Prediction of life satisfaction from resting-state functional connectome

Takashi Itahashi1,# | Neda Kosibaty2,# | Ryu-Ichiro Hashimoto1,3 | Yuta Y. Aoki1

1 Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan
2 Graduate School of Advanced Mathematical Sciences, Meiji University, Tokyo, Japan
3 Department of Language Sciences, Graduate School of Humanities, Tokyo Metropolitan University, Tokyo, Japan

Correspondence
Yuta Y. Aoki, Medical Institute of Developmental Disabilities Research, Showa University, 6-11-11 Kita-karasuyama, Setagaya-ku, Tokyo 157-8577, Japan.
Email: yuaoki-tky@umin.ac.jp

#These authors contributed equally to this work.

Funding information
Japan Society for the Promotion of Science, Grant/Award Numbers: 16H06396, 18K03370, 21K15719

Abstract

Background: Better life satisfaction (LS) is associated with better psychological and psychiatric outcomes. To the best of our knowledge, no studies have examined prediction models for LS.

Methods: Using resting-state functional magnetic resonance imaging (R-fMRI) data from the Human Connectome Project (HCP) Young Adult S1200 dataset, we examined whether LS is predictable from intrinsic functional connectivity (iFC). All the HCP data were subdivided into either discovery (n = 100) or validation (n = 766) datasets. Using R-fMRI data in the discovery dataset, we computed a matrix of iFCs between brain regions. Ridge regression, in combination with principal component analysis and 10-fold cross-validation, was used to predict LS. Prediction performance was evaluated by comparing actual and predicted LS scores. The generalizability of the prediction model obtained from the discovery dataset was evaluated by applying this model to the validation dataset.

Results: The model was able to successfully predict LS in the discovery dataset (r = 0.381, p < .001). The model was also able to successfully predict the degree of LS (r = 0.137, 5000-repetition permutation test p = .006) in the validation dataset, suggesting that our model is generalizable to the prediction of LS in young adults. iFCs stemming from visual, ventral attention, or limbic networks to other networks (such as the dorsal attention network and default mode network) were likely to contribute positively toward predicted LS scores. iFCs within ventral attention and limbic networks also positively contributed to predicting LS. On the other hand, iFCs stemming from the visual and cerebellar networks to other networks were likely to contribute negatively to the predicted LS scores.

Conclusion: The present findings suggest that LS is predictable from the iFCs. These results are an important step toward identifying the neural basis of life satisfaction.

KEYWORDS
Human Connectome Project, intrinsic functional connectivity, life satisfaction, predictive model
1 | INTRODUCTION

Good subjective well-being (SWB) is protective against the psychological impacts of major life events and is related to longevity and a lower chance of physical illness (Diener & Chan, 2011; Luhmann et al., 2012; Wood & Joseph, 2010). Because of these advantages, achieving better SWB has been an important issue for policymakers since the 1980s (Diener, 1984; Layard, 2010). However, because of its complex multifaceted nature, improving SWB is not easy. One of the most important components of SWB is life satisfaction (LS), defined as “a cognitive and global evaluation of the quality of one’s life as a whole” (Pavot & Diener, 2008). LS is associated with a wide variety of psychological and psychiatric sequelae, such as self-esteem (Diener & Marissa, 2009) and even suicide (Koivumaa-Honkanen et al., 2001).

Despite such importance, LS itself is complex and changes with life course; thus, only a few studies have examined its neural basis (Kong, Ding, et al., 2015; Waldinger et al., 2011). A morphometric study showed that LS is positively associated with regional gray matter volume in the right parahippocampal gyrus and negatively correlated with gray matter volume in the left precuneus and ventromedial prefrontal cortex (Kong, Ding, et al., 2015). A functional magnetic resonance imaging (fMRI) study showed that higher LS is associated with stronger functional connectivity within the emotion-processing network (Waldinger et al., 2011). These prior studies showed that unique neural correlates of LS can be found within their respective participant groups. However, a generalizable common neural basis for LS has not yet been conclusively demonstrated.

Given the importance of LS to SWB, the consequences of low LS, and the relevance of LS to individuals worldwide, whether or not LS has a generalizable neural basis should be determined. In the present study, we examine whether a predictive model for LS can be constructed from intrinsic functional connectivity (iFC) data and subsequently examine whether the neural basis of LS in one group of healthy adults is generalizable to a larger, separate healthy adult group.

2 | MATERIALS AND METHODS

2.1 | Human Connectome Project dataset

We used the Human Connectome Project (HCP) Young Adult dataset, comprising MRI data collected from 1200 young adult participants (22–35 years old) (Glasser et al., 2013). We then removed subjects who did not complete two resting-state fMRI (R-fMRI) sessions, for whom psychological well-being measures were not obtained, or who exhibited excessive head motions during the MRI scans. After these exclusion criteria, 866 subjects remained. Because this dataset contained biological siblings, there was a possibility for biasing the prediction model due to the genetic and shared environmental factors among siblings. The HCP dataset provided a list of subjects unrelated to other subjects. To avoid undesirable information leakage in constructing the prediction models, we used 100 unrelated subjects (male: \( n = 46 \), female: \( n = 54 \); LS: 54.47 ± 9.00 [mean ± SD]) out of the 866 for the discovery dataset. The remaining 766 subjects were used for the validation dataset (male: \( n = 363 \), female: \( n = 403 \); LS: 54.79 ± 8.89 [mean ± SD]). The degree of general LS in HCP participants was measured using the National Institutes of Health (NIH) toolbox (Salsman et al., 2014). We note that the overarching aim of the current study is building the generalizable model to predict LS. Thus, we eschewed adjusting potential confounding demographic factors, such as race and income. In other words, we intended to accept heterogeneity of the participants, while aiming to achieve the acceptable prediction performance.

2.2 | R-fMRI data processing and network construction

We obtained minimally preprocessed R-fMRI data from the publicly available HCP database (Glasser et al., 2013). Using an in-house MATLAB code, we applied additional processing procedures to the R-fMRI data, including the removal of the first 10 s of data for each run and nuisance regression on the data. Nuisance regressors comprised linear detrending, six head motion parameters, and averaged signals from subject-specific white matter, ventricle, and gray matter masks, as well as their derivatives. A band-pass filter (0.008–0.1 Hz) was then applied to the residuals. Frame-wise displacement (FD) was calculated to identify motion-contaminated volumes, and a scrubbing method with an FD threshold of 0.5 mm was applied to reduce spurious changes in iFCs due to subtle head motion during the scans (Power et al., 2012).

We characterized each individual’s whole-brain functional connectome using 427 regions of interest (ROIs); 400 surface-based cortical regions (Schaefer et al., 2018), 17 subcortical regions (Fischl et al., 2002), and 10 functionally parcellated cerebellar regions (King et al., 2019). Pearson correlation coefficients were calculated between all possible pairs of ROIs, yielding a 427 × 427 iFC matrix for each participant. Fisher’s r-to-z transform was further applied to each correlation coefficient. These procedures yielded 90,951 unique iFCs, excluding the diagonal elements of the iFC matrix. The lower triangular portion of the iFC matrix was vectorized and concatenated across participants, resulting in a 100 × 90,951 feature matrix. To improve the interpretability of our findings, we characterized the 427 ROIs into nine resting-state network (RSN) labels comprising seven RSNs previously described (Yeo et al., 2011), as well as basal ganglia (BG) and cerebellar (CER) networks.

2.3 | Constructing the prediction model for LS using iFCs

To test the association between iFC and the degree of LS, we constructed a prediction model for LS using ridge regression in combination with principal component (PC) analysis (PCA) and a 10-fold cross-validation (CV) approach (Hoerl & Kennard, 1970). As shown in Figure 1, we first applied PCA to the discovery dataset to reduce the dimensionality of features into 99 PC scores. The transformation
matrix obtained from the discovery dataset was also applied to the validation dataset, which, it should be noted, had not been used to obtain the transformation matrix. Then, the discovery dataset was divided into 10 groups, or folds, for 10-fold CV. We used nine folds to construct the ridge regression prediction model, and the remaining fold was used to test the efficacy of the constructed model. The hyperparameter was optimized within the internal loop; that is, the test fold data were not used to optimize the hyperparameter. To evaluate the prediction performance of the obtained model, we calculated the Pearson correlation coefficient between averaged predicted and actual scores.

### 2.4 Permutation test for assessment of prediction performance

To assess the statistical significance of the performance of the model in predicting LS, we used a permutation test with 5000 iterations. Briefly, we shuffled LS scores at each iteration and constructed a prediction model with 10-fold CV as described above. To assess the null models’ prediction performance, we computed the Pearson correlation coefficient between predicted and actual scores, yielding a null distribution from which the $p$ value could be calculated. Statistical significance was set to $p < .05$.

### 2.5 Generalization of the prediction model for the validation dataset

The generalizability of the prediction model was tested using the validation dataset ($n = 766$). As we had obtained 10 prediction models from 10-fold CV, we applied these prediction models to the validation dataset and computed the Pearson correlation coefficient between the averaged predicted and actual scores. Again, we constructed the null distribution using prediction models obtained by permutation test with 5000 iterations. Statistical significance was set to $p < .05$.

### 2.6 Contribution of iFCs to predicted LS

We measured the extent to which iFCs contributed to the prediction of LS by calculating the weight contribution of each iFC. As both PCA and ridge regression are linear methods, the contribution of each iFC was calculated by multiplying the PCA-derived transformation matrix and the regression coefficient matrix of our prediction models, yielding 10 weights for each iFC. The contribution of each iFC was assessed using one-sample $t$-tests, with statistical significance set to $p < .05$ adjusted with a Bonferroni correction for multiple comparisons (90,951 comparisons).

To further improve the interpretability of our findings, we examined network anatomy in a similar manner to previous studies (Barron et al., 2020; Lake et al., 2019). Briefly, we computed the probability that iFCs are shared between the networks identified by our prediction model and within or between nine canonical RSNs. Statistical significance was determined using a hypergeometric cumulative distribution function after applying Bonferroni correction for 45 comparisons.

### 3 RESULTS

#### 3.1 Prediction performance in the discovery dataset

The model was able to successfully predict the degree of LS reported in healthy young adults by using PC scores derived from iFCs ($r = 0.381$, 0.381,
FIGURE 2  Scatter plots of actual versus model-predicted life satisfaction (LS) scores. For the discovery (a) and validation (b) datasets, Pearson correlation coefficients were calculated between actual and predicted LS scores. Permutation tests with 5000 iterations were conducted to examine the statistical significance. Statistical threshold was set to $p < .05$

5000-repetition permutation test $p < .001$) in the discovery dataset of 100 individuals (Figure 2a).

3.2 Prediction performance in the validation dataset

Having determined that our obtained model could successfully predict the degree of LS in the discovery dataset, we applied this prediction model to the validation dataset. As we obtained 10 models through the 10-fold CV, we applied each of these models to the validation dataset and considered their averaged scores as the predicted scores for the validation dataset. The model was also able to successfully predict the degree of LS ($r = 0.137, 5,000$-repetition permutation test $p = .006$) in the validation dataset of 766 individuals, suggesting that our model is generalizable to the prediction of LS in young adults (Figure 2b).

3.3 Contribution of iFCs to the prediction performance

We computed the contribution of each iFC by multiplying the PCA-derived transformation matrix and the weights of prediction models. Statistical analyses revealed that 6164 iFCs were significant contributors to the prediction of LS ($p < .05/90,951$), comprising 3479 iFCs contributing positively to predicted LS scores and 2685 iFCs contributing negatively to predicted LS scores (Figure 3). Qualitatively, iFCs stemming from the default mode network (DMN) were the dominant contributors to the prediction of LS.

To further examine the model’s network anatomy, we computed $p$ values between our network and within or between nine canonical RSNs (with a significance level of $p < .025/45$) for iFCs contributing positively and negatively to LS prediction. Among iFCs contributing to LS, iFCs stemming from visual, ventral attention (VAN), or limbic networks to other networks (such as the dorsal attention network [DAN] and DMN) were likely to contribute positively toward predicted LS scores (Figure 4a). iFCs within VAN and limbic networks also positively

FIGURE 3  Overview of intrinsic functional connectivities (iFCs) contributing to the prediction of life satisfaction. The left panel shows the matrix, each element of which represents $t$-statistics. The red and blue colors indicate that iFCs positively and negatively contributed to predict the life satisfaction (LS) score, respectively. The middle and left upper panels show the connectogram of iFCs contributing to the prediction of LS scores. The middle and left bottom panels show the brain maps.
CONTRIBUTED TO PREDICTING LS. ON THE OTHER HAND, IFCS STEMMING FROM THE VISUAL AND CEREBELLAR NETWORKS TO OTHER NETWORKS (SUCH AS THE DAN, DMN, BG, AND LIMBIC NETWORKS) WERE LIKELY TO CONTRIBUTE NEGATIVELY TO THE PREDICTED LS SCORES (FIGURE 4B).

4 DISCUSSION

ANALYZING A SIZABLE R-fMRI DATASET IN ADULTS, WE DEMONSTRATED THAT LS SCORES IN HEALTHY YOUNG ADULTS COULD BE PREDICTED BY A MODEL BASED ON IFCS. FURTHERMORE, THE LS PREDICTION MODEL CONSTRUCTED USING DATA FROM ONE GROUP OF HEALTHY PEOPLE COULD BE USED TO PREDICT THE LS OF ANOTHER GROUP OF HEALTHY PEOPLE, SUGGESTING A GENERALIZABLE NEURAL BASIS FOR LS.

TO ACHIEVE THIS NOVEL FINDING, WE FIRST CONSTRUCTED A PREDICTION MODEL USING IFCS OF 100 HEALTHY PARTICIPANTS. WE APPLIED PCA AND 10-FOLD CV IN THE PREDICTION MODEL, WHICH WERE THEN APPLIED TO THE VALIDATION DATASET. ALTHOUGH THE CORRELATION BETWEEN THE ACTUAL AND PREDICTED SCORES WAS WEAKER IN THE VALIDATED DATASET THAN THAT IN THE DISCOVERY DATASET (FIGURE 2), THE FINDINGS SHOWED HIGH GENERALIZABILITY OF THE PREDICTION MODEL.

SIGNIFICANTLY, THE PRESENT STUDY ESTABLISHED A PROTOCOL TO PREDICT LS FROM R-fMRI. THIS RESEARCH WAS INSPIRED BY CONNECTOME-BASED PREDICTIVE MODELING (CPM), WHICH USES LARGE-SCALE NEUROIMAGING DATA TO PREDICT INDIVIDUAL DIFFERENCES IN TRAITS AND BEHAVIOR (SHEN ET AL., 2017). UTILIZING CPM, A NUMBER OF STUDIES HAVE SUCCESSFULLY PREDICTED ATTENTION, ANXIETY, AND MOTHER–INFANT BONDING (REN ET AL., 2021; RUTHERFORD ET AL., 2020; YOO ET AL., 2018). IN LINE WITH THE CONCLUSIONS OF THESE PRIOR STUDIES, THE PRESENT FINDINGS SUGGEST THAT CHANGES IN FUNCTIONAL CONNECTIVITY MAY IMPROVE LS. ADDITIONALLY, THE PRESENT RESULTS PROVIDE INSIGHT INTO THE FUNCTIONAL CONNECTIVITY PATHWAYS ON WHICH WE SHOULD FOCUS.

FUTURE RESEARCH COULD INVESTIGATE POTENTIAL MECHANISMS OF INTERVENTION TO ENHANCE LS.

THE CURRENT FINDINGS SHOULD BE INTERPRETED IN THE CONTEXT OF SEVERAL PRIOR NEUROIMAGING STUDIES THAT EXAMINED THE BRAIN-LS RELATIONSHIPS. KONG, WANG, ET AL. (2015) SHOWED THAT LS CAN BE PREDICTABLE WITH REGIONAL HOMOGENEITY (ReHo). KONG ET AL. AND THE CURRENT STUDY SHARE A PART OF FINDINGS. NAMELY, THEY IDENTIFIED THAT DORSAL ANTERIOR CINGULATE CORTEX (ACC) WAS RELATED TO LS, WHILE THE DORSAL ACC WAS INCLUDED IN THE CURRENT LIST OF NODES. HOWEVER, THERE ARE SOME IMPORTANT DIFFERENCES IN THE ASSUMPTION. FIRST, WE DID NOT INTEND TO IDENTIFY ONE BRAIN REGION OR A FEW AS A NEURAL CORRELATE FOR LS, BUT WE DID ASSUME THAT IFCS PATTERNS WOULD SERVE AS A NEURAL CORRELATE OF LS. SECOND, WE ADOPTED IFCS AS A MARKER FOR THE LS BECAUSE WE ASSUMED THAT NOT ONLY SHORT-DISTANCE BUT ALSO LONG-DISTANCE CONNECTIVITY WOULD ALSO WORK AS NEURAL CORRELATES. IN CONTRAST, ReHo FOCUSES ON THE LOCAL CONNECTIVITY. ANOTHER STUDY BY WALDINGER ET AL. (2011) ADOPTED FUNCTIONAL CONNECTIVITY, BUT...
they focused on some brain regions by setting ROIs. However, they split the participants based on LS into high LS and low LS groups. In contrast, the current study assumed that LS was not categorical but dimensional. In this context, we identified a large number of iFCs contributing significantly to the prediction of LS (Figure 3), despite employing stringent corrections for multiple comparisons. The number of iFCs significantly involved in LS prediction is most extensive in the DMN, which may be driven by the fact that the DMN possesses the largest number of ROIs. Indeed, at the network level, the DMN did not frequently reach statistical significance (Figure 4). Instead, the VAN and limbic system both showed a significant contribution to LS. Given that the VAN and limbic system play roles in reorienting the attention (Vossel et al., 2014) and in motivation and emotional processing (Mogenson et al., 1980), these components of cognition may be associated with LS. It should be noted that these networks comprise several structures, that is, nodes in the present analyses. However, we do not delve into which nodes in these networks are involved, as this is beyond the scope of the present study; additionally, the large-scale brain network concept assumes that the network per se serves as the neural basis for these cognitive components. The present findings are indirectly consistent with prior psychological studies that reported a link between LS and emotion and attention (Bastian et al., 2014; Diener et al., 2012). However, LS is not a simple combination of emotion and attention but is more complex and is related to culture, society, and the environment (Chen et al., 2015; Schimmack et al., 2002). In this regard, we need to emphasize that although the present study aimed to show the generalizability of our prediction model for LS, the current study did not overcome the impact of the difference in culture, society, and the environment as the data were obtained in a single-center in the states. Future large-scale international collaborative study is expected to overcome the limitation. The present findings have some more limitations. First, we used a cross-sectional dataset to examine the associations between iFCs and LS. Although we used the validation dataset to confirm the main findings, the causal relationship between iFCs and LS remains unclear. Future longitudinal investigations are necessary to elucidate causal relationships. Second, LS scores were self-reported, which may introduce bias. However, LS is intrinsically self-reported as a measure of how he or she subjectively feels, regardless of objective evaluation by others. Finally, although we paid much attention to avoiding undesirable information leakage in constructing the prediction model, demographic characteristics that might potentially be related to LS, such as life events (Luhmann & Eid, 2009), medical status and religiosity (Levin et al., 1995), in the present participant groups were not available. However, given that the present study aimed to show the generalizability of the prediction model, the potential heterogeneity of the present participants would further support the generalizability of our model.

5 | CONCLUSION

We successfully predicted LS in unrelated healthy young adults using a prediction model constructed from R-fMRI data. Additionally, the obtained model successfully predicted the LS scores in a further validation dataset, suggesting that LS has generalizable neural basis. The present findings are a step toward future intervention strategies to enhance LS, which could potentially bring the wide variety of benefits associated with good SWB to people worldwide.

ACKNOWLEDGMENTS

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. This work was partially supported by the JSPS KAKENHI (16H06396 to RH, 19K03370, and 19H04883 to TI, and 18K15493 and 21K15719 to YYA).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The authors declare that the data supporting the findings of this study are available within the article with DOI 10.1016/j.neuroimage.2013.04.127

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/brb3.2331

ORCID

Yuta Y. Aoki https://orcid.org/0000-0001-5617-4948

REFERENCES

Barron, D. S., Gao, S., Dadashkarimi, J., Greene, A. S., Spann, M. N., Noble, S., Lake, E. M. R., Krystal, J. H., Constable, R. T., & Scheinost, D. (2020). Trans-diagnostic, connectome-based prediction of memory constructs across psychiatric disorders. Cerebral Cortex, 31, 2523–2533. https://doi.org/10.1093/cercor/bhaa371

Bastian, B., Kuppens, P., De Roover, K., & Diener, E. D. (2014). Is valuing positive emotion associated with life satisfaction? Emotion, 14(4), 639–645. https://doi.org/10.1037/a0036466

Chen, X., Morin, A. J. S., Parker, P. D., & Marsh, H. W. (2015). Developmental investigation of the domain-specific nature of the life satisfaction construct across the post-school transition. Developmental Psychology, 51(8), 1074–1085. https://doi.org/10.1037/a0039477

Diener, E. (1984). Subjective well-being. Psychological Bulletin, 95(3), 542–575. https://doi.org/10.1037/0033-2909.95.3.542

Diener, E., & Chan, M. Y. (2011). Happy people live longer: Subjective well-being contributes to health and longevity. Applied Psychology: Health and Well-Being, 3(1), 1–43. https://doi.org/10.1111/j.1758-0854.2010.00145.x

Diener, E., Inglehart, R., & Tay, L. (2012). Theory and validity of life satisfaction scales. Social Indicators Research, 112(3), 497–527. https://doi.org/10.1007/s11205-012-0076-y

Diener, E., & Marissa, D. (2009). Cross-cultural correlates of life satisfaction and self-esteem. Journal of Personality and Social Psychology, 68(4), 653–663.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain...
