

Examining the language of solitude versus loneliness in tweets

Journal of Social and
Personal Relationships
1–15

© The Author(s) 2021




Article reuse guidelines:

sagepub.com/journals-permissions

DOI: [10.1177/0265407521998460](https://doi.org/10.1177/0265407521998460)

journals.sagepub.com/home/spr



Will E. Hipson¹ , Svetlana Kiritchenko²,
Saif M. Mohammad², and Robert J. Coplan¹

Abstract

People have widely different experiences of being alone. Sometimes being alone is relaxing and restorative other times it gives way to feelings of loneliness. Researchers conceptually distinguish between *solitude*, which tends to be viewed more positively, and *loneliness*, which is more negative. However, it is unclear whether these terms are used differently in everyday language. We sought to compare the emotional content of over 19 million tweets containing the terms *solitude* and *lonely/loneliness*. Using a computational linguistics approach, we found that *solitude* tends to be used in more positive and less emotionally activated (i.e., lower arousal) contexts compared to *lonely*. We also found that the word *alone* tends to be used somewhat differently from *solitude* and *lonely*. These results have implications both for how we understand different experiences of time alone in general and for what kind of language we should use when discussing these experiences.

Keywords

Emotion, language, loneliness, solitude, text analysis

Much of our waking day (approximately 33% in adults and up to 50% in older adults) is spent in some form of solitude (Shelley, 2005). Whether it is hiking alone in the woods, sitting by oneself on the bus, or savoring a brief 15 minutes of “quiet time” during a busy

¹ Carleton University, Canada

² National Research Council-Canada, Canada

Corresponding author:

Will E. Hipson, Department of Psychology, Carleton University, 1125 Colonel By Drive, 550 Loeb Building, Ottawa, Ontario K1S 5B6, Canada.

Email: william.hipson@carleton.ca

workday, solitude features prominently in our day-to-day experiences. However, we also have very different experiences of being alone. The psychological study of solitude demonstrates that our experience of being alone can range from blissful and rejuvenating to bleak and dispiriting (Hawkey & Cacioppo, 2010; Long et al., 2003; Nguyen et al., 2018).

We know that spending some time alone offers a welcome respite from extended periods of social interaction, yet conversely, spending too much time alone is a risk factor for loneliness and depression (Hawkey & Cacioppo, 2010). Indeed, some have invoked the idea of a “Goldilocks Hypothesis” for solitude, such that for each person there exists an optimal amount of solitude and spending too much or too little time alone can be cause for dissatisfaction (Coplan et al., 2019; Larson, 1990).

Taken together, people spend time alone for many different reasons and some experiences of being alone are more conducive, and others less conducive to positive emotions. At the most basic level though, researchers distinguish between *solitude* and *loneliness* (Coplan & Bowker, 2014), where the former often describes an intrinsic desire to be alone, while the latter is a dissatisfaction with being alone. Notwithstanding, it is a different question as to whether these words are understood to denote different phenomena in everyday usage. That is, when someone uses the word *solitude* is it in a more positive context than the word *lonely*? We sought to explore this question by analyzing tweets containing these words and comparing the affective properties (e.g., valence, arousal) of neighboring words. Differences in the affective “climate” of words that accompany these target constructs offer us a unique window into how everyday language reflects our understanding of these phenomena. Answering this question would not only give us a deeper understanding of how people view different solitary experiences, but could help in wording questionnaires about time alone, or in developing programs to identify people at risk of severe loneliness.

Distinguishing among alone, lonely, and solitude

We distinguish between the concepts *alone*, *lonely*, and *solitude* as denoting different motivations toward and experiences of solitude (Galanaki, 2004). *Alone* is construed as a state in which one is removed from social interaction, *lonely* (loneliness) is a dissatisfaction with the quality of our social interactions and *solitude* denotes a positive state of voluntary aloneness (Coplan et al., 2018; Galanaki, 2004; Hawkey & Cacioppo, 2010).

In the psychological literature, these concepts are conceived as distinct but sharing some conceptual overlap. The concept *alone* (or more formally *aloneness*) refers to a state of social separation relative to others (Goffman, 1971; Larson, 1990). It is not entirely clear whether this should be confined to situations in which one is physically apart from others or whether it makes sense to include instances in which we are around other people but not interacting (Larson, 1990). Nor do we know whether to categorize computer-mediated communication in the physical absence of other people as something other than solitude (Wang et al., 2012). Despite these questions, we ascribe to the rather convenient view that *alone* simply refers to a neutral state of physical separation from other people (Goffman, 1971).

However, everyday language is rife with ambiguity. The phrase “I am sitting alone right now” matches the conceptual definition of alone as a neutral state, but the phrase “I lost a friend today and I feel alone” uses *alone* to describe a negative feeling in the same way that one would use the word *lonely*. Further complicating matters, *alone* may even be used for hyperbole or emphasis, as in the phrase “Evolution alone is the most important discovery in biology.” Thus, although we conceptualize *alone* as a neutral state, but acknowledge that its meaning depends greatly on the context in which it is used.

In contrast, *lonely* and *solitude* carry with them somewhat clearer affective meanings. Researchers ascribe to *lonely* and *loneliness* a negative feeling arising from a perceived lack of affiliation and closeness (Hawkley & Cacioppo, 2010). Loneliness is a dissatisfaction with the quality of one’s social relationships, and therefore may occur even when one is physically surrounded by others (Rokach, 2004). Moreover, loneliness is recognized as a major mental and physical health concern (Qualter et al., 2015), and some have gone so far as to claim that we are in the midst of a *loneliness epidemic* (Bergland, 2015). Governments are increasingly taking concrete steps to combat loneliness, such as the United Kingdom’s appointment of a Minister for Loneliness (Cox, 2017). On the whole, we see that the terms *lonely* and *loneliness* refer uniquely to the negative emotional state arising from perceived deficiency in one’s social relationships.

Finally, we turn to the term *solitude*. Perhaps most commonly, we might think of solitude as a more whimsical alternative to *alone*, whereby it describes a state of separation from others, but casts it in a more positive light. For instance, researchers who study solitude often use it in the context of being alone in nature (Korpela et al., 2001) or seeking out time alone for leisure and restoration (Lay et al., 2019). Similarly, from a motivational perspective, many researchers either explicitly or implicitly use *solitude* to refer to an intrinsic motivation to be alone (Coplan et al., 2019; Nguyen et al., 2019). However, this is not an established convention and *solitude* is frequently used interchangeably with *alone*. This is partly it is unwieldy to say “the psychological study of alone.” Notwithstanding, even if we substitute *alone* for *aleness*, the theorized motivational and emotional underpinnings remain the same.

Thus, researchers often use these terms to carve the universal experience of aloneness at its joints. But there is scant evidence as to whether people use these words to differentiate “positive” and “negative” solitary experiences. For example, it would be useful to know whether people generally use the word *lonely* in more negative contexts than *alone* or *solitude*. Previous research has found that social context influences the kind of words that people use. For instance, people induced to be in an “alone” state of mind were more likely to use words related to the past or future as opposed to the present (Uziel, 2021). However, it remains unclear how different solitary concepts relate to people’s usage of emotion words.

This research has both conceptual and practical applications in psychology and natural language processing. For one, we can be more confident in our conceptual distinction between *solitude* and *lonely* if there is evidence that people use these words differently. As well, having a clearer understanding of how these words are used in everyday language can improve assessments of these constructs that rely heavily on language (e.g., self-report measures) and also help clinicians interpret language used by

their clients. It may be possible to extend this work toward using natural language processing to identify people who may be at risk for suffering the negative consequences of loneliness. Although this research is a long way from using language to detect loneliness, we sought to take initial steps in understanding the affective context in which different solitary terms are used.

Exploring the affective context of words

People communicate their thoughts, feelings, and emotions through language (Lindquist et al., 2015), and psychology researchers are increasingly turning to natural language to gain insights into these thoughts and emotions. Osgood et al. (1957) were the first to map the semantic aspects of language along three dimensions: good vs. bad (valence), active vs. inactive (arousal), and strong vs. weak (dominance). Russell and colleagues (Mehrabian & Russell, 1974; Russell, 2003) would later use these dimensions (focusing predominately on valence and arousal) to describe the entire range of subjective feeling. Although most researchers continue to describe affect along valence and arousal, linguistic approaches have often relied on a three-dimensional view, which includes dominance (Fontaine et al., 2007; Warriner et al., 2013).

Under the assumption that aspects of cognition and emotion are reflected in language, Pennebaker and colleagues (Pennebaker & King, 1999; Pennebaker et al., 2001; Rude et al., 2004; Tausczik & Pennebaker, 2010) spearheaded the use of computerized text analysis programs applicable to research in psychology. More recently, larger lexicons with tens of thousands of words have been developed using various combinations of crowdsourcing and machine learning techniques (Mohammad, 2018; Mohammad & Turney, 2010; Tang et al., 2014). These lexicons are specific to emotion with some developed to address discrete emotions (e.g., anger, sadness) (Aman & Szpakowicz, 2007; Mohammad & Turney, 2010) and others to address affect dimensions (e.g., valence, arousal) (Mohammad, 2018; Warriner et al., 2013). Researchers in computational linguistics and psychological disciplines have used these openly available lexicons to examine questions such as how emotions change at the *micro* timescale (Kross et al., 2019) and how emotions change at the *macro*, developmental timescale (Hipson, 2019), as well as such varied topics as attitudes toward products, politicians, and professional sports teams (Gratch et al., 2015; Maynard & Funk, 2011; Zhang et al., 2010).

Here, our focus is on words used denote solitary experiences. Specifically, we want to know whether the terms *lonely*, *alone*, and *solitude* arise in different emotional contexts. We turned to tweets to obtain a vast quantity of text containing specified key terms. Twitter is in many ways an ideal source of textual data because tweets are public domain and relatively concise (280 character limit). As well, Twitter is a naturalistic context and thus may be highly representative of how people use language in everyday life.

We can measure the emotional context of solitary terms by examining the words that tend to co-occur with each of the terms across a massive collection of tweets. Presumably, if a word such as *lonely* occurs in more negative contexts it will co-occur with lower valence words, such as *sad* or *depressed*. Similarly, we can look at the arousal rating of co-occurring terms for evidence of whether *solitude* is used in more deactivated (e.g., calm, bored, sleepy) contexts relative to *lonely* which we would expect to arise in

more activated (e.g., stressed, tense) contexts (Hawkey & Cacioppo, 2010; Nguyen et al., 2018; Seeman, 1996). Moreover, dominance may provide some insight into motivations as it reflects the extent to which one feels in control versus out of control.

Beyond affect dimensions, there are also more specific emotion concepts, such as anger, fear, joy, and sadness. The emotion literature is strongly divided as to whether concepts like anger and joy refer to physiologically distinct processes (as is argued by proponents of basic emotions theory; see Ekman & Cordaro, 2011), or are constructed by the brain to help organize and understand events (Barrett, 2013). A conciliatory view suggests that dimensions like valence and arousal capture one's underlying mood or *affect*, whereas discrete emotions organize behavior, cognition, and affect during unambiguous emotional events (see Ortony et al., 1990). For example, people use terms like *frustrated*, *embarrassed*, and *exuberant* to interpret and convey more specific emotional situations. Thus, at least from a linguistics perspective it makes sense to examine words that relate to concepts such as anger and sadness as well as affect dimensions.

The current study

The goal of the current study was to explore how the terms *lonely*, *alone*, and *solitude* are used in online language. We focused especially on contrasting *solitude* with *lonely* but included *alone* as a more neutral baseline. We analyzed which words are most likely to co-occur with these terms and then assessed the sentiment (e.g., valence, arousal, dominance, anger, etc.) of these words to get a sense of whether different solitude terms arise in different emotional contexts. Although this research was largely exploratory, we generally expected *solitude* to occur in the context of higher valence, lower arousal, and higher dominance terms compared to *lonely* because *solitude* is theorized to reflect a positive, intrinsic motivation toward solitude. We had fewer expectations for comparisons with *alone*, but given that it supposedly denotes a neutral state, it seemed likely that it would populate the middle ground between *solitude* and *lonely*. We also did not have specific hypotheses for discrete emotions (e.g., anger), but instead, included these as wholly exploratory variables.

Method

Corpus

We obtained 19,277,359 tweets from the Twitter API containing any of the following key terms: “solitude,” “lonely” (“loneliness,” “lonesomeness”), and “alone” (“aloneness”) from August 2018 to July 2019. We refer to this corpus as the SOLO (*State of Being Alone*) corpus (for more details see Kiritchenko et al., 2020). We discarded duplicate tweets, tweets with fewer than three words, tweets containing external URLs, and kept up to the first three tweets from the same user (to reduce influence of frequent Tweeters). Twitter does not provide demographics about its users through its API, but a Pew Survey (Pew Research Center, 2019) reported that 32% of users were between the ages of 13–17, 38% were between 18–29, 26% were between 30–49, 17% were between 50–64, and 7% were 65+.

Table 1 shows a breakdown of the number of tweets containing the relevant key terms. Not surprisingly, *alone* occurred more frequently than *lonely* and *solitude*, probably

Table 1. Number of tweets and users by term.

	# of tweets	# of users
Solitude	191,643	158,878
Lonely	3,829,399	2,685,596
Alone	15,256,317	9,094,772
Alone (filtered by phrases)	4,238,746	3,061,916

because *alone* occurs in a wider variety of contexts and meanings. To ensure that the tweets in our sample were relevant to the topic of solitude, we manually checked 100 tweets within each term. We concluded that 92% of *solitude* tweets and 97% of *lonely* tweets were relevant; however, only 57% of *alone* tweets were actually about being alone. To rectify this issue, we narrowed our search to phrases such as “I am alone,” “I feel alone,” “forever alone,” etc. (see Appendix for complete list of phrases).

Text analysis

Our focus in this analysis was the words that co-occur with solitary terms. We used pointwise mutual information (PMI) as an index of word co-occurrence (Church & Hanks, 1990):

$$PMI(x; y) = \log \frac{p(x, y)}{p(x)p(y)}$$

where $p(x)$ is the number of times word x occurs across all tweets, $p(y)$ is the number of times word y occurs across all tweets, and $p(x, y)$ is how often word x and word y occur in the same tweet.

We calculated PMIs for each solitary term (*alone*, *lonely*, and *solitude*) with each word in our emotion lexicon. This resulted in a new dataset of individual words and their likelihood of co-occurring with each of the solitary terms. In order to compare words that tend to co-occur more with one solitary term over another, we simply computed difference scores between each PMI, resulting in three sets of pairwise differences (*solitude* vs. *lonely*, *solitude* vs. *alone*, and *lonely* vs. *alone*).

Valence, arousal, and dominance. We used the NRC Valence Arousal Dominance lexicon (NRC VAD; Mohammad, 2018) to identify the valence, arousal, and dominance of co-occurring words. The NRC VAD lexicon contains over 20,000 commonly used English words that have been scored on valence (0 = extremely unpleasant, 1 = extremely pleasant), arousal (0 = calm, sleepy, 1 = active, intense), and dominance (0 = weak, 1 = powerful). Examples of high and low words for each dimension include: high valence = *love*, low valence = *toxic*; high arousal = *explosion*, low arousal = *napping*; high dominance = *conquer*, low dominance = *empty*.

Anger, fear, joy, and sadness. We used the NRC Emotion Lexicon Mohammad and Turney (2010) to identify anger, fear, joy, and sadness words. The NRC Emotion Lexicon

contains just over 14,000 commonly used English words with binary ratings corresponding to whether the word reflects the emotion label or not. For example, the word *isolate* has a score of 1 on the sadness label but 0 everywhere else.

Analytic plan

To reiterate, our goal was to examine the affective context in which solitary terms arise. Thus, we sought to determine the relation between word co-occurrence and dimensions of affect/discrete emotion labels. We limited our analysis to words that appeared at least 500 times across all tweets in the sample. Our reasoning for this was to reduce noise arising from words that are used less often.

Our choice of cut-off was completely arbitrary, so we examined different cut-offs to ensure our conclusions did not differ. Although the magnitude of the relation varies somewhat depending on this selection, the overall conclusions remain the same, thus we present our analyses here using the 500-word cut-off.

We performed a series of linear regressions, each one pitting two solitary terms against another with the independent variable being PMI difference (interpreted as the tendency for a word to co-occur with one term over the other) and the dependent variable being one of the three dimensions of affect or four emotion labels. This resulted in a total of 9 linear regressions and 12 logistic regressions, respectively.

Results

Word co-occurrences

We first examined word co-occurrence on its own. Of particular interest, was to see which words co-occurred most with *solitude* versus *lonely*. There is an interesting connection to popular culture references among some of these strongly co-occurring words. For instance, *solitude* co-occurred strongly with *fortress* and *superman* (presumably, as in Superman's "Fortress of Solitude"), whereas *lonely* co-occurred strongly with *sergeant* and *pepper* (presumably as in the Beatles' album "Sergeant Pepper's Lonely Hearts Club Band"). More germane co-occurrences with *solitude* were *recharge*, *tranquility*, and *meditation*, while, consistent with loneliness as a negative state, *lonely* co-occurred with *horny*, *depressed*, and *bored*, among others.

We performed similar comparisons with *solitude* versus *alone* and *lonely* versus *alone*, although *alone* seems to be less uniquely linked to relevant words even when we limited the sample to tweets using *alone* in the proper context. This likely reflects the fact that the word *alone* is used more commonly and in a wider variety of contexts.

Modeling affect of co-occurring words

Valence, arousal, and dominance. Our main analysis was to contrast word co-occurrence among solitary terms along three dimensions of affect: valence, arousal, and dominance. Results are presented in Table 1 and depicted graphically in Figure 1. First comparing *solitude* versus *lonely*, valence was significantly and positively associated with the tendency to co-occur with *solitude* over *lonely*. In other words, the more pleasant the

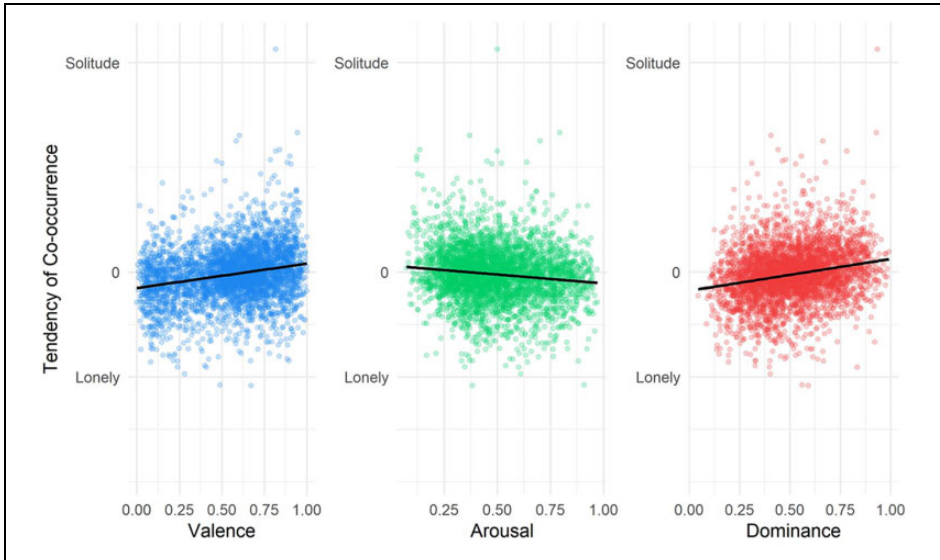


Figure 1. Scatterplot showing tendency of word co-occurrence of *solitude* and *lonely* with dimensions of affect. Each point is a word. *Note.* Word co-occurrence is a continuous variable, but the Y-axis is labeled with markers to aid interpretation.

word, the greater its likelihood of co-occurring with *solitude* instead of *lonely*. For arousal, the association was reversed, whereby arousal was significantly and negatively associated with the tendency to co-occur with *solitude*. As arousal decreases, its likelihood of co-occurring with *solitude* over *lonely* increases. Finally, dominance was significantly and positively associated with the tendency to co-occur with *solitude*. As dominance increases, its likelihood of co-occurring with *solitude* over *lonely* increases.

Comparing *solitude* versus *alone*, only valence (not dominance) was positively associated with the likelihood of co-occurring with *solitude* over *alone* and arousal was negatively associated. However, when we limited this comparison to tweets with key phrases such as “I am alone” or “forever alone” we found that valence and dominance were once again positively associated with the likelihood of co-occurring with *solitude*. As for the comparison with *lonely* versus *alone*, arousal and dominance were negatively associated with the likelihood of co-occurring with *lonely*, but there was no association with valence.

Anger, fear, joy, sadness. Analyses with the NRC Emotion Lexicon largely mirror those obtained using the VAD Lexicon. Results are presented in Table 2. Comparing *solitude* versus *lonely*, words that co-occurred more with *solitude* were less likely to be labeled as angry, fearful, and sad, but more likely to be labeled as joyful. A similar pattern of results was obtained comparing *solitude* with *alone*.

Comparing *lonely* versus *alone*, words that co-occurred more with *lonely* were more likely to be labeled as sad (see Table 3). However, there were no other significant relations when we compared *lonely* versus *alone*.

Table 2. Results from linear regressions predicting valence, arousal, and dominance from PMI.

Solitude vs. Lonely					
	<i>b</i>	<i>t</i>	<i>df</i>	<i>R</i> ²	<i>p</i>
Valence	0.072	13.900	3,651	0.050	<0.001
Arousal	-0.030	-7.492	3,651	0.015	<0.001
Dominance	0.051	13.380	3,651	0.047	<0.001
Solitude vs. Alone					
	<i>b</i>	<i>t</i>	<i>df</i>	<i>R</i> ²	<i>p</i>
Valence	0.057	10.800	3,687	0.030	<0.001
Arousal	-0.033	-8.441	3,687	0.019	<0.001
Dominance	0.022	5.520	3,687	0.008	<0.001
Lonely vs. Alone					
	<i>b</i>	<i>t</i>	<i>df</i>	<i>R</i> ²	<i>p</i>
Valence	-0.024	-3.767	5,305	0.002	<0.001
Arousal	-0.008	-1.734	5,305	0.001	0.083
Dominance	-0.057	-12.360	5,305	0.028	<0.001

Table 3. Results from generalized linear regressions predicting anger, fear, joy, and sadness from PMI.

Solitude vs. Lonely				
	<i>b</i>	<i>z</i>	<i>df</i>	<i>p</i>
Anger	-0.477	-5.627	2,928	<0.001
Fear	-0.235	-3.093	2,928	0.002
Joy	0.430	-5.565	2,928	<0.001
Sadness	-0.398	-5.145	2,928	<0.001
Solitude vs. Alone				
	<i>b</i>	<i>t</i>	<i>df</i>	<i>p</i>
Anger	-0.417	-4.820	2,949	<0.001
Fear	-0.252	-3.242	2,949	0.002
Joy	0.433	5.581	2,949	<0.001
Sadness	-0.021	-2.677	2,949	0.007
Lonely vs. Alone				
	<i>b</i>	<i>t</i>	<i>df</i>	<i>p</i>
Anger	0.039	0.404	4,193	0.698
Fear	-0.084	-0.928	4,193	0.353
Joy	0.075	0.733	4,193	0.463
Sadness	0.247	2.759	4,193	0.006

Discussion

Paradoxically, being alone is portrayed as undesirable and isolating, yet at times restorative and rejuvenating (Coplan et al., 2018). The terms *lonely* and *solitude* are commonly used in the literature to capture these divergent experiences of time alone. In this study, we set out to explore how these words are used in more everyday language, paying attention to the emotional context in which they arise. We collected a massive sample of tweets and applied computational linguistics methods to identify and compare the emotional content of words that co-occur with either *solitude*, *lonely*, or *alone*.

The primary goal of this study was to examine whether sentiment differs among words that co-occur more strongly among *solitude* versus *lonely*. Consistent with expectations, words that co-occurred more with *solitude* and less with *lonely* tended to be higher in valence and dominance and lower in arousal. This fits with our conceptualization of the term *solitude* as denoting a more pleasant, restorative, and intrinsically motivated experience of time alone (Galanaki, 2004; Nguyen et al., 2018, 2019), whereas *lonely* denotes a more unpleasant, stressful, and externally imposed experience of being alone (Hawley & Cacioppo, 2010; Qualter et al., 2015). We found additional support for these associations among discrete emotion labels, including anger, fear, joy, and sadness. Specifically, *solitude* co-occurred more with words corresponding to joy and less to words corresponding to anger, fear, and sadness.

That *solitude* co-occurs with more pleasant words than *lonely* is perhaps not surprising. Loneliness is increasingly recognized as a cause and consequence of mental health problems (Hawley & Cacioppo, 2010; Seeman, 1996) and has been labeled a major threat to public well-being (Bergland, 2015). Thus, *solitude* may be *relatively* more positive than *lonely*, but it does not necessarily tell us that it is viewed positively on its own. Instead, we can gain some insight into this question by looking at the comparison between *solitude* and *alone*, which shows a similar trend of *solitude* co-occurring with more pleasant words. Overall, compared to the conceptually negative term *lonely* and the neutral term *alone*, *solitude* seems to reflect greater valence and, specifically, greater joy.

Perhaps more noteworthy is that words co-occurring with *solitude* tended to be lower in arousal compared to those co-occurring with *lonely* or *alone*. Recall that arousal is a dimension of affect referring to physiologically activation (Russell, 2003; Russell & Mehrabian, 1977). For example, high arousal states are often described as tense, excited, or agitated, whereas low arousal states are calm, drowsy, and dull. Low arousal, in and of itself, is not necessarily good or bad in terms of implications for well-being. But, there are parallels between the benefits of solitude and mildly positive low arousal states (e.g., restful, relaxing, calming). People tend to experience reduced arousal when they spend time alone (Nguyen et al., 2018) and when people claim to be enjoying time alone, they report less cognitive effort and activity (e.g., restful) (Lay et al., 2019). In contrast, loneliness is associated with both momentary and chronic stress, which may explain why *lonely* occurs among higher arousal terms (Matias et al., 2011; Seeman, 1996).

We also found that *solitude* tended to co-occur with higher dominance words compared to *lonely* and *alone*. Dominance ranges from feeling a complete lack of control to feeling in command (Russell & Mehrabian, 1977). Subsequent theory and research extended dominance to refer to a sense of agency or control over a situation (Fontaine et al., 2007).

In light of this interpretation, our findings may reflect that *solitude* refers to a desire for time alone, whereas *lonely* captures time alone as externally imposed and undesirable (Galanaki, 2004). Loneliness is strongly associated with depression (Beutel et al., 2017), which itself is a pervasive feeling of helplessness (i.e., low dominance). However, the interpretation of dominance as reflecting intrinsic motivation remains speculative given that motivation is a complex concept involving both affective and cognitive elements.

Our results also shed some light on how the term *alone* is used relative to *solitude* and *lonely*. The pattern of results suggests that *alone* has an affective context closer to *lonely* than *solitude*. Although we did not formally test whether *lonely* and *alone* are equivalent, we found that many of the differences in affective context between *solitude* and *alone* were not replicated in the comparison of *alone* and *lonely*. This could suggest that *alone* is a less neutral concept relative to the other terms, which deviates from how it is often conceptualized in the psychological literature (Larson, 1990).

What do these findings tell us about the psychological study of solitude? We focused on three words that capture three related, but conceptually different ways of experiencing time alone. Our results lend credence to the view that *solitude* describes an emotional context that is distinct from *loneliness*. Although this distinction is not a novel one, this study was the first to demonstrate evidence of this in natural language. Overall, we should be somewhat cautious about inferring the intent behind these words, but our results do support the conceptual distinctions between the positive experience of solitude and its negative, lonely counterpart.

Beyond bolstering our conceptual ideas of solitude, are there practical applications for these findings? We previously alluded to the possibility that these findings could help refine our measurement of different states of being alone. For instance, researchers developing questionnaires to measure aloneness, or solitude, or loneliness should be careful in their wording of questions to ensure that the meaning is conveyed accurately. Choosing the word *lonely* instead of *solitude* in a questionnaire could have a profound impact on how a participant responds to that question. Separately, clinicians may gain a deeper understanding of how people describe different states of being alone. Paying careful attention to when a client uses the word *lonely* could facilitate more efficient treatment. Extrapolating even further, social media platforms could further harness such algorithms to identify people potentially at risk of loneliness and help them by “nudging” them to engage in social activities.

This final point brings up the question of whether we *should* develop applications to identify people who are lonely and whether we ought to intervene in these cases. This is a complicated issue and deserves full treatment of its own. Briefly, there are already applications designed to identify people at risk of suicide or depression based on their social media activity (de Andrade et al., 2018). But, extending this into chronic loneliness would prove more challenging given that loneliness is more pervasive. Even more problematic is the question of how identification would lead to intervention. Prompting people with messages to connect with others is unlikely to be fruitful, although if someone is already meeting regularly with a therapist, it may be beneficial for the therapist to have some indication of whether their client uses *loneliness* instead of *solitude* on their social media. This of course introduces more ethical issues regarding data confidentiality and security. Research in this area should continue to investigate the ethical implications of natural language processing in real life.

Limitations and future directions

In this study, we applied natural language processing techniques to compare the emotional context of different solitary terms. It is important, however, to be cautious against interpreting these results too strongly. Just because a tweet contains the word *solitude* or *lonely*, it does not mean that the tweeter is experiencing one of these states. Similarly, a word that is labeled as high valence, for example, may not actually be used to express high valence in some contexts. All of this is to say that our findings only speak to the sentiment of words that co-occur with solitude terms and not the overall sentiment of a tweet. It would be worth further examining the narrative context of tweets to determine whether the author is writing about feeling lonely themselves, or decrying loneliness as a public health issue.

It is worth mentioning too that although tweets are naturalistic representations of online language, they are still different from how people converse in other domains (e.g., face-to-face). tweets serve a different purpose than everyday conversational language. Thus, we should consider how these results might differ if they were applied to conversational language or even more private social media platforms (e.g. Facebook). A final limitation is that our choice of key terms may not have exhausted all words used to refer to the state of being alone, and may therefore have omitted relevant tweets. Taken together, these results give us a glimpse into how people understand and use words like *solitude* and *lonely*. We see these findings as strengthening our conceptual distinction between positive and negative experiences of time alone. Moving forward, these results may represent a starting point for understanding the language surrounding loneliness and provide a means of identifying individuals at risk of chronic loneliness.

Authors' note

A version of this paper appears as a conference proceedings from the *International Conference on Language Resources and Evaluation* (LREC). Marseilles, France (2020).


Data sharing statement

Data and scripts to reproduce results in Tables and Figures are here: <https://github.com/whipson/solitude-tweets-jspr>. Raw tweets will not be shared as this violates Twitter's API policies.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Will E. Hipson  <https://orcid.org/0000-0002-3931-2189>

Open research statement

As part of IARR's encouragement of open research practices, the author(s) have provided the following information: This research was not pre-registered. The data used in the research can be publicly posted. The data and materials can be obtained at: <https://github.com/whipson/SPSP2020-solitude-tweets>.

References

- Aman, S., & Szpakowicz, S. (2007, September 3–7). *Identifying expressions of emotion in text* [Conference session]. International Conference on Text, Speech and Dialogue, Pilsen, Czech Republic.
- Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review*, 5(4), 379–389.
- Bergland, C. (2015). Loneliness: Perceived social isolation is public enemy no. 1. *Psychology Today*.
- Beutel, M. E., Klein, E. M., Brahler, E., Reiner, I., Junger, C., Michal, M., Wiltink, J., Wild, P. S., Münzel, T., Lackner, K. J., & Tibubos, A. N. (2017). Loneliness in the general population: Prevalence, determinants and relations to mental health. *BMC Psychiatry*, 17(1), 97.
- Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1), 22–29.
- Coplan, R. J., & Bowker, J. C. (2014). All alone: Multiple perspectives on the study of solitude. In R. J. Coplan & J. C. Bowker (Eds.), *The handbook of solitude: Psychological perspectives on social isolation, social withdrawal, and being alone* (pp. 1–13). Wiley Online Library.
- Coplan, R. J., Hipson, W. E., Archbell, K. A., Ooi, L. L., Baldwin, D., & Bowker, J. C. (2019). Seeking more solitude: Conceptualization, assessment, and implications of aloneliness. *Personality and Individual Differences*, 148, 17–26.
- Coplan, R. J., Zelenski, J., & Bowker, J. (2018). Well enough alone? The costs and benefits of solitude. In J. E. Maddux (Ed.), *Subjective well-being and life satisfaction* (pp. 129–147). Routledge.
- Cox, J. (2017). *Commission on loneliness: A call to action*. Report, Jo Cox Commission on Loneliness, 15.
- de Andrade, N. N. G., Pawson, D., Muriello, D., Donahue, L., & Guadagno, J. (2018). Ethics and artificial intelligence: Suicide prevention on Facebook. *Philosophy & Technology*, 31(4), 669–684.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3(4), 364–370.
- Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, 18(12), 1050–1057.
- Galanaki, E. (2004). Are children able to distinguish among the concepts of aloneness, loneliness, and solitude? *International Journal of Behavioral Development*, 28(5), 435–443.
- Goffman, E. (1971). *The presentation of self in everyday life/Erving Goffman*. Penguin.
- Gratch, J., Lucas, G., Malandrakis, N., Szablowski, E., Fessler, E., & Nichols, J. (2015, September 21–24). Goallll!: Using sentiment in the world cup to explore theories of emotion [Conference session]. 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, China.
- Hawkey, L. C., & Cacioppo, J. T. (2010). Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine*, 40(2), 218–227.
- Hipson, W. E. (2019). Using sentiment analysis to detect affect in children's and adolescents' poetry. *International Journal of Behavioral Development*, 43(4), 375–382.
- Kiritchenko, S., Hipson, W. E., Coplan, R. J., & Mohammad, S. M. (2020). *SOLO: A corpus of tweets for examining the state of being alone*. arXiv preprint arXiv:2006.03096.
- Korpela, K. M., Hartig, T., Kaiser, F. G., & Fuhrer, U. (2001). Restorative experience and self-regulation in favorite places. *Environment and Behavior*, 33(4), 572–589.

- Kross, E., Verduyn, P., Boyer, M., Drake, B., Gainsburg, I., Vickers, B., Ybrra, O., & Jonides, J. (2019). Does counting emotion words on online social networks provide a window into people's subjective experience of emotion? A case study on Facebook. *Emotion, 19*(1), 97.
- Larson, R. (1990). The solitary side of life: An examination of the time people spend alone from childhood to old age. *Developmental Review, 10*(2), 155–183.
- Lay, J. C., Pauly, T., Graf, P., Biesanz, J. C., & Hoppmann, C. A. (2019). By myself and liking it? Predictors of distinct types of solitude experiences in daily life. *Journal of Personality, 87*(3), 633–647.
- Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology, 6*, 444.
- Long, C. R., Sebum, M., Averill, J. R., & More, T. A. (2003). Solitude experiences: Varieties, settings, and individual differences. *Personality and Social Psychology Bulletin, 29*(5), 578–583.
- Matias, G. P., Nicolson, N. A., & Freire, T. (2011). Solitude and cortisol: Associations with state and trait affect in daily life. *Biological Psychology, 86*(3), 314–319.
- Maynard, D., & Funk, A. (2011, May). Automatic detection of political opinions in tweets. In *Extended Semantic Web Conference* (pp. 88–99). Springer, Berlin, Heidelberg.
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. The MIT Press.
- Mohammad, S. M. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In I. Gurevych & Y. Miyao (Eds.), *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 174–184). Association for Computational Linguistics.
- Mohammad, S. M., & Turney, P. D. (2010). Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26–34). Association for Computational Linguistics.
- Nguyen, T.-v. T., Ryan, R. M., & Deci, E. L. (2018). Solitude as an approach to affective self-regulation. *Personality and Social Psychology Bulletin, 44*(1), 92–106.
- Nguyen, T.-v. T., Werner, K. M., & Soenens, B. (2019). Embracing me-time: Motivation for solitude during transition to college. *Motivation and Emotion, 43*(4), 571–591.
- Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge University Press.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois press.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*(6), 1296.
- Pew Research Center. (2019). <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/>
- Qualter, P., Vanhalst, J., Harris, R., Van Roekel, E., Lodder, G., Bangee, M., Maes, M., & Verhagen, M. (2015). Loneliness across the life span. *Perspectives on Psychological Science, 10*(2), 250–264.
- Rokach, A. (2004). Loneliness then and now: Reflections on social and emotional alienation in everyday life. *Current Psychology, 23*(1), 24–40.

- Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion, 18*(8), 1121–1133.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145.
- Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality, 11*(3), 273–294.
- Seeman, T. E. (1996). Social ties and health: The benefits of social integration. *Annals of Epidemiology, 6*(5), 442–451.
- Shelley, K. J. (2005). Developing the American Time Use Survey activity classification system. *Monthly Labour Review, 128*, 3.
- Tang, D., Wei, F., Qin, B., Zhou, M., & Liu, T. (2014). Building large-scale Twitter-specific sentiment lexicon: A representation learning approach. In J. Tsujii & J. Hajic (Eds.), *Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical papers* (pp. 172–182). Dublin City University and Association for Computational Linguistics.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24–54.
- Uziel, L. (2021). The language of being alone and being with others. *Social Psychology, 52*(1), 13–22. <https://doi.org/10.1027/1864-9335/a000430>
- Wang, Z., Tchernev, J. M., & Solloway, T. (2012). A dynamic longitudinal examination of social media use, needs, and gratifications among college students. *Computers in Human Behavior, 28*(5), 1829–1839.
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207.
- Zhang, L., Liu, B., Lim, S. H., & O'Brien-Strain, E. (2010). Extracting and ranking product features in opinion documents. In *Proceedings of the 23rd international conference on computational linguistics: Posters* (pp. 1462–1470). Association for Computational Linguistics.

Appendix A

List of “alone” phrases

For tweets classified as *alone*, we limited our search to the following strings (note that tweet text is converted to lowercase prior to filtering):

be alone, being alone, am alone, i'm alone, im alone, you are alone, you're alone, youre alone, i'm all alone, im all alone, am all alone, be all alone, me all alone, you're all alone, youre all alone, you are all alone, not alone, you aren't alone, you arent alone, never alone, so alone, forever alone, completely alone, always alone, more alone, less alone, feel alone, feeling alone, felt alone, live alone, living alone, lived alone, die alone, dying alone, suffer alone, suffering alone, suffered alone, better off alone, end up alone, sad and alone