Estimation of Happiness Changes through Longitudinal Analysis of Employees' Texts

Junko Hayashi[1], **Kazuhiro Ito**[1], **Masae Manabe**[2], **Yasushi Watanabe**[2], **Masataka Nakayama**[2], **Yukiko Uchida**[2], **Shoko Wakamiya**[1], and **Eiji Aramaki**[1]

[1]Nara Institute of Science and Technology, Nara, Japan
[2]Kyoto University, Kyoto, Japan
{hayashi.junko.hh5, ito.kazuhiro.ih4, wakamiya, aramaki}@is.naist.jp
{manabe.masae.4y, nakayama.masataka.4v, uchida.yukiko.6m}@kyoto-u.ac.jp
yasushi.watanabe.77@gmail.com

Abstract

Measuring happiness as a determinant of well-being is increasingly recognized as crucial. While previous studies have utilized free-text descriptions to estimate happiness on a broad scale, limited research has focused on tracking individual fluctuations in happiness over time owing to the challenges associated with longitudinal data collection. This study addresses this issue by obtaining longitudinal data from two workplaces over two and six months respectively. Subsequently, the data is used to construct a happiness estimation model and assess individual happiness levels. Evaluation of the model performance using correlation coefficients shows variability in the correlation values among individuals. Notably, the model performs satisfactorily in estimating 9 of the 11 users' happiness scores, with a correlation coefficient of 0.4 or higher. To investigate the factors affecting the model performance, we examine the relationship between the model performance and variables such as sentence length, lexical diversity, and personality traits. Correlations are observed between these features and model performance.

1 Introduction

Well-being, happiness, and mental quality of life (referred to as "happiness" herein) are key factors affecting workplace performance (Diener and Biswas-Diener, 2002). Low happiness tends to cause issues such as decreased productivity (Shi et al., 2013) and increased turnover (Hurtz and Donovan, 2000). Hence, monitoring happiness is essential in order to make interventions, such as having them seek consultation at the appropriate time. Consequently, several studies regarding happiness measurement have been conducted. However, these studies primarily utilized questionnaire-based survey methods, which do not adequately examine short-term changes in individual happiness in the workplace. Thus, studies suffer from capturing short-term happiness, because most of the questionnaires, such as the Satisfaction with Life Scale (SWLS) (Diener et al., 1985), are not designed for daily assessment.

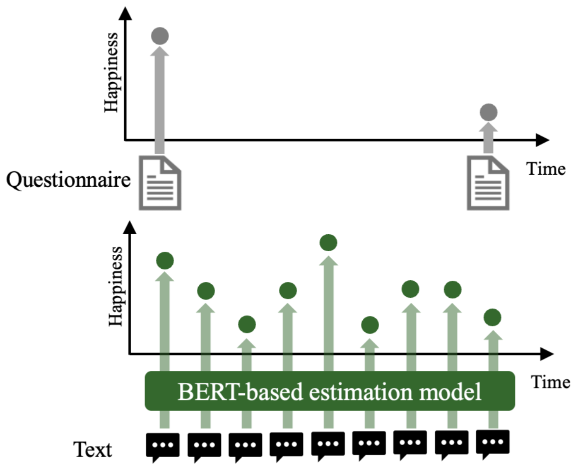


Figure 1: Overview of questionnaire-based survey methods (above) and BERT-based estimation model (below). To perform short-term surveys that cannot be easily conducted using questionnaires, this study estimates changes in happiness levels based on free-text responses.

This study addresses the challenge of capturing short-term happiness. To achieve this, we leveraged free-text diaries, which can be reported daily and offers a more flexible alternative to traditional questionnaires. Previous studies have demonstrated the feasibility of estimating happiness from text data (Schwartz et al., 2015; Wu et al., 2017; Coşkun and Ozturan, 2018; Jaidka et al., 2020; Kjell et al., 2022; Song and Zhao, 2023). However, these studies primarily focused on estimating overall happiness levels at the mass level, instead of monitoring individual changes in happiness over time.

This study aims to capture longitudinal individual happiness in the workplace (see Figure 1). Thus, we obtained daily self-report texts with happiness scores from employees in their workplaces and then

Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, pages 294–304

constructed a model based on bidirectional encoder representations from a transformer (BERT) (Devlin et al., 2018) to estimate happiness from the text. To generalize the model, we constructed a model that estimates happiness using only one diary for a specific day.

The contribution of this study is two-fold:

- A new task is proposed that employs happiness estimation model in the workplace (Section 5.1).
- Suggestions are provided for mitigating performance issues using a happiness estimation model (Section 5.2).

The proposed model is considerably simpler than conventional questionnaires, thereby facilitating future psychological research.

2 Related work

We describe studies that solved the task of estimating happiness from text. Previous studies that constructed estimation models can be categorized into two approaches: one that utilizes datasets containing texts annotated by individuals other than the original authors, and another that employs datasets containing texts and happiness levels obtained from the same individuals.

First, we introduce an approach that uses a dataset in which happiness is assigned to a text by a person who is not the author of the text. There are studies that estimate happiness on a large scale from social media posts (Sametoğlu et al., 2023). These studies (Mitchell et al., 2013; Schwartz et al., 2015; Jaidka et al., 2020) use dictionaries such as Linguistic Inquiry and Word Count (LIWC)[1] and LabMT (Dodds et al., 2011).

Second, we introduce an approach employing datasets where both the texts and happiness levels are collected from the same individuals (Wu et al., 2017; Kjell et al., 2022; Song and Zhao, 2023). In this case, no specific annotators are involved and no particular criterion is set for happiness. However, the happiness levels of individuals exhibit commonalities, as shown in the texts. These studies used methods such as support vector machines and decision trees. In a study that estimated happiness using BERT with textual responses to life-satisfaction questions (Kjell et al., 2022), higher performances were obtained compared with other studies. Therefore, in this study, we used BERT as a training method to estimate happiness in the workplace.

[1] https://www.liwc.app

3 Materials

3.1 Data collection

We obtained two types of data: personality traits, which were obtained simultaneously prior to performing a survey; and daily reports, which were obtained throughout the experimental period, i.e., during the survey.

All participants completed a questionnaire at the beginning of the study period. Each item was answered based on a 5-point Likert scale. These scales are customized for investigating cooperativeness and autonomy in the workplace and have been proposed in the papers discussed for each item. We were the only researchers to verify the obtained data.

Cooperativeness The question pertaining to cooperativeness comprised 10 items, such as “I’m concerned about others’ perception of me at the workplace.” These questions refer to the following scale: (Hitokoto and Uchida, 2014). All the questions are listed in Table 4.

Autonomy Each question comprised 10 items, such as “I always try to form my own opinion within the company.” Such questions are based on the scales reported by (Watanabe et al., 2023; Domae et al., 2024). All the questions are listed in Table 5.

Next, we obtained two items: daily report texts and self-reported happiness scores.

Daily report text The daily reports encompass free descriptions in Japanese detailing the participants’ daily lives, with no character limits. Examples of these data are presented in Table 2.

Self-reported happiness score An 11-point scale ranging from 0 (extremely unhappy) to 10 (extremely happy) was used. This study focused on the cognitive aspects of workplace happiness. Questions pertaining to happiness were based on the Cantril ladder (Cantril, 1965), which assesses happiness via comparison to an 11-step ladder.

(a) Company A (b) Company B

Figure 2: Distribution of self-reported happiness scores based on data obtained.

We used the data of company A, which were used in a previous study to investigate the happiness levels of employees at the company(Ito et al., 2023). Furthermore, we obtained additional data from company B to evaluate the estimation model. The name of each company has been changed to preserve anonymity.

An original web browser-based application was used to obtain the daily reports. The instructions provided to the participants included: “The input timing is typically once or more per working day, thus reflecting the day’s events at the end of the work day.”

3.2 Data Statistics and Examples

Data were obtained from company A, which is a major advertising and marketing company in Japan. In this company, 94 members participated in the input of daily reports over two months (from September 1 to October 31, 2022). Consequently, 1,728 data points were obtained from company A (Table 1). Figure 2 (a) shows the distribution of the data obtained.

Additionally, data were obtained from company B, which is a major Japanese electronics manufacturer. In this company, 11 employees provided daily report texts over six months (from December 19, 2022, to May 18, 2023). Consequently, 652 data points were collected from company B (Table 1). Figure 2 (b) shows the distribution of the data obtained.

4 Experiments

4.1 Modeling

We constructed a model in which the input was the daily report text and the output was the estimated happiness score. The training data were obtained from company A. In this experiment, we built a bidirectional encoder representation from a transformer (BERT) (Devlin et al., 2018)-based estimation model. Specifically, we constructed a regression model of the AutoModel for Sequence Classification[2] using Tohoku University’s pre-trained BERT[3]. The learning rate was set to 0.000002, the number of epochs to 20, and AdamW (Loshchilov and Hutter, 2017) was used as the optimizer. We did not conduct a detailed validation to select the best model from the existing models because our focus was on whether the task can be solved, not on achieving a high-performance model.

4.2 Evaluation metrics

The model constructed using company A’s data was tested using company B’s data. By testing models with data from a completely different population, we successfully verified the generality of the model under strict conditions. The purpose of this study is not so much to estimate the correct level of happiness as to assess whether changes in happiness are captured. For the evaluation index, we employed the Pearson correlation coefficient, which has been used in previous studies (Song and Zhao, 2023). Subsequently, the correlation coefficient between the estimated happiness score derived from the model (hereafter referred to as “estimated happiness score”) and the self-reported happiness score by employees was computed.

4.3 User-based Happiness Estimation

For each user, we calculated the correlation coefficient between the self-reported and estimated happiness scores. We tested the performance of the model trained on company A’s data in predicting the estimated level of happiness from the texts of individuals in company B. The eleven users in company B are U1 to U11, respectively. In this study, the model was applied to a completely different population.

5 Results

5.1 User-based happiness estimation

The correlation coefficients of the happiness estimation for each user in company B are shown in Table 3 and the time series is shown in Figure 3. The model performance is defined as the correlation coefficient between the self-reported and estimated happiness scores. In Table 3, we observed that the performance of the model in estimating the

[2] https://huggingface.co/docs/transformers/model_doc/auto
[3] https://github.com/cl-tohoku/bert-japanese

	Users	Reports	Average number of characters	Collection period
Company A	94	1,728	58.8	09/01/2022 ∼ 10/31/2022 (2-months)
Company B	11	652	72.2	12/19/2022 ∼ 05/18/2023 (6-months)

Table 1: Data Statistics

	Daily report text (Translated from Japanese)	Self-reported HS
(1)	The most significant advantage of remote work is being able to nap. Although rather conspicuous, I would like to be able to nap in the office as well.	7
(2)	Rushing on Monday mornings. Addressing complaints is difficult.	2
(3)	Although I had much work to perform since morning, I participated in a fun drinking party! I enjoyed interacting with some people I have not met before! The delicious food at a standing bar and the tasty wine were a great start to the week! It was incredibly fun.	10
(4)	Insufficient time... I am exhausted...	1
(5)	I had new insights for marketing. Additionally, I am continuing with the accident response.	5

Table 2: Examples of daily report text and self-reported happiness score (Self-reported HS). The data were confidential, so the text has been slightly modified. After translating the Japanese into English, parts of the diary have been deleted and expressions have been changed.

happiness of each user varied. The variations in the correlation coefficients reflect the efficacy of the estimation model in capturing fluctuations in happiness over time.

We discuss the results for the characteristic users. First, the correlation coefficient for U1 in Table 3 could not be calculated because the self-reported happiness score of U1 remained consistently at 8 (Figure 3, U1). For users who answered the same number in the questionnaire but not in the free-response form, different information can be obtained daily. Additionally, incorporating the behaviors of users who responded in such manner into the model may facilitate happiness estimation.

Next, one user indicated a negative correlation coefficient, although not significant (Figure 3, U3). We examined the daily report text of this user and discovered that the user frequently used the expression, “A nice day that...” in the daily report. This suggests that the performance of the estimation model may be affected by differences in the writing style of each user.

As shown in Figure 3, only one user (U5) answered 0 after April. Although the user initially provided diverse texts, after April, his entries were primarily zeros and the content became uniform. Some users could not easily provide accurate estimates after such environmental changes. Although the occurrence of events can complicate data acquisition, inadequate data may reflect an individual’s happiness level.

user	corr		user	corr	
U1	-		U7	0.63	**
U2	0.64	**	U8	0.63	**
U3	-0.05		U9	0.48	**
U4	0.59	**	U10	0.55	**
U5	0.40	*	U11	0.59	**
U6	0.40	*			

Table 3: Correlation coefficients for each user. U1 to U11 represent the 11 users of company B. Significant positive correlations were observed across most users, except for U1 and U3. Because the self-reported happiness score of U1 was always 8, we could not calculate the correlation coefficient and p-value. $^{*}p < .005$ and $^{**}p < .001$.

Finally, we assessed the happiness estimation for all users. Inter-individual variations in the factors influencing happiness can affect prediction accuracy. Because the training and test datasets used in this study encompass various distinct user profiles, the determinants of happiness may vary considerably. This suggests that incorporating user personality traits and writing habits into the model may improve estimation performance.

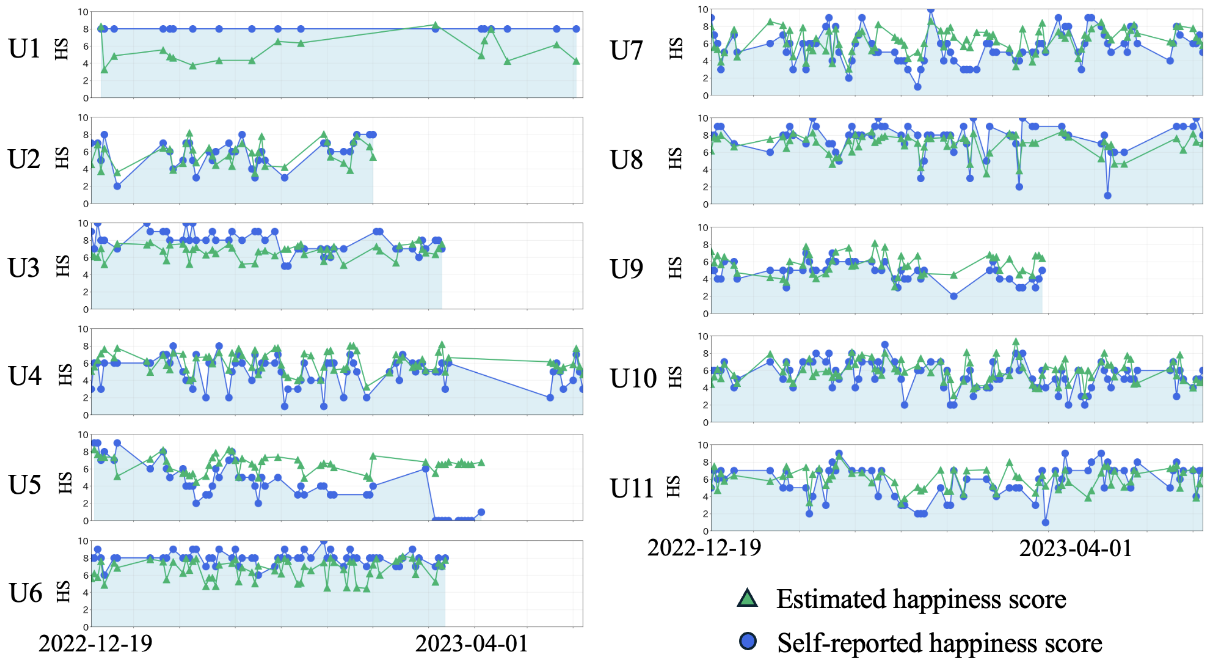


Figure 3: Estimation of happiness level of company B's user. Blank periods indicate that the individual did not complete the diary. Vertical axis represents happiness score (HS), whereas horizontal axis represents time. Background color in graphs highlight time series representing transition of self-reported happiness score.

5.2 Analysis

The performance of the model depends on the content of the user's daily report text. Considering user differences as intervening factors may increase the effectiveness of the happiness estimation models. Therefore, we explored the following features for users from U2 to U11, for which correlation coefficients could be computed.

Average length of sentences The model performance may have been influenced by the length of the diary entries. Previous studies showed that the length of sentences differs from author to author (Yule, 1939).

Lexical diversity The model performance may have been influenced by the lexical diversity. Research has demonstrated that lexical diversity differs among individuals (Gregori-Signes and Clavel-Arroitia, 2015).

Personality Differences in personality may cause differences in diaries. In this study, the following personality traits were collected during the experiment. This study examines items of **cooperativeness** and **autonomy**.

The results are visualized using box and scatter plots (Figure 4). In the scatter plots, the horizontal and vertical axes represent the model performance and feature, respectively. The box plot partitions the model performance into two sections: above- and below-median section.

Additionally, we obtained the correlation coefficients and p-values for each of the characteristics. The results showed a trend, although insignificant, which is likely due to the small sample size of 10 participants. Thus, the following discussion is limited to trends in the data, and we will attempt to obtain additional data to validate our analysis.

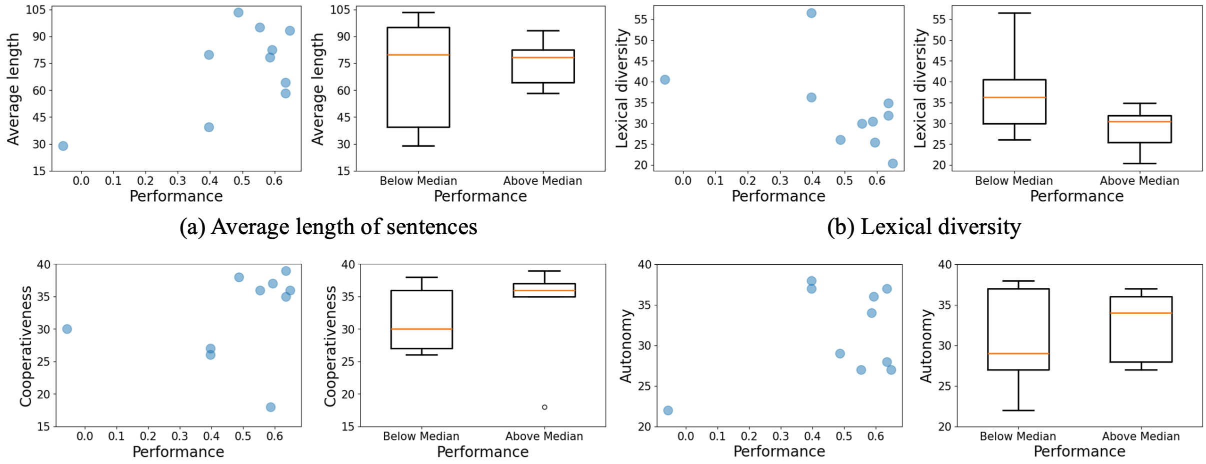
5.2.1 Average length of sentences

Longer sentences in the diary contain more information. We examined the effect of the average sentence length per user on the model performance (Figure 4 (a)). The results showed that the group with a longer average sentence performed better in estimating happiness. The median values remained consistent across both groups in the box plots.

The results suggest that a longer user diary corresponds to the likelihood of the model performing better. Furthermore, the longer the diary sentences, the more likely they are to contain more information concerning happiness.

5.2.2 Lexical diversity

We examined the relationship between the lexical diversity of each user's diary and the model performance. Lexical diversity is an evaluation



(a) Average length of sentences

(b) Lexical diversity

(c) Cooperativeness

(d) Autonomy

Figure 4: Scatter plot and box plot of factors and performance. Scatter plot and box plot of average length of sentences (a). Correlation coefficient was 0.63 ($p = 0.052$). Scatter plot and box plot of lexical diversity (b). Correlation coefficient was -0.51 ($p = 0.135$). Scatter plot and box plot of cooperativeness (c). Correlation coefficient was 0.28 ($p = 0.435$). Scatter plot and box plot of autonomy (d). Correlation coefficient was 0.38 ($p = 0.274$). In this figure, U1 is excluded due to the inability to calculate its model performance.

of whether a user writes different content in each diary. The diversity of each user's diary was evaluated using the Self-BLEU (Zhu et al., 2018) metric. Self-BLEU uses the BLEU score (Papineni et al., 2002) to evaluate the overall similarity between a group of three or more sentences. A high score indicates that the daily report texts are similar, whereas a low score indicates that the Daily report texts are diverse. In this experiment, we computed BLEU scores using SacreBLEU (Post, 2018).

The results showed that users with low self-BLEU scores performed better (Figure 4 (b)). This finding indicates that greater lexical diversity in the diaries is associated with improved happiness estimation. Moreover, this indicates that individual motivation may affect the model performance. Users who repeatedly inputted the same text from April onward (see Figure 3, U5) had the highest self-BLEU scores and low model performance. Given the diversity of the daily report text before April 1, the self-BLEU score was affected by the events after April. In April, the number of members changed and the number of employees increased. Self-BLEU scores are affected by events that complicate continuous daily reporting.

5.2.3 Cooperativeness

The relationship between cooperativeness and model performance was investigated. For cooperativeness, we used the total score of the answers to the 10 questions.

Four of the five users whose model performances were above the median were relatively highly cooperative within their group (Figure 4 (c)). Cooperative users are more likely to engage with others in the workplace while working. Increased interactions with many people suggest that there is more diversity in work and a greater likelihood of variation in diary content. Additionally, more cooperative users may have been more willing to participate in the experiment. Such users can employ strategies to enhance the content richness. However, users with the lowest cooperativeness were in the above-median model performance group, whereas users with the highest cooperativeness were in the below-median model performance group. Nonetheless, the trend suggests that the model performed well when the users were cooperative.

5.2.4 Autonomy

Next, we investigated the relationship between autonomy and model performance. We used the total score of the answers to the 10 autonomy questions.

U3, located at the bottom left of scatter plot in Figure 4 (d), has the lowest autonomy score and model performance. The lack of autonomy in creating new sentences each time may result in only similar sentences written, thus adversely affecting the model performance.

However, recognizing and accomplishing tasks contribute minimally to their annotation in journals.

As shown in the box plot, the users with the greatest and least autonomy belonged to the group of users exhibiting low performance. This suggests that autonomy does not affect happiness estimation as much as cooperativeness.

6 Conclusion and future works

In this study, we estimated happiness scores from diaries obtained longitudinally over six months using a transformer-based model. The analysis revealed a significant correlation between the estimated and self-reported happiness scores of 9 among 11 participants.

The findings of this study suggest three key avenues for future research and application.

First, happiness can be enhanced via targeted interventions. This study demonstrated the feasibility of estimating happiness levels in real time using free-text descriptions. Furthermore, when a shift in happiness occurs, free-text descriptions can provide insights into the underlying events or experiences.

Second, this study focused on estimating happiness scores from texts, as free texts contain abundant information beyond happiness. A nuanced examination of texts can provide insight into the determinants of happiness, the author's experiences, and the related emotions. Analyzing the characteristics of texts enables the measurement of happiness and various aspects that contribute to happiness, such as cooperativeness and autonomy.

Third, the efficacy of happiness estimation from texts is affected by individual traits, including writing style and personality. Understanding each user's diary patterns may enable a preliminary assessment of the happiness estimation performance of the model. This may determine whether another instruction is to be employed to predict the happiness of users with shorter or fewer diary entries more accurately. For example, setting requirements such as the minimum character count for diary entries can enhance the performance. The findings of this study emphasize the importance of considering individual user characteristics when estimating happiness from texts, which can provide a better assessment of happiness.

Limitation

Using this task as a substitute for conventional questionnaires presents five major challenges.

First, cases exist where the relationship between a specific event and the happiness level is different for each individual. For example, the statement "Today is a holiday."may elicit happiness in some people but not in others. To understand the relationship between holidays and happiness, a significant amount of information, such as the author's beliefs, values, and current situation, is required, in addition to sufficient texts.

Second, the model is not generalized. The model used in this study is based on data obtained from experiments conducted by two companies. The participants of this study were from company A, which is an advertising and marketing business company, and company B, which is an electronics manufacturer. Because the vocabulary varies with the company characteristics, more data should be obtained from diverse industries and companies.

Third, the data acquisition was biased. Because both datasets were obtained simultaneously, the respondents were able to describe their happiness level and free texts consistently.

Fourth, this study was conducted using a limited sample size. The findings were validated based on data obtained from 10 participants. To enhance the robustness and applicability of these findings, more data should be obtained from larger cohorts of users.

Fifth, our experiments were restricted to the Japanese language.

Acknowledgments

This research was supported by the Japan Science and Technology Agency, Japan Society for the Future, and JPMJMI21J2.

References

Hadley Cantril. 1965. *The Pattern of Human Concerns*. Rutgers University Press.

Mustafa Coşkun and Meltem Ozturan. 2018. europehappinessmap: A framework for multi-lingual sentiment analysis via social media big data (a twitter case study). *Information*, 9(5):102.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Ed Diener and Robert Biswas-Diener. 2002. Will money increase subjective well-being? *Social Indicators Research*, 57(2):119–169.

Ed Diener, Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. 1985. The satisfaction with life scale. *Journal of Personality Assessment*, 49(1):71–75.

Peter Sheridan Dodds, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PLoS ONE*, 6(12):e26752.

Hiina Domae, Masataka Nakayama, Kosuke Takemura, Yasushi Watanabe, Matthias S. Gobel, and Yukiko Uchida. 2024. Antecedents and consequences of telework during the covid-19 pandemic: a natural experiment in japan. *Humanities and Social Sciences Communications*, 11(1).

Carmen Gregori-Signes and Begoña Clavel-Arroitia. 2015. Analysing lexical density and lexical diversity in university students' written discourse. *Procedia - Social and Behavioral Sciences*, 198:546–556.

Hidehumi Hitokoto and Yukiko Uchida. 2014. Interdependent happiness: Theoretical importance and measurement validity. *Journal of Happiness Studies*, 16(1):211–239.

Gregory M. Hurtz and John J. Donovan. 2000. Personality and job performance: The big five revisited. *Journal of Applied Psychology*, 85(6):869–879.

Kazuhiro Ito, Shoko Wakamiya, Masae Manabe, Yasushi Watanabe, Masataka Nakayama, Yukiko Uchida, and Eiji Aramaki. 2023. Mood sharedness: Well-being indicator on the team status, rather than individual factors. Preprint on webpage at `https://www.researchsquare.com/article/rs-3666850/v1`.

Kokil Jaidka, Salvatore Giorgi, H. Andrew Schwartz, Margaret L. Kern, Lyle H. Ungar, and Johannes C. Eichstaedt. 2020. Estimating geographic subjective well-being from twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences*, 117(19):10165–10171.

Oscar N. E. Kjell, Sverker Sikström, Katarina Kjell, and H. Andrew Schwartz. 2022. Natural language analyzed with ai-based transformers predict traditional subjective well-being measures approaching the theoretical upper limits in accuracy. *Scientific Reports*, 12(1).

Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. *arXiv preprint*.

Lewis Mitchell, Morgan R. Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M. Danforth. 2013. The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, 8(5):e64417.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Matt Post. 2018. A call for clarity in reporting BLEU scores. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191, Belgium, Brussels. Association for Computational Linguistics.

S. Sametoğlu, D.H.M. Pelt, J.C. Eichstaedt, L.H. Ungar, and M. Bartels. 2023. The value of social media language for the assessment of wellbeing: A systematic review and meta-analysis. *The Journal of Positive Psychology*, 19(3):471–489.

H. Andrew Schwartz, Maarten Sap, Margaret L. Kern, Johannes C. Eichstaedt, Adam Kapelner, Megha Agrawal, Eduardo Blanco, Lukasz Dziurzynski, Gregory Park, David Stillwell, Michal Kosinski, Martin E.P. Seligman, and Lyle H. Ungar. 2015. Predicting individual well-being through the language of social media. In *Biocomputing 2016*. WORLD SCIENTIFIC.

Yuyan Shi, Lindsay E. Sears, Carter R. Coberley, and James E. Pope. 2013. Classification of individual well-being scores for the determination of adverse health and productivity outcomes in employee populations. *Population Health Management*, 16(2):90–98.

Mengyao Song and Nan Zhao. 2023. Predicting life satisfaction based on the emotion words in self-statement texts. *Frontiers in Psychiatry*.

Yasushi Watanabe, Kosuke Takemura, Masataka Nakayama, Matthias S Gobel, Hiina Domae, and Yukiko Uchida. 2023. Person-organization fit in japan: A longitudinal study of the effects of clan culture and interdependence on employee well-being. *Current Psychology*, 43(17):15445–15458.

Jiaqi Wu, Marilyn Walker, Pranav Anand, and Steve Whittaker. 2017. Linguistic reflexes of well-being and happiness in echo. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Association for Computational Linguistics.

G. Udny Yule. 1939. On sentence-length as a statistical characteristic of style in prose: With application to two cases of disputed authorship. *Biometrika*, 30(3/4):363.

Yaoming Zhu, Sidi Lu, Lei Zheng, Jiaxian Guo, Weinan Zhang, Jun Wang, and Yong Yu. 2018. Texygen: A benchmarking platform for text generation models. In *The 41st International ACM SIGIR Conference on Research amp; Development in Information Retrieval*, SIGIR '18. ACM.

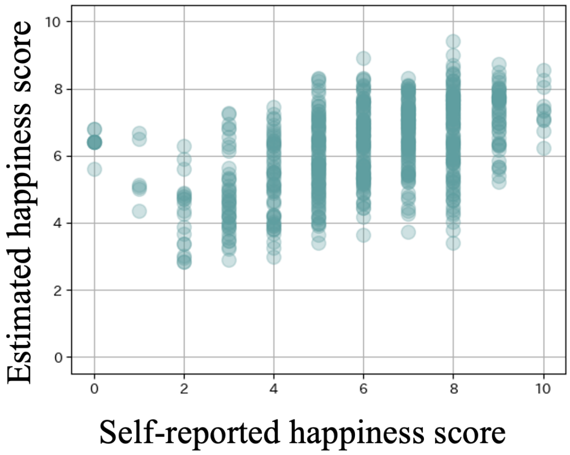


Figure 5: Results of workplace happiness estimation experiments. Vertical axes represent estimated and self-reported happiness scores, respectively. Correlation coefficient was 0.49 ($p < 0.001$).

A Appendix

A.1 Workplace-based Happiness Estimation

To estimate workplace happiness, the happiness level was estimated for each workplace and evaluated by determining the correlation coefficient. The correlation coefficient between the self-reported and estimated happiness scores was 0.49 (Figure 5).

	Questionnaire (Translated from Japanese)
(1)	I am concerned about others' perception about me.
(2)	When interacting with people in my company, I am concerned about the relationships and statuses between me and them.
(3)	I think it is important to maintain harmony among my work colleagues.
(4)	Avoiding disagreements with others in the workplace.
(5)	When I disagree with someone in my company, I typically accept that person's opinion.
(6)	I respect people who have a sense of humility in the workplace.
(7)	I would sacrifice my own interests for the good of the company to which I belong.
(8)	I typically feel that my social interactions with others at the workplace are more important than my accomplishments.
(9)	I feel that my destiny is intertwined with the destiny of others' at the workplace.
(10)	I may change my attitude or behavior at the workplace depending on another person or the situation.

Table 4: Questionnaire pertaining to cooperativeness. Each item was answered on a 5-point Likert scale ranging from 1 (not applicable at all) to 5 (extremely applicable). All questionnaires are in Japanese and have been translated into English.

	Questionnaire (Translated from Japanese)
(1)	I always try to form my own opinion within the company.
(2)	I do not mind being the sole recipient of accolades at the workplace.
(3)	I think the best decisions at the workplace are those made by myself.
(4)	I typically make decisions on my own at the workplace.
(5)	I behave in the same manner regardless of my status in the company.
(6)	I do not mind if my thoughts or actions differ from those of others at the workplace.
(7)	I always speak up for myself at the workplace.
(8)	Being independent at the workplace is extremely important to me.
(9)	I enjoy being unique and different from others at the workplace.
(10)	My actions at the workplace are not governed by others' perceptions.

Table 5: Questionnaire pertaining to autonomy. Each item was answered on a 5-point Likert scale ranging from 1 (not applicable at all) to 5 (extremely applicable). All questionnaires are in Japanese and have been translated into English.

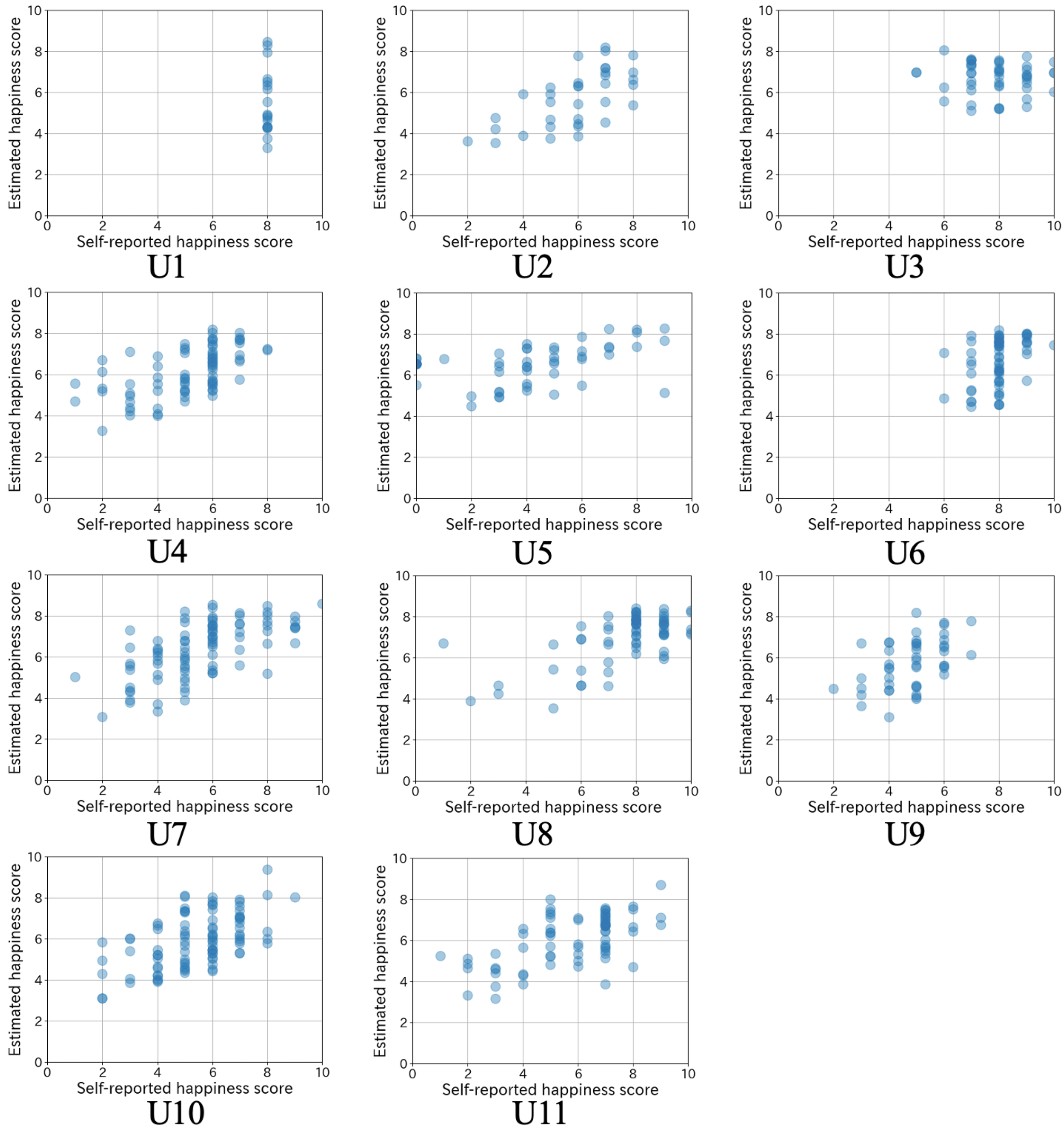


Figure 6: Scatter plot of company B user-based happiness estimation. The vertical axis is the estimated happiness score and the horizontal axis is the self-reported happiness score. The correlation coefficients for each user are in Table 3.