

The Impact of Single Status on Life Satisfaction: A Study Based on Zhihu Data

Authors: Zhang Zhiyong, Ma Weiwei, Zhang Shiyu, Qiu Yi, Xu Chenglong,
Zhang Hongyu, Zhu Tingshao, Zhao Nan, Zhu Tingshao

Date: 2022-03-06T14:54:17Z

Abstract

To investigate the differences between single individuals and individuals in romantic relationships regarding life satisfaction and the proportion of emotional word frequency, this study leveraged large-scale web data, employing a Python crawler program to extract user posts from the Zhihu platform under singlehood and romantic relationship themes, and obtained the corresponding populations' emotional word frequency proportions and life satisfaction scores through the 'Wenxin' system and a life satisfaction prediction model. The results revealed that the romantic relationship group exhibited significantly higher life satisfaction than the single group ($t=4.415$, $p < 0.001$); the proportions of positive emotional words ($t=-9.061$, $p < 0.001$) and anxiety words ($t=1.844$, $p < 0.001$) in the romantic relationship group were significantly lower than those in the single group, while the word frequency proportion of anger words ($t=5.101$, $p < 0.001$) was significantly higher than that in the single group. The results indicate that while individuals in romantic relationships obtain higher life satisfaction and lower anxiety levels, they also need to address emotion-related issues associated with partners.

Full Text

The Impact of Singlehood on Life Satisfaction: A Study Based on Zhihu Data

ZHANG Zhiyong^{1,2}, MA Weiwei^{1,2}, ZHANG Shiyu^{1,2}, QIU Yi^{1,2}, XU Chenglong^{1,2}, ZHANG Hongyu^{1,2}, ZHU Tingshao^{1,2}, ZHAO Nan^{1,2}

¹Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101

²Department of Psychology, University of Chinese Academy of Sciences, Beijing, 100049

Abstract

This study investigates differences in life satisfaction and emotional word frequency between single and married/romantically involved individuals using large-scale online data. We collected user posts from Zhihu, China's largest Q&A platform, on topics related to singlehood and romantic relationships using a Python web crawler. Emotional word frequencies and life satisfaction scores were obtained through the "Wenxin" text analysis system and a life satisfaction prediction model. Results revealed that the married/romantic relationship group exhibited significantly higher life satisfaction than the single group ($t = 4.415$, $p < 0.001$). The relationship group also showed significantly lower proportions of positive emotion words ($t = -9.061$, $p < 0.001$) and anxiety words ($t = 1.844$, $p < 0.001$), but significantly higher proportions of anger words ($t = 5.101$, $p < 0.001$) compared to the single group. These findings suggest that while individuals in romantic relationships experience higher life satisfaction and lower anxiety levels, they must also navigate emotion-related challenges associated with partnership.

Keywords: single population; married population; word frequency analysis; life satisfaction; machine learning

Since 2013, the number of marriage registrations in China has declined annually, with an increasing number of eligible men and women choosing to remain single, either voluntarily or involuntarily. By 2021, China's single population reached approximately 300 million, with projections indicating continued growth over the next decade. Single individuals have become a significant demographic component of society, and their life status and satisfaction represent important factors affecting social stability and harmony. A recent survey reported by CCTV on February 15, 2021, revealed that single individuals' happiness and life satisfaction have surpassed those of married individuals, challenging mainstream research findings that have traditionally reported lower satisfaction among singles. While existing studies have examined differences in life satisfaction between single and married individuals, results remain inconsistent. As romantic and marital values evolve, emerging research suggests that single individuals may not be as unhappy as previously assumed (Chopik et al., 2021), and scholars have begun to differentiate subtypes within the single population based on gender, culture, and personality traits. Verbrugge (1979) argued that married individuals experience higher well-being and lower anxiety compared to their unmarried counterparts. However, a study of 553 young adults aged 20-30 found that singlehood did not directly affect life satisfaction, though higher loneliness and lower perceived social support among singles indirectly compromised their satisfaction (Adamczyk & Segrin, 2015). Other studies have reached different conclusions: Girmé et al. (2016) found that single individuals with high avoidance goals reported life satisfaction equivalent to those in relationships, while research on Hong Kong women showed that marital status did not affect female life satisfaction, with married women actually reporting lower satisfaction than unmarried women (Markey et al., 2005; Yang & Yen, 2014).

Life satisfaction represents an individual's comprehensive subjective evaluation of their quality of life, providing a more complete cognitive assessment of living conditions typically measured through questionnaires. Life satisfaction is influenced by multiple subjective and objective factors. Subjective factors primarily include interpersonal relationships, personality traits, values, and other cognitive and emotional dimensions (Veenhoven, 1984), while objective factors have focused on income, culture, living environment, and social policies and welfare (Zhu, 2016).

The impact of singlehood on life satisfaction operates through distinct mechanisms. On one hand, romantic relationships exhibit "selection effects" and "protection effects." The selection effect refers to the tendency for partners to be attracted to each other based on shared characteristics, and compatibility with a like-minded partner enhances mutual happiness. The protection effect suggests that romantic relationships benefit psychological and physical health while providing emotional and financial support that helps individuals overcome life difficulties, thereby increasing life satisfaction (Gao, 2021). On the other hand, compared to those in relationships who often reduce other social activities due to partnership commitments, single individuals have more time to engage with family, friends, and colleagues, potentially deriving greater happiness from diverse relationships. Those who actively choose singlehood often possess economic advantages, higher education, social status, and stronger self-identity, enabling them to satisfy their desires and pursue self-actualization, which may lead to higher life satisfaction (Song & Wang, 2016).

Given these theoretical perspectives and contradictory empirical findings, this study adopts a novel methodological approach using online platform data to examine differences in life satisfaction between single and partnered individuals through analysis of their online expressions. Investigating the relationship between relationship status and life satisfaction among young adults not only provides positive guidance for establishing healthy romantic and marital values but also offers theoretical and empirical references for national policy-making on marriage and relationships, thereby promoting stable partnerships. This research contributes to the literature for sociologists, demographers, and psychologists while pioneering the use of social media data in this domain.

2.1 Participant Selection

This study utilized responses from Zhihu users under relevant topics. As China's largest platform for in-depth content, Zhihu allows users to answer and discuss questions of interest. Selecting Zhihu responses for comparing life satisfaction between single and partnered individuals offers several advantages.

Compared to traditional questionnaire research, social media studies provide distinct benefits: First, they demonstrate high ecological validity. Zhihu's anonymous response feature reduces social desirability bias common in questionnaire studies, enhancing data authenticity. Second, they yield larger sample sizes at

lower cost. Zhihu's textual data also demonstrate high validity: users can describe their experiences around specific themes in response format, facilitating relatively easy extraction of target populations. This format encourages comprehensive reflection in a relatively rational state, producing objective and thorough descriptions. Most responses meet requirements for describing life status and remain highly relevant to research questions. Additionally, leveraging Zhihu's question descriptions and user profile information enables partial control for variables such as gender and occupation.

This study crawled responses under questions related to self-descriptions of life status among single or partnered individuals. All collected data were publicly available and did not involve users' personal identity information.

2.2 Web Data Crawling

For selected Zhihu questions, we used a Python program to simulate browser login status, request Zhihu API links, and sequentially crawl response data under each question for screening and storage.

2.3 Data Cleaning

Data cleaning involved both programmatic and manual stages.

2.3.1 Programmatic Cleaning

The web crawler collected approximately 950 user posts each from singlehood and relationship topics. During crawling, we conducted initial screening to retain only responses containing 300 or more characters. Additionally, we programmatically removed texts containing external links or the term "public account" to eliminate most advertisements or posts from marketing and institutional accounts that do not reflect individual life status. Finally, by comparing whether the first 50 characters of consecutive responses were identical, we eliminated duplicate texts resulting from users updating their responses, retaining only the original version.

2.3.2 Manual Cleaning

Following programmatic cleaning, data underwent further manual screening to retain qualified responses. Manual screening criteria included: (1) Relevance to topic: eliminating completely off-topic responses. Additionally, some single individuals described their single lives under relationship topics while explicitly identifying as single, or vice versa; such data were reclassified accordingly. (2) Removing responses from marketing or institutional accounts not identified by the program. (3) Eliminating commentary responses rather than self-descriptions. (4) Removing fictional responses presented in novel or story formats. (5) Excluding texts from a bystander perspective. (6) Removing meaningless texts.

Following these criteria, we obtained 1,739 valid responses, including 861 self-descriptions from single individuals and 878 from partnered individuals.

2.4 Content Feature Extraction

The “Wenxin” Chinese Psychological Analysis System (TextMind), developed by the Computational Cyber-Psychology Laboratory at the Institute of Psychology, Chinese Academy of Sciences, includes Chinese word segmentation tools and psychological analysis dictionaries to analyze the degree and preference of different language categories in texts. This study used the Wenxin system to extract content features from cleaned data, employing its Simplified Chinese LIWC dictionary to calculate emotional word frequency proportions for single and partnered populations.

2.5 Life Satisfaction Prediction Model

Life satisfaction reflects individuals’ subjective satisfaction with their own status. This study used volunteers’ life description texts and life satisfaction questionnaire scores as labeled data to establish a life satisfaction prediction model through machine learning algorithms. Cleaned data were then input into this model to obtain corresponding life satisfaction scores.

3.1 Word Frequency Analysis

We conducted independent samples t-tests using SPSS 26.0 to examine differences in emotional word frequency proportions between the two groups. As shown in Table 1, the partnered group exhibited significantly lower proportions of positive emotion words ($t = -9.061$, $p < 0.001$) and anxiety words ($t = 1.844$, $p < 0.001$) compared to the single group, but significantly higher proportions of anger words ($t = 5.101$, $p < 0.001$).

Table 1 Descriptive Statistics and Independent Samples t-Test Results for Word Frequency Analysis

Variable	Partnered Group (M \pm SD)	Single Group (M \pm SD)	t-value	p-value
Positive Emotion Words	0.028 \pm 0.012	0.034 \pm 0.016	-9.061	0.000***
Anxiety Words	0.017 \pm 0.008	0.017 \pm 0.010	1.844	0.000***
Anger Words	0.002 \pm 0.002	0.003 \pm 0.004	5.101	0.000***
Negative Emotion Words	0.004 \pm 0.004	0.003 \pm 0.004	-	-

Variable	Partnered Group (M \pm SD)	Single Group (M \pm SD)	t-value	p-value
Other	0.003 \pm 0.003	0.003 \pm 0.004	-	-

Note: $p < .05$, $p < .01$, $p < .001$

3.2 Life Satisfaction Index

We compared life satisfaction data between partnered and single groups using independent samples t-tests after processing through the prediction model. As shown in Table 2, the partnered group demonstrated significantly higher life satisfaction than the single group ($t = 4.415$, $p < 0.001$).

Table 2 Descriptive Statistics and Independent Samples t-Test Results for Life Satisfaction

Variable	Partnered Group (M \pm SD)	Single Group (M \pm SD)	t-value	p-value
Life Satisfaction	0.613 \pm 0.064	0.598 \pm 0.072	4.415	0.000***

Note: $p < .05$, $p < .01$, $p < .001$

4 Discussion and Outlook

To investigate the impact of singlehood on life satisfaction, this study used a Python crawler to collect user posts from Zhihu topics related to singlehood and romantic relationships, employing mirrored questions to reduce sample selection bias to some extent. After programmatic and manual data cleaning, we obtained 1,739 valid responses from an initial pool of 1,936, including 861 texts from single individuals and 878 from partnered individuals. We used the Wenxin system and life satisfaction prediction model to obtain word frequency features and life satisfaction indices for group comparison. Results showed that the partnered group exhibited significantly higher life satisfaction than the single group. Word frequency analysis further revealed that the partnered group had significantly lower proportions of positive emotion words and anxiety words, but significantly higher proportions of anger words compared to the single group.

The finding that partnered individuals show higher life satisfaction aligns with our expectations and some previous research. On one hand, single individuals face greater economic and social pressures, and loneliness generates substantial psychological stress. We speculate that these multiple internal and external pressures may deplete psychological resources, resulting in relatively lower life satisfaction. On the other hand, marriage offers numerous potential benefits,

including children, stable sexual relationships, family and economic support, and companionship in old age. The social support provided by marriage benefits physical and mental health and enhances subjective well-being (Yuan & Li, 2017), while selection and protection effects help individuals overcome life difficulties and obtain emotional and financial assistance, thereby increasing happiness (Gao, 2021), leading to higher life satisfaction among partnered individuals. Word frequency results showing lower positive emotion and anxiety words but higher anger words among partnered individuals may appear contradictory to increased life satisfaction. However, Diener (2006) decomposed subjective well-being into three components: (1) life satisfaction—overall and domain-specific evaluations of life aspects such as marriage, work, income, housing, and leisure; (2) emotions—including positive emotions like pleasure and contentment and negative emotions like depression and anger; and (3) value judgments—such as sense of achievement and meaning in life. Thus, entering romantic or marital relationships may enhance life satisfaction while having mixed effects on the emotional component of subjective well-being.

This study's use of Zhihu data offers advantages including high ecological validity, large sample size, and low cost. However, several limitations warrant consideration. First, single individuals, having more free time and energy, may use online platforms more frequently to post self-descriptions. Whether this usage difference affects our results remains unclear and requires further investigation. Second, motivations for singlehood vary—some individuals remain single due to life or economic pressures, while others prioritize career and self-actualization. Life satisfaction likely differs across these motivation-based subgroups, necessitating deeper research. Third, our word frequency analysis shows that partnered individuals exhibit significantly higher anger word proportions and lower positive emotion word proportions, reflecting negative aspects of romantic relationships, though the reasons for these differences require further analysis. Finally, this study only distinguished between single and non-single groups without fully exploiting publicly available Zhihu user information for more granular segmentation by gender, age, education level, and other variables. Future research should leverage this information to control for relevant variables and clarify their impact on life satisfaction.

This study selected Zhihu as a data source, used a Python crawler to collect user posts under relevant topics, and obtained text data through programmatic and manual screening. We then used the Wenxin system and life satisfaction prediction model to acquire emotional word frequency proportions and life satisfaction scores for corresponding populations. Analysis revealed that partnered individuals exhibit higher life satisfaction and lower anxiety levels compared to singles, but also show significantly lower positive emotion word frequencies and higher anger word frequencies.

In conclusion, while individuals in romantic relationships experience higher life satisfaction and lower anxiety, they must also navigate emotion-related challenges associated with partnership. These findings provide valuable insights for

young adults making decisions about romantic relationships.

References

- Zhu, D. (2016). Market competition, collective consumption, and environmental quality: An analysis of urban residents' life satisfaction and its influencing factors. *Sociological Studies*, (03), 193-217+246. doi:10.19934/j.cnki.shxyj.2016.03.009.
- Gao, X. (2021). A study on emotional attitudes and adjustment of unmarried women from an individualization perspective. (2019-4), 100-112.
- Song, J., & Wang, J. (2016). Marital status and subjective well-being among Chinese youth. *China Youth Study*, (9).
- Gao, X. (2021). A study on emotional attitudes and adjustment of unmarried women from an individualization perspective. (2019-4), 100-112.
- Yuan, Z., & Li, L. (2017). Marriage and happiness: Based on WVS Chinese micro-data. *China Economic Issues*, (01), 24-35. doi:10.19365/j.issn1000-4181.2017.01.03.
- Verbrugge, L. M. (1979). Marital status and health. *Journal of Marriage and the Family*, 267-285.
- Adamczyk, K., & Segrin, C. (2015). Direct and indirect effects of young adults' relationship status on life satisfaction through loneliness and perceived social support. *Psychologica Belgica*, 55(4), 196.
- Girme, Y. U., Overall, N. C., Faingataa, S., & Sibley, C. G. (2016). Happily single: The link between relationship status and well-being depends on avoidance and approach social goals. *Social Psychological and Personality Science*, 7(2), 122-130.
- Oh, J., Chopik, W. J., & Lucas, R. E. (2021). Happiness Singled Out: Bidirectional Associations Between Singlehood and Life Satisfaction. *Personality and Social Psychology Bulletin*.
- Veenhoven, R. (1984). *Conditions of happiness*. D. Reidel Pub.
- Markey, C. N., Markey, P. N., Schneider, C., & Brownlee, S. (2005). Marital status and health beliefs: Different relations for men and women. *Sex Roles*, 53(5), 443-451.
- Yang, W. S., & Yen, P. C. (2014). Single and happier? A comparative study of marital status, socioeconomic security and wellbeing of women in Hong Kong and Taiwan. In *The family and social change in Chinese societies* (pp. 81-97). Springer, Dordrecht.
- Diener, Ed (2006). Guidelines for National Indicators of Subjective Well-Being and Ill-Being. *Journal of Happiness Studies*, 7.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.