced by content related to liberal political beliefs. Openness as a trait is often connected to beliefs such as political ideology (Jost et al., 2003; Ozer & Benet-Martínez, 2006) and users tend to follow accounts that align with their views (Himelboim et al., 2013). In opposition, Openness most negatively influenced the effect of sports content and mainstream celebrity content. Though these features had an overall negative impact on account interest, these effects were especially negative for those high in Openness. High levels of Openness predict interest in activities such as visiting museums, reading literature, and creating art (McManus & Furnham, 2006), which could be seen in opposition to more mainstream interests such as sports or celebrities. These results indicate that those high in Openness are less likely to follow along with mainstream interests than their counterparts low in Openness.

Similar to Extraversion, those high in Conscientiousness were more negatively influenced by features related to negative mood states, such as swear words, anger, and negative affect. While these features all have a slightly negative impact on account interest, this effect was more negative for those high in Conscientiousness. These results also align with Schwartz's (2013) finding that use of swear words in Facebook statuses was predictive of being low in Conscientiousness, again demonstrating parallel connections between personality and linguistic features that people produce as well as follow. Self-regulation and planning ahead are attributes of Conscientiousness (McCrae & Löckenhoff, 2010) which may allow those high in this trait to foresee potential emotional consequences of following an account with a lot of negative content.

Taking a step back to consider these interpretations as a whole, we can draw conclusions across traits about social media behavior. Personal interests and their previous association with personality motivated this research to explore if traits influence the interests we pursue in a social media context. However, my interpretation of these results hints at something deeper than the surface level of how we typically think about interests. Rather than following accounts just to seek information, these connections with personality indicate that people may also be considering what type of experience they want to have on a social media platform when they consider who to follow. Extraverts want to feel connected to popular accounts and seek content on topics that lots of other people care about. Those high in Openness seek to broaden their worldview by following accounts with less mainstream content. Neuroticism is associated with following accounts that conform to gender and age norms. Agreeableness, Honesty-Propriety, and Conscientiousness are associated with curating an environment that is low in social norm-violating language. Taken together, these patterns can inform broad theories about the influence of personality on how people approach social media platforms and what they want to get out of their experience.

Generalizing the Relationship between Personality Traits and Twitter Profile Features

To what extent might the conclusions of this study apply in other situations or circumstances? Though a number of interesting connections were drawn between personality and account features when using hypothetical following decisions in a controlled laboratory environment, a goal of this research is to understand behavior in real-life social media environments. To test this possibility, the generalizability of the association between personality and Twitter account features was examined in a set of real-world following accounts. Overall, the effects between most personality traits and profile features found in real-world followed accounts demonstrated remarkable similarity to results found when using hypothetical following decisions.

While examining the translation of hypothetical following decisions to real-world followed accounts is a natural first step in understanding the generalizability of these findings, there are a number of additional factors to address in exploring how far these results extend. When considering the relationship between personality and Twitter profile features, how much are the effects universal vs. the impact of a particular cultural or historical context? The results of this dissertation were interpretable in relation to existing knowledge on personality, suggesting that they may have been picking up on at least some universally relevant connections. Additionally, testing the generalizability of results between Study 2 and Study 3 did incorporate a somewhat expanded age range, indicating these effects may translate across age groups. However, there are other relevant considerations. In Study 2, Honesty-Propriety was found to be most negatively influenced by content related to Rappers and R&B artists, which primarily feature black artists. While rap and R&B content generally had a positive influence on interest in following Twitter accounts, this influence became negative for those higher in Honesty-

Propriety. With a primarily white sample, this result raises the question of what stereotypes may be influencing these relationships and how stable these results may be across populations with different racial and ethnic make ups. Further testing these effects across different demographic groups and cultural contexts would better indicate the universality of these relationships.

A continual issue in considering generalizability of social media behaviors is the historical context in which behaviors occur. Larger world events may influence the kind of experiences that people seek on social media. In considering the data used in this dissertation, tumultuous political events or global health crises may have led some people to seek comforting content on social media and avoid negative emotional content. For example, Conscientiousness is associated with a desire for predictability (Berembaum et al., 2008) and uncertain times such as the COVID-19 pandemic may heighten the use of self-regulation strategies like avoiding content high in negative affect. Additionally, results found in data-driven approaches connecting personality and social media behaviors may change as platforms continually modulate based on user-generated content. Language in particular has evolved rapidly due to social media platforms like Twitter, incorporating both new words as well as new forms of expression such as emojis (Androutsopoulos, 2011; Dimson, 2015). In connection to personality, the use of emojis and abbreviations was found to have a particularly negative impact on following decisions for those high in Honesty-Propriety, likely because they are seen as less socially proper forms of communication. However, as social media drives the evolution of language over time and netspeak permeates other forms of communication, social norms may evolve and lessen the negative relationship with Honesty-Propriety. These ongoing changes present difficulty in distilling generalizable relationships and future research may benefit from longitudinal data collection to better identify universal relationships.

Limitations and Future Research

While this research has taken steps to establish foundational knowledge about how who we are is reflected in social media environments, there are limitations to these findings that can inform future research. A number of psychological processes were speculated to drive the relationship between personality and the effect of Twitter profile features on following decisions, though none of these processes were tested directly in these studies. Future research can utilize these theories to build testable hypotheses about specific relationships between personality and Twitter account following decisions. For example, it was speculated that social anxiety may be a driving influence on Twitter following decisions for those high in Neuroticism. Additional studies could collect participant levels of social anxiety to examine a potential pathway from social anxiety to Neuroticism to Twitter following decisions.

Additionally, future research could also work toward distilling the effect of these profile features even more by creating artificial stimuli that have differing levels of a given feature. For example, negative affect of text was found to be an influential feature on following decisions and showed relationships with several personality traits. This dissertation took the approach of collecting text from existing Twitter accounts and assessing the sentiment. However, correlations among Twitter profiles features indicated strong relationships between negative affect and the use of power or disgust words as well as political content. This multicollinearity of features makes it difficult to determine if there is a unique effect of negative affect on following decisions or if negative affect is just present in profiles where political content is actually driving following decisions. While the data-driven approach taken in this dissertation was necessary to narrow down promising features, future research creating artificial accounts with low and high levels of negative affect while controlling for other features such as content topics would further distill the unique contribution of that feature on Twitter following decisions.

My examination of Twitter accounts in this study was limited to those with a high number of followers which also generally meant that they were well-known people or companies. While popular accounts are useful to study because they are likely to appeal to this particular sample, they only represent a small percentage of accounts on Twitter. Many Twitter users follow accounts of users they know in real life such as friends or family members and some may use Twitter for other purposes such as building their professional network (e.g., academic and medical professionals) or connecting with others going through similar experiences (e.g., such as chronic illness communities, Bedford-Petersen & Weston, 2021). Popular accounts on Twitter tend to focus on a particular topic or style of content in order to develop an identifiable brand, which could lead to overinflation of the effect of interests or language style on the relationship between personality and account following decisions. Future research could examine other types of Twitter accounts and incorporate user reasons for following accounts to more fully understand the breadth of following behaviors on Twitter.

Relatedly, these studies did not account for the effect that a user's prior knowledge of a Twitter account may have on following decisions. Given that accounts used in these studies were relevant to the participant community, it is likely that participants already were familiar with some of the people or brands behind the accounts. For example, in qualitative responses, several participants indicated that they would not follow beauty influencer James Charles due to ongoing sexual misconduct allegations. This information was not represented anywhere in the screenshot of his profile, so participants that indicated no interest in following him may have done so due to prior perceptions of him rather than an evaluation of his profile. Future studies examining popular accounts should incorporate participant familiarity with an account into inferences about Twitter following decisions.

Practical Implications

Beyond an academic context, what are the practical implications of these findings? Social media companies are a prime audience that can stand to benefit from this information. With their aim of keeping users engaged with their platform, these companies are interested in learning more about attributes of their users and how it influences their behavior. Typically, their analyses utilize behavior metrics as well as demographic variables to better understand user experiences and predict future behavior, such as account following. The results of this dissertation suggest that personality does impact Twitter following decisions and that traits align with characteristics of Twitter accounts. Though personality as a whole demonstrates substantial association with following behaviors, when we unpack relationships with account features, a more complicated story is revealed. Small effect sizes of these results indicate that the way personality influences online behavior is through many small effects. In turn, these findings suggest that using personality metrics on their own wouldn’t revolutionize the predictive power of current algorithms aimed at predicting user behavior. With similar or slightly less predictive power than demographic variables, these traits may be seen as a contributor to the larger understanding of user attributes that preempt behaviors on social media.

Beyond the goal of just using personality to predict account following, social media companies could leverage personality to provide context for patterns in behaviors of their users. For example, Twitter may be interested in why particular groups of accounts have a lot of followers in common to better understand user engagement with those accounts. A natural first step might be to look at the network connections of those particular profiles and draw conclusions based on those features. Motamedi et al. (2016) examined network connections among highly followed Twitter accounts to identify groupings of communities with cohesive themes. Their findings generally indicated groupings based on language or geography. However, they also found a grouping based on content centered around United States based actors and celebrities similar to content groupings found Study 1 and Study 2 results. While personality traits aren't necessary to draw surface level connections, incorporating personality in relation to its larger body of research can provide context for patterns in findings. For example, the results of this dissertation indicated that celebrity accounts are typically followed by those high in Neuroticism, which has been connected to a potentially emotionally unhealthy desire for parasocial relationships with famous people (Maltby et al. 2003, 2011). Understanding this broader connection can provide context for psychological processes that may influence groupings of network connections.

Finally, contextual information from personality could be utilized by social media companies to shape experiences of users, providing them more positive interactions with their platform. While accounts with high negative affect may share a lot of followers, incorporating personality indicates that these followers also tend to be high in Neuroticism. Drawing on personality literature provides context that people who are high in Neuroticism are vulnerable to negative mood states (Gomez et al., 2000). This deeper understanding of user behavior could prompt a social media company to modify user experiences by suggesting content that would lessen negative mood states or reminding users to take a break from the platform.

Conclusion

The increasing digitization of our social world has opened up new possibilities for people to curate their online social experiences, which presents researchers a unique opportunity to explore the relationship between who we are and our online behaviors. Examining this relationship has provided a better understanding of how characteristics, such as personality traits, drive the selection and modification of our social environments. As a relatively new area of research, the lack of strong theory to guide testing of specific hypotheses highlighted the importance of data-driven exploratory research to establish foundational information about how who we are is reflected in our social media environments.

Overall, findings indicate that personality does influence Twitter account following, though there is some heterogeneity among traits in both strength and consistency of this relationship. Personality traits of users also align with characteristics of Twitter accounts and moderate the effect of different Twitter profile features on our following decisions, highlighting potential psychological processes that drive following decisions. Remarkable generalizability was demonstrated for the effects between most personality traits and profile features when tested on real-world followed accounts, though additional cultural and historical factors should be considered in fully understanding the universality of these findings.

Future research can use psychological processes indicated in the interpretation of these results to build testable hypotheses and further distill the effects of individual Twitter profile features. Additionally, future research can better understand a user’s journey to following an account by investigating how different types of Twitter accounts and prior knowledge of those accounts affect the relationship between personality and following decisions. Finally, the impact of these results can be extended beyond an academic context to provide social media companies context for behaviors of their users. Taken together, these findings begin to disentangle the complicated relationship between who we are and how we construct our social environments, as well as uncover some potential underlying psychological processes that drive these relationships. Though this research is a first step in exploring the influence of personality on the vast number behaviors that occur on social media, these findings establish foundational knowledge and inform the development of theory.

APPENDIX A

STUDY 1 PRINCIPAL COMPONENTS ANALYSIS RESULTS

Figure A1

Scree Plot of Eigenvalues in Study 1

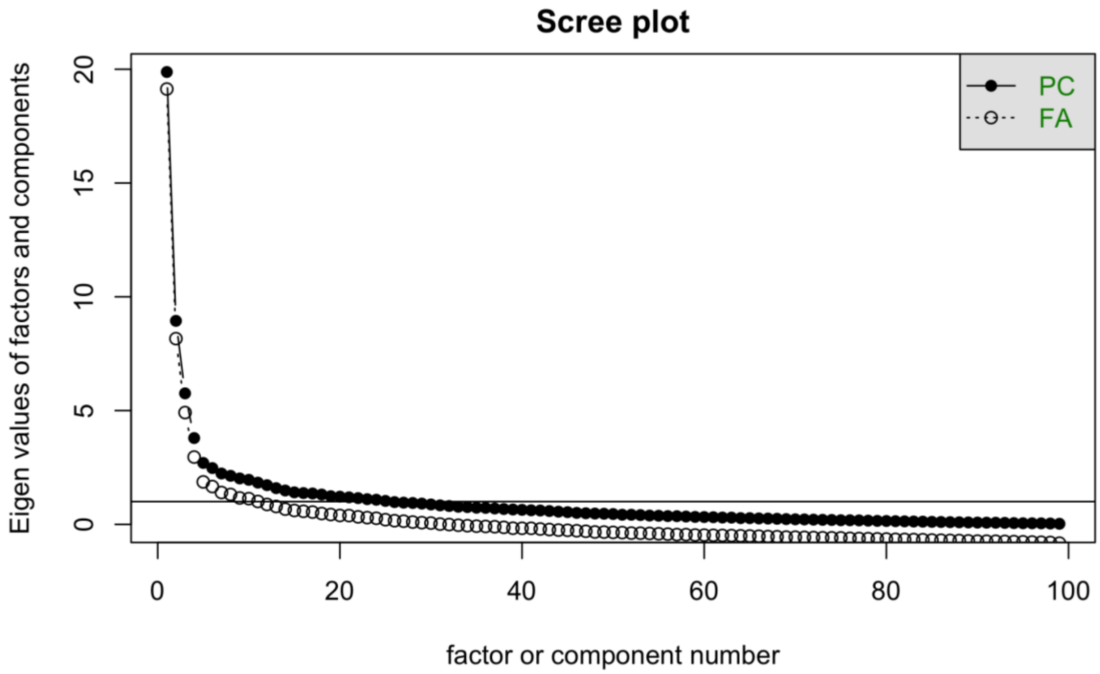


Table A1

Variance Accounted for in a 4 Component Solution with Varimax Rotation in Study 1

	RC1 (Sports)	RC2 (Gaming)	RC3 (Actors)	RC4 (Mom Bloggers)
SS loadings	13.77	9.99	7.38	7.23
Proportion Var	.14	.10	.07	.07
Cumulative Var	.14	.24	.31	.39
Proportion Explained	.36	.26	.19	.19
Cumulative Proportion	.36	.62	.81	.00

Table A2

Loadings for a 4 Component Solution with Varimax Rotation in Study 1

Twitter Account	RC1 (Sports)	RC2 (Gaming)	RC3 (Actors)	RC4 (Mom Bloggers)
McShay13	.84	.06	.01	.06
SportsCenter	.83	-.06	-.08	.04
KrisBryant_23	.82	-.13	-.15	.07
MelKiperESPN_2	.81	.14	.08	-.02
NFL	.76	-.09	-.01	.13
obj	.76	-.11	.05	.02
AdamSchefter	.74	.24	.06	-.01
mortreport	.74	.10	.08	.25
YahooSportsNBA	.74	.09	-.09	.24
NFL_Memes	.74	.02	.06	.23
Arrieta34	.73	-.05	.17	.26
Rotoworld_FB	.73	.24	-.06	.19
Ken_Rosenthal	.70	.17	.11	.20
PatMcAfeeShow	.68	.25	.06	.02
StuartScott	.66	.19	.00	.04
DangeRussWilson	.65	-.08	.21	.18
PFTCommenter	.64	.33	.05	-.06
RickieFowler	.62	-.01	.19	.12
TeamUSA	.60	-.11	.14	.16
bubbawatson	.56	.00	.19	.25
jimrome	.55	.41	.05	.26
BarstoolBigCat	.54	.19	.08	.02

Table A2 Continued

Twitter Account	RC1 (Sports)	RC2 (Gaming)	RC3 (Actors)	RC4 (Mom Bloggers)
McllroyRory	.53	.09	.19	.05
solecollector	.46	.22	.12	.03
FamilyGuyOnFOX	.36	.10	.31	.09
stoolpresidente	.32	.04	.19	-.18
michaelianblack	.29	.12	.24	.17
elgatogaming	.22	.74	-.01	.09
LogitechG	.15	.73	-.03	.13
IronsidePC	.13	.73	-.05	.14
ZOWIEbyBenQUSA	.15	.70	-.14	.12
teksyndicate	.09	.66	-.01	.10
NorthernlionLP	.09	.66	.10	-.01
SMITEGame	.21	.65	-.11	.08
notch	.05	.61	.15	.01
Monstercat	.10	.60	.15	.23
ElderScrolls	-.07	.58	.03	-.01
feedme	-.03	.53	.34	.12
RiffTrax	.16	.50	.24	.27
verge	.17	.47	.13	.07
TheWookieeRoars	.10	.45	.21	.16
RBReich	-.04	.43	.25	-.04
GroovyBruce	.06	.41	.36	.26
Totalbiscuit	.12	.39	.32	.33
thesulk	-.21	.39	.36	.00

Table A2 Continued

Twitter Account	RC1 (Sports)	RC2 (Gaming)	RC3 (Actors)	RC4 (Mom Bloggers)
snopes	-.04	.39	.14	.04
newbelgium	.19	.39	.28	.13
newcastle	.29	.36	.16	.13
drewmagary	.18	.36	.25	.10
TomHall	.30	.34	.17	.30
NaNoWriMo	-.10	.33	.15	.19
ABFalecbaldwin	.29	.33	.30	.13
AmznMovieRevws	-.18	.32	.23	.20
InternetHippo	-.25	.31	.27	-.02
TheWalkingDead	.19	.30	.29	.07
MrCraigRobinson	.20	.22	.64	-.17
RobLowe	.20	-.10	.59	.14
arnettwill	.13	.28	.58	.13
ElizabethBanks	.20	-.22	.56	.23
ThatKevinSmith	.04	.40	.56	.01
CobieSmulders	-.03	.06	.55	.03
GilianA	.01	-.02	.53	.18
SteveCarell	.28	.06	.51	.12
melissamccarthy	.21	-.30	.47	.30
batemanjason	.23	.18	.45	.13
PrettyLights	.10	.25	.45	.14
BobsBurgersFOX	-.03	.10	.43	.14
ders808	.03	.20	.43	.08

Table A2 Continued

Twitter Account	RC1 (Sports)	RC2 (Gaming)	RC3 (Actors)	RC4 (Mom Bloggers)
omgthatspunny	.11	.08	.42	.26
EdwardNorton	-.09	.37	.42	-.10
lustrelux	.03	-.15	.41	.32
Chrisspymakeup	-.23	.06	.38	.29
joelmchale	.17	.29	.38	.18
JLo	.34	-.30	.35	.28
BBCAMERICA	.07	.31	.35	.16
Lilpeep	-.02	.21	.33	.00
StateDept	.17	.18	.31	.00
Lovesmytwoboys	.13	-.04	.04	.66
ConservamomE	.08	-.07	.08	.66
lifewithheidig	.11	.09	.06	.66
FSOC2011	.06	.28	.10	.63
KellysLuckyYou	.06	.16	.23	.62
jonbonjovious	.18	.07	.00	.60
SeeMomClick	.15	.21	.13	.58
CouponsFreebie	.16	.22	.09	.58
NIVEAUSA	.15	.10	.24	.51
twokidsandacoupon	.06	.31	.08	.49
MissingLynxx	.03	-.06	.23	.48
PBnWhine	.07	.24	.13	.47
SanitySuburbia	.04	.10	.28	.41
Sleepopolis	.13	.40	.05	.40

Table A2 Continued

Twitter Account	RC1 (Sports)	RC2 (Gaming)	RC3 (Actors)	RC4 (Mom Bloggers)
TomiLahren	.33	.06	.01	.36
BestBuy_Deals	.27	.28	.03	.34
SpiderManMovie	.26	.11	.18	.26
ToysRUs	.22	.18	.15	.23

APPENDIX B

STUDY 2 PRINCIPAL COMPONENTS ANALYSIS RESULTS

Figure B1

Scree Plot of Eigenvalues in Study 2

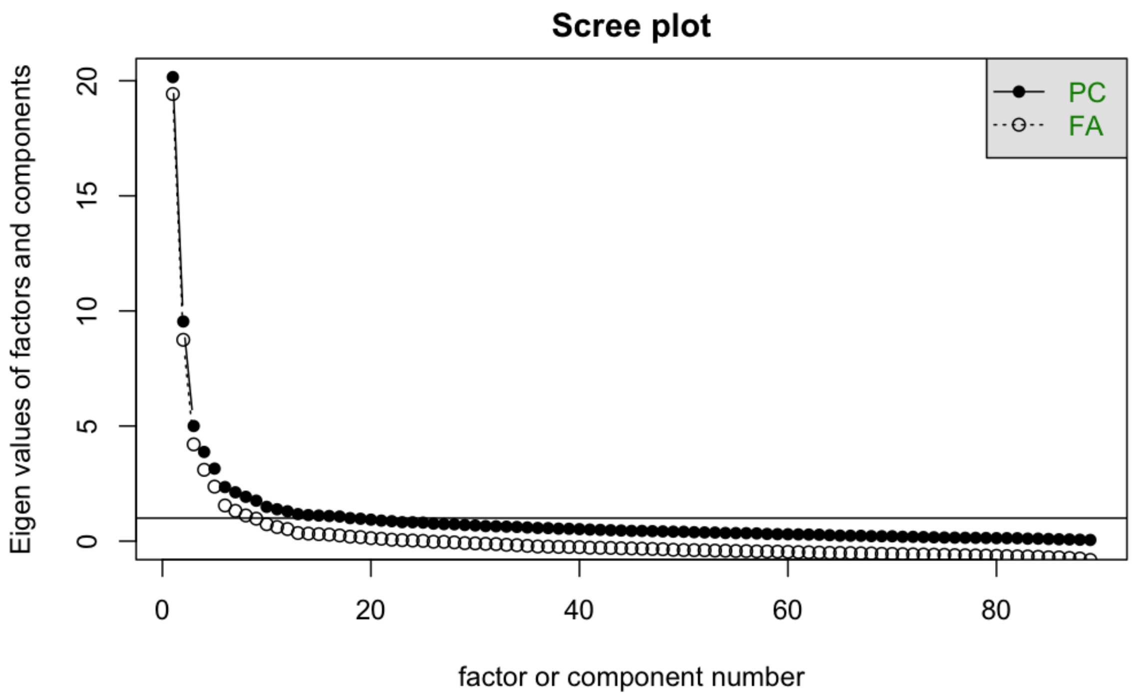


Table B1

Variance Accounted for in a 7 Component Solution with Varimax Rotation in Study 2

	RC1 (Celebrities)	RC2 (Sports)	RC4 (Rap/R&B)	RC6 (Mainstream Influencers)	RC5 (Liberal Politicians)	RC3 (Traditional Comedy	RC7 (Social Media Personalities)
SS loadings	8.39	8.11	8.09	7.62	6.30	4.36	3.32
Proportion Var	.09	.09	.09	.09	.07	.05	.04
Cumulative Var	.09	.19	.28	.36	.43	.48	.52
Proportion Explained	.18	.18	.18	.17	.14	.09	.07
Cumulative Proportion	.18	.36	.53	.70	.83	.93	1.00

Table B2

Loadings for a 7 Component Solution with Varimax Rotation in Study 2

Twitter Account	RC1 (Celebrities)	RC2 (Sports)	RC4 (Rap/R&B)	RC6 (Mainstream Influencers)	RC5 (Liberal Politicians)	RC3 (Traditional Comedy)	RC7 (Social Media Personalities)
NiallOfficial	.80	.05	-.06	.16	.02	.11	.11
onedirection	.77	.00	-.02	.21	.11	-.02	.09
Harry_Styles	.72	-.10	.01	.29	.21	-.06	.13
Louis_Tomlinson	.71	.09	-.02	.23	.10	.02	.06
Luke5SOS	.69	.06	-.02	.01	.04	.16	.17
zaynmalik	.67	.06	.04	.34	.03	.12	-.02
dylanobrien	.61	-.08	.15	.18	.16	.20	.17
LiamPayne	.57	.11	-.08	.36	.07	.13	-.07
MileyCyrus	.53	-.18	.10	.31	.34	.00	.12
troyesivan	.53	-.19	.04	-.08	.17	.06	.22
TomHolland1996	.51	.10	.15	.20	.16	.41	-.11
Zendaya	.50	-.12	.34	.27	.32	.03	-.01
dylansprouse	.48	.03	.04	.30	.25	.27	.03
theestallion	.46	-.12	.43	-.08	.21	-.15	.15

Table B2 Continued

Twitter Account	RC1 (Celebrities)	RC2 (Sports)	RC4 (Rap/R&B)	RC6 (Mainstream Influencers)	RC5 (Liberal Politicians)	RC3 (Traditional Comedy)	RC7 (Social Media Personalities)
thegreatkhalid	.42	.05	.40	.28	.16	.09	-.03
SportsCenter	.03	.88	.14	.14	.03	.01	.00
espn	.02	.86	.10	.15	.04	-.02	.00
BleacherReport	-.03	.86	.19	.05	-.01	-.02	.03
wojespn	-.02	.85	.13	-.01	-.03	.09	.02
KingJames	.02	.76	.28	.19	.11	.00	-.08
StephenCurry30	.07	.74	.27	.20	.09	.08	-.09
Dame_Lillard	.01	.71	.32	.03	.04	.02	.00
stephenasmith	.01	.71	.19	.04	.01	.18	.13
WorldWideWob	-.04	.68	.02	.00	-.06	.18	.11
elonmusk	-.13	.47	.14	.31	-.08	.33	-.01
BrotherNature	.12	.24	.12	.01	.16	.16	.21
asvpxrocky	-.04	.24	.72	.13	-.05	.11	.10
youngthug	-.12	.37	.71	.17	-.02	.09	.01
kendricklamar	.11	.35	.70	.08	.11	.10	.08

Table B2 Continued

Twitter Account	RC1 (Celebrities)	RC2 (Sports)	RC4 (Rap/R&B)	RC6 (Mainstream Influencers)	RC5 (Liberal Politicians)	RC3 (Traditional Comedy)	RC7 (Social Media Personalities)
tylerthecreator	.11	.03	.70	.05	.15	.08	.14
LILUZIVERT	-.17	.27	.66	.07	-.07	.05	.19
lilyachty	-.08	.20	.64	.08	-.02	.04	.16
sza	.33	-.10	.56	.16	.21	-.07	.08
trvisXX	-.13	.37	.55	.36	.00	.03	-.01
rihanna	.41	.01	.54	.31	.27	-.17	-.01
Drake	-.03	.43	.54	.33	.06	.06	-.04
KidCudi	.00	.31	.53	-.02	-.04	.28	.13
kanyewest	-.17	.29	.53	.27	-.06	.14	.14
chancetherapper	.23	.27	.50	.29	.08	.17	.07
DemetriusHarmon	.17	.08	.49	-.04	.03	.11	.17
theweeknd	.22	.18	.48	.24	.06	.20	-.03
jaden	.27	.05	.42	.11	.23	.21	.08
WORLDSTAR	-.15	.37	.39	.17	-.03	.30	.04
PostMalone	.25	.00	.33	.20	.05	.22	.13

Table B2 Continued

Twitter Account	RC1 (Celebrities)	RC2 (Sports)	RC4 (Rap/R&B)	RC6 (Mainstream Influencers)	RC5 (Liberal Politicians)	RC3 (Traditional Comedy)	RC7 (Social Media Personalities)
khloekardashian	.27	.09	.21	.74	.12	-.15	.09
KylieJenner	.21	.15	.27	.72	.09	-.19	.02
KendallJenner	.23	.15	.22	.70	.14	-.22	.07
KimKardashian	.21	.09	.32	.68	.14	-.16	.08
TheEllenShow	.18	.12	.01	.67	.01	.12	.01
justinbieber	.33	.17	.18	.59	-.02	.05	.02
DavidDobrik	.16	.23	.14	.57	-.08	.17	.09
jamescharles	.10	.02	-.05	.56	.05	.08	.21
tanamongeau	.11	-.02	.24	.54	.04	-.05	.30
chrissyteigen	.30	.12	.08	.53	.19	-.02	.04
jimmyfallon	.30	.28	.15	.52	.16	.24	-.01
colesprouse	.38	.02	.07	.48	.24	.26	-.04
BuzzFeed	.28	.09	.23	.44	.23	.19	-.07
shanedawson	-.01	-.04	-.02	.42	-.04	.33	.07
BarackObama	.21	.15	.14	.19	.78	.14	.01

Table B2 Continued

Twitter Account	RC1 (Celebrities)	RC2 (Sports)	RC4 (Rap/R&B)	RC6 (Mainstream Influencers)	RC5 (Liberal Politicians)	RC3 (Traditional Comedy)	RC7 (Social Media Personalities)
MichelleObama	.32	.05	.09	.15	.77	.06	.04
KamalaHarris	.18	.09	.05	.27	.76	.01	.00
SenSanders	.04	-.10	.04	-.09	.76	.11	.22
BernieSanders	.07	-.11	.06	-.09	.75	.06	.22
JoeBiden	.12	.14	-.03	.31	.74	.11	.07
POTUS44	.18	.16	.14	.16	.74	.20	.04
AOC	.23	-.12	.00	-.05	.71	.01	.16
johnkrasinski	.20	.14	.12	.15	.16	.51	.09
ericandre	-.15	.17	.28	-.03	.05	.51	.21
MrBeast	.03	.26	.06	.11	-.15	.49	.00
VancityReynolds	.28	.13	.06	.29	.24	.45	-.12
YourAnonCentral	-.04	.11	.13	-.08	.21	.44	.27
RobertDowneyJr	.31	.22	.15	.11	.19	.43	-.11
JordanPeele	.13	.10	.39	-.14	.28	.43	.07
archillect	.05	.32	.09	-.03	-.05	.42	.06

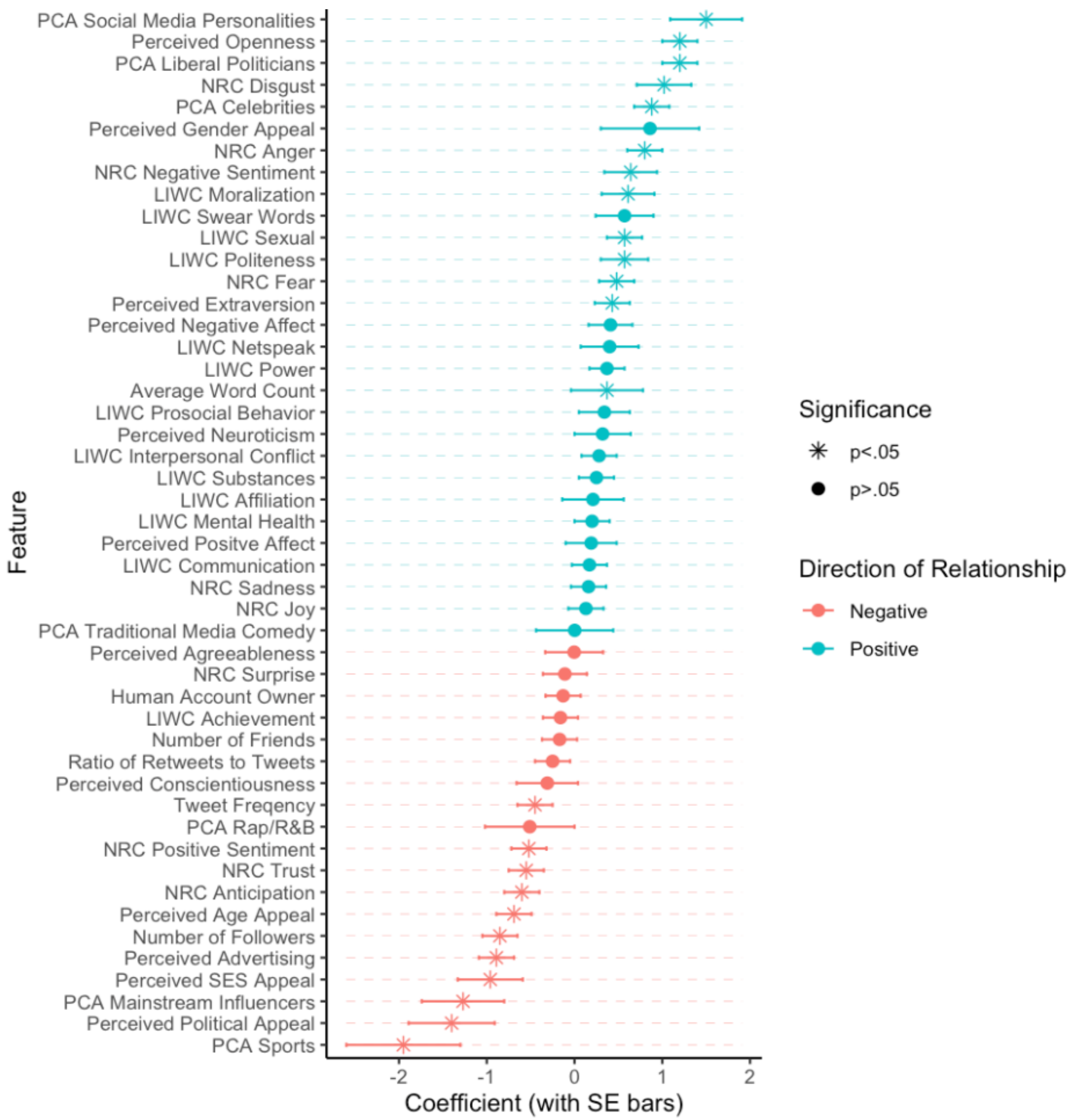
Sethrogen	.08	.23	.39	.13	.16	.39	.21
Lin_Manuel	.34	.07	.03	.21	.24	.39	.06
TheOnion	-.06	-.01	-.08	-.22	.28	.38	.35
netflix	.22	.03	.13	.36	.11	.38	-.07
dog_rates	.15	-.16	.01	.00	.25	.36	.21
richbrian	.24	-.03	.17	-.11	.03	.33	.16
snitchery	.12	-.17	.01	-.16	.04	.32	.08
thenoelmiller	.19	.04	.17	.07	.13	.05	.66
CaucasianJames	.03	.02	.08	.08	.12	.13	.62
codyko	.26	.02	.21	.13	.08	-.01	.59
caseykfrey	.06	.05	.21	.10	.06	.14	.57
Nick_Colletti	.06	.25	.17	.15	.08	.26	.50
SarahBaska	.31	-.14	.12	.28	.16	.04	.42
quenblackwell	.30	-.05	.34	.14	.28	-.13	.37

APPENDIX C

STUDY 2 GENDER AND NEUROTICISM

Figure C1

Moderating Effect of Gender-centered Neuroticism on the Relationship Between Z-scored Twitter Profile Features and Study 2 Account Interest Scores



APPENDIX D

STUDY 2 FEATURES INTERCORRELATIONS

Table D1

Top Positive Intercorrelations among Twitter Profile Features

Feature 1	Feature 2	r
Perceived Positive Affect	Perceived Agreeableness	.78
PCA Liberal Politicians	Average Word Count	.73
Swear Words	Netspeak	.72
PCA Liberal Politicians	Power Words	.66
Average Word Count	Power Words	.65
Perceived Negative Affect	Perceived Neuroticism	.62
Prosocial Words	Polite Words	.62
PCA Celebrities	Perceived Gender Appeal	.60
Power Words	Moral Words	.60
Positive Sentiment	Joy	.60

Table D2

Top Negative Intercorrelations among Twitter Profile Features

Feature 1	Feature 2	r
Perceived Neuroticism	Perceived Conscientiousness	-.78
Perceived Negative Affect	Perceived Agreeableness	-.73
Perceived Neuroticism	Perceived Agreeableness	-.71
PCA Liberal Politicians	Perceived Political Appeal	-.69
PCA Sports	Perceived Gender Appeal	-.65
Perceived Positive Affect	Perceived Neuroticism	-.64
Perceived Age Appeal	Perceived Openness	-.60
PCA Sports	Perceived Openness	-.60
Power Words	Perceived SES Appeal	-.58
Perceived Positive Affect	Perceived Negative Affect	-.58

REFERENCES CITED

Allport, G. W. (1937). *Personality: A psychological interpretation* (pp. xiv, 588). Holt.

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology*, *84*(5), 1054. https://doi.org/10.1037/0022-3514.84.5.1054

Androutsopoulos, J. (2011). Language change and digital media: a review of conceptions and evidence. *Standard languages and language standards in a changing Europe*, *1*, 145-159.

Arnoux, P.-H., Xu, A., Boyette, N., Mahmud, J., Akkiraju, R., & Sinha, V. (2017). 25 Tweets to Know You: A New Model to Predict Personality with Social Media. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1), 472–475.

Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, *124*, 150–159. https://doi.org/10.1016/j.paid.2017.12.018

Baldwin, D., & Meyer, M. (2007). How inherently social is language? In *Blackwell handbook of language development* (pp. 87–106). Blackwell Publishing. https://doi.org/10.1002/9780470757833.ch5

Becker, J., Chan, C., Chan, G. C., Leeper, T. J., Gandrud, C., MacDonald, A., Zahn, I., Stadlmann, S., Williamson, R., Kennedy, P., Price, R., Davis, T. L., Day, N., Denney, B., & Bokov, A. (2021). *rio: A Swiss-Army Knife for Data I/O* (0.5.29). https://CRAN.R-project.org/package=rio

Bedford-Petersen, C., & Weston, S. J. (2021). Mapping Individual Differences on the Internet: Case Study of the Type 1 Diabetes Community. *JMIR Diabetes*, *6*(4), e30756. https://doi.org/10.2196/30756

Bengtsson, H. (2022). *future: Unified Parallel and Distributed Processing in R for Everyone* (1.27.0). https://CRAN.R-project.org/package=future

Berenbaum, H., Bredemeier, K., & Thompson, R. J. (2008). Intolerance of uncertainty: Exploring its dimensionality and associations with need for cognitive closure, psychopathology, and personality. *Journal of anxiety disorders*, *22*(1), 117-125.

Bjornsdottir, R. T., & Rule, N. (2017). The Visibility of Social Class From Facial Cues. *Journal of Personality and Social Psychology*, *113*, 530–546. https://doi.org/10.1037/pspa0000091

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Boyce, P., Parker, G., Barnett, B., Cooney, M., & Smith, F. (1991). Personality as a Vulnerability Factor to Depression. *British Journal of Psychiatry*, *159*(1), 106–114. https://doi.org/10.1192/bjp.159.1.106

Boyd, R., Ashokkumar, A., Seraj, S., & Pennebaker, J. (2022). *The Development and Psychometric Properties of LIWC-22*. https://doi.org/10.13140/RG.2.2.23890.43205

Bresin, K., & Robinson, M. D. (2015). You Are What You See and Choose: Agreeableness and Situation Selection. *Journal of Personality*, *83*(4), 452–463. https://doi.org/10.1111/jopy.12121

Brewin, C. R., Rose, S., Andrews, B., Green, J., Tata, P., McEvedy, C., Turner, S., & Foa, E. B. (2002). Brief screening instrument for post-traumatic stress disorder. *The British Journal of Psychiatry*, *181*(2), 158–162. https://doi.org/10.1192/bjp.181.2.158

Bruner, J., & Tagiuri, R. (1954). The perception of people. In *The handbook of social psychology* (1st ed., Vol. 2, pp. 634–654). Addison-Wesley.

Buss, D. M. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology*, *53*(6), 1214–1221. https://doi.org/10.1037/0022-3514.53.6.1214

Cohen, J. (1982). Set Correlation As A General Multivariate Data-Analytic Method. *Multivariate Behavioral Research*, *17*(3), 301–341. https://doi.org/10.1207/s15327906mbr1703_2

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The Problem of Units and the Circumstance for POMP. *Multivariate Behavioral Research*, *34*(3), 315–346. https://doi.org/10.1207/S15327906MBR3403_2

Cohen, S., Mermelstein, R., Kamarck, T., & Hoberman, H. M. (1985). Measuring the Functional Components of Social Support. In I. G. Sarason & B. R. Sarason (Eds.), *Social Support: Theory, Research and Applications* (pp. 73–94). Springer Netherlands. https://doi.org/10.1007/978-94-009-5115-0_5

Costello, C. K. (2020). *Human and Computerized Personality Inferences from Digital Footprints on Twitter* [Ph.D., University of Oregon]. https://www.proquest.com/docview/2457729886/abstract/DE63DDACC16C41DBPQ/1

Costello, C. K., Bedford-Petersen, C., & Srivastava, S. (2022). *You are Who You Follow: Personality is Revealed in the Accounts we Follow on Twitter*.

Costello, C., Srivastava, S., Rejaie, R., & Zalewski, M. (2021). Predicting Mental Health From Followed Accounts on Twitter. *Collabra: Psychology*, *7*(1), 18731. https://doi.org/10.1525/collabra.18731

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2. ed). Univ. of Illinois Press.

DesJardins, N. M. L., Srivastava, S., Küfner, A. C. P., & Back, M. D. (2015). Who Attains Status? Similarities and Differences Across Social Contexts. *Social Psychological and Personality Science*, *6*(6), 692–700. https://doi.org/10.1177/1948550615580171

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Wirtz, D., Biswas-Diener, R., Tov, W., Kim-Prieto, C., Choi, D.-W., & Oishi, S. (2009). New Measures of Well-Being. In *Social Indicators Research Series* (Vol. 39, pp. 247–266). https://doi.org/10.1007/978-90-481-2354-4_12

Dimson, T. (2015). *Emojineering Part 1: Machine Learning for Emoji Trends*. Medium. https://instagram-engineering.com/emojineering-part-1-machine-learning-for-emoji-trendsmachine-learning-for-emoji-trends-7f5f9cb979ad

Duggan, C. F., Lee, A. S., & Murray, R. M. (1990). Does Personality Predict Long-Term Outcome in Depression? *British Journal of Psychiatry*, *157*(1), 19–24. https://doi.org/10.1192/bjp.157.1.19

Feiler, D. C., & Kleinbaum, A. M. (2015). Popularity, Similarity, and the Network Extraversion Bias. *Psychological Science*, *26*(5), 593–603.

Fellows, I. (2018). *wordcloud: Word Clouds* (2.6). https://CRAN.R-project.org/package=wordcloud

Firke, S., Denney, B., Haid, C., Knight, R., Grosser, M., & Zadra, J. (2021). *janitor: Simple Tools for Examining and Cleaning Dirty Data* (2.1.0). https://CRAN.R-project.org/package=janitor

Fox, J., Weisberg, S., Price, B., Friendly, M., Hong, J., Andersen, R., Firth, D., Taylor, S., & R Core Team. (2022). *effects: Effect Displays for Linear, Generalized Linear, and Other Models* (4.2-2). https://CRAN.R-project.org/package=effects

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 652–670.

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Furnham, A. (1981). Personality and activity preference. *British Journal of Social Psychology*, *20*(1), 57–68. https://doi.org/10.1111/j.2044-8309.1981.tb00474.x

Gamer, M., Lemon, J., & Singh, I. F. P. (2019). *irr: Various Coefficients of Interrater Reliability and Agreement* (0.84.1). https://CRAN.R-project.org/package=irr

Gill, A. J., & Oberlander, J. (2019). Taking Care of the Linguistic Features of Extraversion. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society* (1st ed., pp. 363–368). Routledge. https://doi.org/10.4324/9781315782379-99

Gohel, D., Jager, C., Fazilleau, Q., output), M. N. (rmarkdown for docx, Robert, T., footnotes), M. B. (inline, reference), A. Y. (support for bookdown cross, objects), P. J. (support for gam, & system), S. B. (work on footnote positioning. (2022). *flextable: Functions for Tabular Reporting* (0.7.2). https://CRAN.R-project.org/package=flextable

Golbeck, J., Robles, C., & Turner, K. (2011). *Predicting personality with social media*. 10.

Gomez, R., Cooper, A., & Gomez, A. (2000). Susceptibility to positive and negative mood states: Test of Eysenck's, Gray's and Newman's theories. *Personality and Individual Differences*, *29*(2), 351-365.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*(3), 379–398. https://doi.org/10.1037/0022-3514.82.3.379

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Gross, J. J. (1998). The Emerging Field of Emotion Regulation: An Integrative Review. *Review of General Psychology*, 29.

Grün, B., Hornik, K., CTM), D. M. B. (VEM estimation of L. and, CTM), J. D. L. (VEM estimation of, LDA), X.-H. P. (MCMC estimation of, RNG), M. M. (Mersenne T., RNG), T. N. (Mersenne T., & RNG), S. C. (Mersenne T. (2021). *topicmodels: Topic Models* (0.2-12). https://CRAN.R-project.org/package=topicmodels

Hall, J. A., Mast, M. S., & West, T. V. (2016). *The Social Psychology of Perceiving Others Accurately*. Cambridge University Press.

Hall, J. A., Pennington, N., & Lueders, A. (2014). Impression management and formation on Facebook: A lens model approach. *New Media & Society*, *16*(6), 958–982. https://doi.org/10.1177/1461444813495166

Harms, P. D. (2020). Corresponsive Principle of Personality Development. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 919–921). Springer International Publishing. https://doi.org/10.1007/978-3-319-24612-3_1855

Harrell Jr, F. E. (2020). *Hmisc: Harrell Miscellaneous*. https://CRAN.R-project.org/package=Hmisc

Hester, J., Bryan, J., & RStudio. (2022). *glue: Interpreted String Literals* (1.6.2). https://CRAN.R-project.org/package=glue

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. *Journal of Computer-Mediated Communication*, *18*(2), 40–60. https://doi.org/10.1111/jcc4.12001

Ickes, W., Snyder, M., & Garcia, S. (1997). *Personality influences on the choice of situations*. https://doi.org/10.1016/B978-012134645-4/50008-1

Izrailev, S. (2021). *tictoc: Functions for Timing R Scripts, as Well as Implementations of Stack and List Structures* (1.0.1). https://CRAN.R-project.org/package=tictoc

Jennrich, C. B. and R. (2022). *GPArotation: GPA Factor Rotation* (2022.4-1). https://CRAN.R-project.org/package=GPArotation

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339. https://doi.org/10.1037/0033-2909.129.3.339

Kearney, M. (2019). rtweet: Collecting and analyzing Twitter data. *Journal of Open Source Software*, *4*(42), 1829. https://doi.org/10.21105/joss.01829

Kearney, M. W., Sancho, L. R., Wickham, H., Heiss, A., Briatte, F., & Sidi, J. (2022). *rtweet: Collecting Twitter Data* (1.0.2). https://CRAN.R-project.org/package=rtweet

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Kraus, M. W., Torrez, B., Park, J. W., & Ghayebi, F. (2019). Evidence for the reproduction of social class in brief speech. *Proceedings of the National Academy of Sciences*, *116*(46), 22998–23003. https://doi.org/10.1073/pnas.1900500116

Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.* https://www.tidymodels.org

Kuhn, M., Wickham, H., & RStudio. (2022). *tidymodels: Easily Install and Load the "Tidymodels" Packages* (1.0.0). https://CRAN.R-project.org/package=tidymodels

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & Jensen, S. P. (2020). *lmerTest: Tests in Linear Mixed Effects Models* (3.1-3). https://CRAN.R-project.org/package=lmerTest

Lambiotte, R., & Kosinski, M. (2014). Tracking the Digital Footprints of Personality. *Proceedings of the IEEE*, *102*(12), 1934–1939. https://doi.org/10.1109/JPROC.2014.2359054

Liege), G. R. (Universite du Q. a M. and D. M. (Universite de. (2020). *nFactors: Parallel Analysis and Other Non Graphical Solutions to the Cattell Scree Test* (2.4.1). https://CRAN.R-project.org/package=nFactors

Maltby, J., Houran, J., & McCutcheon, L. E. (2003). A clinical interpretation of attitudes and behaviors associated with celebrity worship. *The Journal of nervous and mental disease*, *191*(1), 25-29.

Maltby, J., McCutcheon, L. E., & Lowinger, R. J. (2011). Brief report: Celebrity worshipers and the five-factor model of personality. *North American Journal of Psychology*, *13*(2), 343-348.

Macrae, C. N., & Martin, D. (2007). A boy primed Sue: Feature-based processing and person construal. *European Journal of Social Psychology, 37*(5), 793–805. https://doi.org/10.1002/ejsp.406

McClain, C., Widjaya, R., Rivero, G., & Smith, A. (2021, November 15). The Behaviors and Attitudes of U.S. Adults on Twitter. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2021/11/15/the-behaviors-and-attitudes-of-u-s-adults-on-twitter/

McCrae, R. R., & Löckenhoff, C. E. (2010). Self-regulation and the five-factor model of personality traits. In *Handbook of personality and self-regulation* (pp. 145–168). Wiley-Blackwell. https://doi.org/10.1002/9781444318111.ch7

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30. https://doi.org/10.1037/1082-989X.1.1.30

McManus, I. C., & Furnham, A. (2006). Aesthetic activities and aesthetic attitudes: Influences of education, background and personality on interest and involvement in the arts. *British Journal of Psychology*, *97*(4), 555–587. https://doi.org/10.1348/000712606X101088

Mischel, W. (1968). Personality and Assessment.(1968). *DJ Bern, Beliefs, Attitudes and Human Affairs*, 41–78.

Mohammad, S. M., & Kiritchenko, S. (2015). Using Hashtags to Capture Fine Emotion Categories from Tweets. *Computational Intelligence*, *31*(2), 301–326. https://doi.org/10.1111/coin.12024

Moritz, S., Gatscha, S., & Wang, E. (2021). *imputeTS: Time Series Missing Value Imputation* (3.2). https://CRAN.R-project.org/package=imputeTS

Motamedi, R., Rezayi, S., Rejaie, R., Light, R., & Willinger, W. (2016). Characterizing Twitter Elite Communities: Measurement, Characterization, and Implications.

Müller, K., & Bryan, J. (2020). *here: A Simpler Way to Find Your Files* (1.0.1). https://CRAN.R-project.org/package=here

Mulyanegara, R. C., Tsarenko, Y., & Anderson, A. (2007). The Big Five and brand personality: Investigating the impact of consumer personality on preferences towards particular brand personality. *Journal of Brand Management*, *16*(4), 234–247. https://doi.org/10.1057/palgrave.bm.2550093

Muris, P., Roelofs, J., Rassin, E., Franken, I., & Mayer, B. (2005). Mediating effects of rumination and worry on the links between neuroticism, anxiety and depression. *Personality and Individual Differences*, *39*(6), 1105–1111. https://doi.org/10.1016/j.paid.2005.04.005

Naragon-Gainey, K., & Watson, D. (2011). Clarifying the dispositional basis of social anxiety: A hierarchical perspective. *Personality and Individual Differences*, *50*(7), 926–934. https://doi.org/10.1016/j.paid.2010.07.012

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer Palettes* (1.1-3). https://CRAN.R-project.org/package=RColorBrewer

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the Prediction of Consequential Outcomes. *Annual Review of Psychology*, *57*(1), 401–421. https://doi.org/10.1146/annurev.psych.57.102904.190127

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., & Seligman, M. E. P. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology*, *108*(6), 934–952. https://doi.org/10.1037/pspp0000020

Pew Research Center. (2021). *Social Media Use in 2021 | Pew Research Center*. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., & Cella, D. (2011). Item Banks for Measuring Emotional Distress From the Patient-Reported Outcomes Measurement Information System (PROMIS®): Depression, Anxiety, and Anger. *Assessment*, *18*(3), 263–283. https://doi.org/10.1177/1073191111411667

Probst, P., & Boulesteix, A.-L. (2017). To Tune or Not to Tune the Number of Trees in Random Forest. *The Journal of Machine Learning Research*, *18*(1), 6673–6690.

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. https://doi.org/10.1016/j.jrp.2012.08.008

Queiroz, G. D., Fay, C., Hvitfeldt, E., Keyes, O., Misra, K., Mastny, T., Erickson, J., Robinson, D., Silge [aut, J., & cre. (2022). *tidytext: Text Mining using "dplyr", "ggplot2", and Other Tidy Tools* (0.3.3). https://CRAN.R-project.org/package=tidytext

Quercia, D., Kosinski, M., Stillwell, D., & Crowcroft, J. (2011). Our Twitter Profiles, Our Selves: Predicting Personality with Twitter. *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 180–185. https://doi.org/10.1109/PASSAT/SocialCom.2011.26

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*(6), 1236–1256. https://doi.org/10.1037/0022-3514.84.6.1236

Revelle, W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University. https://CRAN.R-project.org/package=psych

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*(3), 582–593. https://doi.org/10.1037/0022-3514.84.3.582

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Roberts, B. W., & Nickel, L. B. (2021). Personality development across the life course: A neo socioanalytic perspective. In *Handbook of personality: Theory and research, 4th ed* (pp. 259–283). The Guilford Press.

Roberts, B. W., & Robins, R. W. (2000). Broad Dispositions, Broad Aspirations: The Intersection of Personality Traits and Major Life Goals. *Personality and Social Psychology Bulletin*, *26*(10), 1284–1296. https://doi.org/10.1177/0146167200262009

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

Rodriguez, A. J., Holleran, S. E., & Mehl, M. R. (2010). Reading Between the Lines: The Lay Assessment of Subclinical Depression From Written Self-Descriptions. *Journal of Personality*, *78*(2), 575–598. https://doi.org/10.1111/j.1467-6494.2010.00627.x

Rosenfeld, L., & Plax, T. (1977). Clothing as Communication. *Journal of Communication*, *27*, 24–31. https://doi.org/10.1111/j.1460-2466.1977.tb01823.x

Rozin, P. (2009). What Kind of Empirical Research Should We Publish, Fund, and Reward?: A Different Perspective. *Perspectives on Psychological Science*, *4*(4), 435–439. https://doi.org/10.1111/j.1745-6924.2009.01151.x

RStudio Team. (2022). *RStudio: Integrated Development for R* (2022.02.2+485). RStudio PBC. http://www.rstudio.com/

Saklofske, D. H., Kelly, I. W., & Janzen, B. L. (1995). Neuroticism, depression, and depression proneness. *Personality and Individual Differences*, *18*(1), 27–31. https://doi.org/10.1016/0191-8869(94)00128-F

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of personality and social psychology*, *94*(1), 168.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., & Ungar, L. H. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE*, *8*(9), e73791. https://doi.org/10.1371/journal.pone.0073791

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.

Sharma, R. S. (1980). Clothing behaviour, personality, and values: A correlational study. *Psychological Studies*, *25*(2), 137–142.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

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), 117–143. https://doi.org/10.1037/pspp0000096

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330–348. https://doi.org/10.1037/a0021717

Tang, C., Garreau, D., & von Luxburg, U. (2018). When do random forests fail? *Advances in Neural Information Processing Systems*, *31*.

Thalmayer, A. G., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of Brief to Medium-Length Big Five and Big Six Personality Questionnaires. *Psychological Assessment*, *23*(4), 995–1009. https://doi.org/10.1037/a0024165

Tokar, D. M., Fischer, A. R., & Mezydlo Subich, L. (1998). Personality and Vocational Behavior: A Selective Review of the Literature, 1993–1997. *Journal of Vocational Behavior*, *53*(2), 115–153. https://doi.org/10.1006/jvbe.1998.1660

Tomasello, M. (2010). *Origins of Human Communication*. MIT Press.

Tracey, T. J. G., Ryan, J. M., & Jaschik-Herman, B. (2001). Complementarity of Interpersonal Circumplex Traits. *Personality and Social Psychology Bulletin*, *27*(7), 786–797. https://doi.org/10.1177/0146167201277002

Trapnell, P., & Paulhus, D. (2012). Agentic and Communal Values: Their Scope and Measurement. *Journal of Personality Assessment*, *94*, 39–52. https://doi.org/10.1080/00223891.2011.627968

Tskhay, K. O., & Rule, N. O. (2014). Perceptions of personality in text-based media and OSN: A meta-analysis. *Journal of Research in Personality*, *49*, 25–30. https://doi.org/10.1016/j.jrp.2013.12.004

Vaughan, D., Dancho, M., & RStudio. (2022). *furrr: Apply Mapping Functions in Parallel using Futures* (0.3.0). https://CRAN.R-project.org/package=furrr

Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality Impressions Based on Personal Websites. *Journal of Personality and Social Psychology*, *87*(1), 123–132. https://doi.org/10.1037/0022-3514.87.1.123

Waggoner, A. S., Smith, E. R., & Collins, E. C. (2009). Person perception by active versus passive perceivers. *Journal of Experimental Social Psychology*, *45*(4), 1028–1031. https://doi.org/10.1016/j.jesp.2009.04.026

Watson, D., Hubbard, B., & Wiese, D. (2000). Self–other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, *78*(3), 546. https://doi.org/10.1037/0022-3514.78.3.546

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., Zemla, J., Freidank, M., Cai, J., & Protivinsky, T. (2021). *corrplot: Visualization of a Correlation Matrix* (0.92). https://CRAN.R-project.org/package=corrplot

Whiting, A., & Williams, D. (2013). Why people use social media: A uses and gratifications approach. *Qualitative Market Research: An International Journal*, *16*(4), 362–369. https://doi.org/10.1108/QMR-06-2013-0041

Wickham, H. (2020). *reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package* (1.4.4). https://CRAN.R-project.org/package=reshape2

Wickham, H. (2022). *tidyverse: Easily Install and Load the "Tidyverse"* (1.3.2). https://CRAN.R-project.org/package=tidyverse

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, *112*(4), 1036–1040. https://doi.org/10.1073/pnas.1418680112