

# Sleep quality, ability to concentrate, and fruit consumption suffice to predict young adult well-being: Findings from Bayesian predictive projection

Adam Bartonicek (✉ [adam.bartonicek@postgrad.otago.ac.nz](mailto:adam.bartonicek@postgrad.otago.ac.nz))

University of Otago <https://orcid.org/0000-0002-4528-5543>

Shay Ruby Wickham

University of Otago

Narun Pat

University of Otago

Tamlin S. Conner

University of Otago

---

## Research article

**Keywords:** Psychological well-being, Young adults, Health habits, Sleep, Diet, Exercise, Variable selection

**Posted Date:** June 1st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-28472/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Public Health on April 9th, 2021. See the published version at <https://doi.org/10.1186/s12889-021-10690-3>.

# Abstract

## Background

The search for reliable lifestyle predictors of psychological well-being is a key research pursuit in psychology and public health. Researchers often gather data on many candidate predictors and then are faced with the difficult task of variable selection – selecting only few “relevant” predictors from the large set of candidates. Many statistical methods that are currently being used for variable selection suffer from known shortfalls. Fortunately, there are newer methods available. The aim of the study is to illustrate the use of predictive projection – a novel Bayesian method for variable selection with many attractive properties – and identify the predictors of psychological well-being in a sample of young adults.

## Methods

The sample consisted of 791 young adults (ages 18 to 25, 27.4% men) from New Zealand who had taken part in the Dunedin Daily Life Study in 2013–2014. For 13 days, participants completed an online daily diary assessment of their well-being and a range of modifiable lifestyle variables. Demographic characteristics were also measured. Participants’ diary data was averaged across days and analyzed cross-sectionally to identify predictors of average differences in flourishing. Predictive projection was used to select as few predictors as necessary to approximate the predictive accuracy of a large reference model with 28 predictors. Lastly, predictive projection was compared with other standard methods of variable selection.

## Results

Three predictors were sufficient to make average daily flourishing to approximate the predictions of the reference model: higher sleep quality, less trouble concentrating, and more servings of fruit. Demographic factors were not selected into the model. The projected submodel performed adequately on the test data. Compared to other variable selection methods, projection had either slightly worse or matching predictive performance, but led to much more parsimonious inference.

## Conclusion

Predictive projection was used to quickly arrive at a parsimonious submodel that approximated the predictive accuracy of the reference model. The variables selected into the submodel – felt refreshed after waking up, had less trouble concentrating, and ate more servings of fruit – were all theoretically meaningful. Our findings have broad implications for well-being research and public policy.

# Background

The search for reliable predictors of well-being is an important research topic in psychology and public health. Related to this topic are questions such as, what factors help people to become happier and experience greater flourishing in their lives? What are the relative strengths of these factors, and are they modifiable and therefore potential targets of intervention? The search for modifiable predictors of well-being is important because subjective well-being has real-world consequences. There is now robust evidence that well-being is strongly linked to better physical health and greater longevity (Diener & Chan, 2011; Kern et al., 2014; Keyes, 2007; Keyes et al., 2010; Keyes & Simoes, 2012; Pressman & Cohen, 2005). Thus, understanding what makes people happy is not only a goal in itself, but also a concern for public health and policy (Tay & Kuykendall, 2013).

Past studies have identified many candidate predictors of well-being, however, most of these may be difficult to modify on an individual level. On a nation-wide level, well-being has been linked with gross socioeconomic trends and policies, such as population density, income inequality, and strength of social welfare systems (Diener et al., 2018; Huppert & So, 2013; Li & Kanazawa, 2016; Oishi et al., 2011; Tay & Kuykendall, 2013). At the community and family level, well-being is associated with childhood socioeconomic status (SES; Pinquart & Sörensen, 2000), and the quality of social relationships (Pinquart & Sörensen, 2000). At the individual level, well-being is linked with achievement and self-efficacy (Howell, 2009; Ouweneel et al., 2011), faith and life purpose (Adams et al., 2000; Koenig, 2009; T. B. Smith et al., 2003), and substance abuse (Fink, 2014; Visser & Routledge, 2007), among others. While a person might potentially try and succeed in changing some of these factors, for example a substance abuse habit, the factors are likely to be resistant to long-term change and so changing them may be challenging (see e.g. Prochaska et al., 2004).

Particularly interesting from the point of psychological interventions are modifiable individual-level health habits that influence well-being, such as sleep, diet, and exercise. Among these health habits, sleep quality has been consistently shown as one of the strongest predictors of well-being (Kawada et al., 2009; Pilcher et al., 1997; Prendergast, Schofield, et al., 2016; Ridner et al., 2016; Wallace et al., 2017).

Particularly interesting is that while sleep quality shows an association with well-being, sleep quantity does not (Pilcher et al., 1997). Further, diet quality has also been linked with well-being. Specifically, fruit and vegetable consumption is associated with greater well-being in observational (Adams & Colner, 2008; Blanchflower et al., 2013; Conner et al., 2015; Jacka et al., 2010; Moreno-Agostino et al., 2019; Piqueras et al., 2011; Prendergast, Schofield, et al., 2016), prospective (Jacka et al., 2011), micro-longitudinal or daily diary (Jacka et al., 2011; Liao et al., 2018; Wahl et al., 2017; White et al., 2013) and experimental studies (Conner et al., 2017; Jacka et al., 2017). Conversely, poor diet as indicated by soft drink and fast-food consumption is also linked with lower well-being (Hong & Peltzer, 2017; Moreno-Agostino et al., 2019; Prendergast, Schofield, et al., 2016). Lastly, more exercise has been linked to greater well-being (Diener et al., 2018; Goodwin, 2003; Hassmén et al., 2000; Prendergast, Schofield, et al., 2016). However, which of these factors is more important, or whether they matter at all when compared to non-modifiable factors such as childhood SES, is not known.

An important challenge when investigating the predictors of well-being is how to identify the truly important and reliable predictors from a large set of candidate predictors. Ideally, the goal of a researcher is to construct the smallest, most parsimonious model with few relevant predictors that can explain as much of the variance in well-being as possible. However, with this goal in mind, researchers run into two complex statistical issues – *variable selection* and *overfitting*.

*Variable selection* is the process of finding a small subset of “relevant” predictors that can maximally predict the outcome measure (Heinze et al., 2018; Höge et al., 2018; Müller et al., 2013). Traditional regression modelling assumes that the exact set of the predictors to be included in the model is known a-priori. However, often the exact set of relevant predictors is *not* known a-priori (as is the case when it comes to searching for predictors of well-being) and so the issue of variable selection commonly arises across many fields (Heinze et al., 2018; Kadane & Lazar, 2004; van Erp et al., 2019). Under perfect circumstances, variable selection should be performed by fitting all possible models (all possible combinations of predictors) and selecting the best one. However, with an increasing number of predictors and model complexity, fitting all possible models quickly becomes computationally unfeasible – with  $K$  predictors there are  $2^K$  possible models to fit (Heinze et al., 2018). Even when an optimal subset of predictors is found, another issue becomes how to do inference with the selected model, i.e. determine the “significance” of model parameters, given that it is now necessary to condition on the selection procedure, and this is known as post-selection inference (Taylor & Tibshirani, 2015). Despite the fact that psychology and public health researchers have been dealing with the variable selection and post-selection inference for a very long time (Heinze et al., 2018), the two issues are by no means “solved” and remain hot areas of research in statistics and machine learning (Benjamini, 2010; Taylor & Tibshirani, 2015).

When doing variable selection, one also always runs the risk of *overfitting* – selecting predictors that fit to pure noise to the data (Höge et al., 2018; James et al., 2013). Overfitting occurs because extra predictors always make a model more flexible, and thus even irrelevant predictors can seemingly improve a model’s fit (James et al., 2013). However, while overfit models seem to explain the data at hand well (as indicated by e.g. high values or low root-mean-squared error; RMSE), they show poor predictive performance on new data (Hawkins, 2004). Additionally, the fact that overfit models often fail to generalize to new data may be one of the less known causes of the replication crisis (Yarkoni & Westfall, 2017). Research findings may fail to replicate not because of any ill-will of the researchers but because they are upheld by fragile trends idiosyncratic to the training data, resulting from overfitting. The issue of overfitting is particularly relevant to psychological and public health research in that the goal often is to find reliable and modifiable individual-level predictors of well-being (or health outcomes), and the results of studies are often evaluated based on in-sample explanatory power ( $R^2$ ).

The issues of *variable selection* and *overfitting* are just as relevant in the well-being literature as in other fields. There are three traditional regression-based approaches for identifying predictors of well-being. The first traditional approach is to fit many simple regression models, regressing the various candidate predictors on well-being scores individually, and then compare the size or significance of the predictor

slopes across these models (e.g. Prendergast, Schofield, et al., 2016). The main drawback of this simple regression approach is that it can only be used to assess the effect of each candidate predictor in isolation, without determining whether the effects remain after controlling for the other predictors. The second approach is to fit a single large multiple regression model with all predictors and then compare the size or significance of the predictor slopes within this one model. However, a challenge with this large multiple regression model approach is that when two or more predictor slopes are correlated (as is almost always the case with many candidate predictors), the slope and their estimated errors can be wildly distorted due to multicollinearity (Alin, 2010; Graham, 2003), and thus become uninterpretable. The last traditional regression approach is to fit a single multiple regression model with several predictors, manually picking the predictors based on theoretical considerations or using an automatic stepwise variable selection, often based on p-values (e.g. Boyes et al., 2011; Etxeberria et al., 2019; Keyes & Simoes, 2012; Sletta et al., 2019). Even with strong theory, however, there are many possible choices as to which predictors to include and exclude (Gelman & Loken, 2013), and, while it may not be common knowledge, stepwise selection does not account for the selection process and as a result fails to adequately control type-I error and tends to overfit (Flom & Cassell, 2007; Kuhn & Johnson, 2013; Smith, 2018; Steyerberg et al., 2001; Whittingham et al., 2006).

*Predictive projection feature selection* is a novel method for variable selection within the Bayesian framework (Piironen et al., 2018) that helps to stem issues related to variable selection and overfitting. The method consists of two steps. At the first step, a flexible reference model is fitted, using all candidate predictors. The reference model can be a “simple” multiple regression model, or a more complex, flexible model such as random forest, or even a neural network (Afrabandpey et al., 2019; Piironen et al., 2018). At the second step, smaller submodels with subsets of the candidate predictors are sequentially fitted via projection to approximate the reference model’s predictions. After that, it is possible to select the smallest possible submodel which makes predictions that are “similar enough” to those of the reference model (i.e. within some uncertainty bound such as standard error; Piironen et al., 2018). To avoid overfitting, the submodels can be compared with the reference model on cross-validated prediction accuracy, most conveniently via the Bayesian approximation of leave-one-out cross-validation – Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO; Vehtari, Gelman, & Gabry, 2015; 2017). The key advantages of predictive projection is that it can select a parsimonious model with good predictive accuracy, is robust to overfitting, and also has a valid posterior and thus it can be used for inference just like any other Bayesian model (Piironen et al., 2018).

In the present study, we used predictive projection feature selection to construct a parsimonious model that would predict well-being in a sample of 791 young adults. The young adults had participated in the 2013 and 2014 waves of the Daily Life Study, a daily diary study of the health and well-being of young adults in Dunedin, New Zealand. Well-being was surveyed every day for 13 days using the 8-item Flourishing Scale (Diener et al., 2010) adapted for daily measurement. We first fitted a large Bayesian multiple regression reference model, predicting average daily flourishing cross-sectionally from 28 candidate well-being predictors, including demographic and background variables (e.g., age, gender, childhood SES, BMI), and an extensive range of lifestyle variables also assessed in the daily diary related

to stress (stress, most stressful event of the day), somatic symptoms (e.g. tiredness, lack of ability to concentrate), diet-related variables (e.g. fruit and vegetable consumption), and health habit variables (e.g. sleep quality/quantity, exercise). Then, we used predictive projection to find a smallest possible submodel that would predict average daily flourishing almost as well as the reference model.

## Method

### *Software and packages*

Throughout our entire workflow, we used the programming language R, version 3.6.3 (R Core Team, 2019). For data wrangling and visualization, we used the tidyverse package (Wickham, 2017). For general-purpose Bayesian modelling, we used the brms (Bürkner, 2017) – a package that provides an accessible interface to the state-of-the-art Bayesian statistical programming language Stan (Carpenter et al., 2017). For predictive projection feature selection, we used the projpred package (Piironen et al., 2018). For Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO), we used the loo package (Vehtari et al., 2017), and for model checking and visualization, we used the bayesplot package (Gabry et al., 2019). Additional packages used include pander (Daróczy & Roman, 2018), naniar (Tierney et al., 2020), cowplot (Wilke, 2019), ggcorrplot (Kassambara, 2019), and labelled (Larmarange, 2020).

### *Participants and procedure*

Participants came from the Daily Life Study, a daily diary study assessing the psychological well-being and daily health habits in a large sample of young adults living in Dunedin, New Zealand (total  $n = 1470$ ). The majority of the participants were University of Otago students, and were all between 17-25 years old. The participants had first completed an initial survey on demographics, and then were tracked across 13 days via an online daily diary completed each night between 3-8 pm. Participants also attended a clinic visit where they were measured for height and weight to compute BMI. Figure 1 presents a flow diagram of participant selection. The Daily Life Study was run across four years from 2011 to 2014; however, we selected only participants from the 2013 and 2014 years because the assessment of daily health habits during those years was more extensive (with more relevant variables measured). We further excluded those participants who had provided fewer than 7 diary records (out of 13 possible) across the course of the study ( $n = 25$ ). Before data cleaning, our sample consisted of 796 participants (72.4% female), aged between 17-25 years ( $m = 19.72$ ,  $sd = 1.73$ ). The majority of participants were New Zealand European/pākehā (77.5%), followed by Asian (10.8%), and Māori/Pacifica participants (5.5%). Participants of all other ethnicities made up only 6.2% of the sample and were thus aggregated into one category.

### *Measures, data cleaning, and preprocessing*

A list of measures with reliabilities and descriptive statistics can be found in the Supplementary Materials. We selected the following demographic variables ( $n = 5$ ) from the initial survey: age, gender, ethnicity, self-reported childhood socioeconomic status (SES; Griskevicius, Delton, Robertson, & Tybur,

2011), and a measure of body mass index (BMI), calculated from height and weight measured during a clinic visit. From the daily survey, we selected the daily flourishing scale items (Diener et al., 2010), as well as all stress self-assessment, somatic self-assessment, diet-related, and health habit variables (total  $n = 22$ ) that were measured for the entirety of 2013-2014 and could be theoretically linked with flourishing. While most of the 22 variables assessed the participants' state on the day of reporting, some of the variables inquired about the participants' state the night before. Of the 22 variables, the stress self-assessment variables ( $n = 2$ ) were: felt stressed today, most stressful event today. The somatic self-assessment variables ( $n = 5$ ) were: felt tired today, felt rundown today, felt cold or flu today, had hangover today, and had trouble concentrating today. The diet-related ( $n = 11$ ) variables were: servings of fruit today, servings of vegetables today, servings of sweets today, servings of soft drink today, servings of chips today, servings of fruit last night, servings of vegetables last night, servings of sweets last night, servings of soft drink last night, servings of chips last night, and standard drinks of alcohol last night. The health habits ( $n = 4$ ) were: hours slept last night (sleep quantity), felt refreshed after waking up today (sleep quality), minutes physically active today, and minutes spent in nature today. Most variables were measured on a 5-point Likert scale, (0 = not at all, 1 = a little, 2 = somewhat, 3 = moderately, 4 = very). Drinking alcohol, sleep quantity, time spent in nature, physical activity, and the diet related variables were freely reported in the specified units (standard drinks, hours, and minutes, respectively, and servings).

Across all daily variables, we dropped the first two days of observations to account for initial elevation bias – the tendency of participants to over-report symptoms in the beginning of longitudinal studies (Shrout et al., 2018). To make sure that dropping the first two days did not bias our results, we ran a sensitivity analysis (see Supplementary Materials). Additionally, there were 0.5% of missing values across all 29 daily variables (8 flourishing items + 22 lifestyle habits). All variables except for minutes spent in nature today were missing less than 1.5% of all values. The minutes spent in nature variable was missing 4.35% of all values, and there were several participants for whom all or majority (> 50%) of values were missing. Because of the number of missing values, we decided to drop it after conducting sensitivity analyses to show that the variable inclusion/exclusion did not affect predictive projection feature selection (see Supplementary Materials). After excluding the time spent in nature variable, the daily ( $n = 21$ ) and demographic ( $n = 5$ ) made up a total of 26 predictor variables. Because the ethnicity variable contained four levels which added two additional dummy predictors, there were in the end 28 predictors. We did not impute 0.4% of missing values that were left after excluding minutes spent in nature today variable, as the values for each variable would be averaged across days for each participant. As for demographic variables, there were four participants missing their BMI information, and one participant missing two out of the three of the childhood SES items. These five participants were dropped from the analysis. Thus, the final sample consisted of 791 participants.

The eight flourishing scale items were averaged into one daily flourishing variable. Following that, all variables from the daily survey were averaged across days (including the newly created daily flourishing variable). Additionally, childhood SES was averaged across items as well. Finally, all continuous variables were centred and scaled to 1-unit standard deviation.

## *Statistical analyses and modeling*

Prior to fitting any models, we randomly split our data into training (75%;  $n = 593$ ) and test set (25%,  $n = 198$ ). The training data was used to fit all models. The test data was used to validate the models' predictions on an independent dataset. All relevant statistics are reported for both the training and test data.

We first fitted a Bayesian multiple regression as the reference model, predicting average daily flourishing from all 28 candidate well-being predictors (21 daily predictors, 5 demographic predictors, 2 additional dummy predictors for ethnicity). We used weakly informative normal priors for the predictor slopes and the intercept (normal, mean = 0, sd = 1). For model standard deviation, we also used a weakly informative prior (half-normal, mean = 0, sd = 1). As for sampling, we ran four chains of 4,000 iterations, with 2,000 iterations of warm-up and 2,000 iterations of sampling each. The chains were run in parallel to speed up convergence.

After fitting the reference model, we used predictive projection feature selection to fit a submodel with fewer predictors that would give similar predictions to the full model. To implement predictive projection, we used the `projpred` package (Pironen et al., 2018). The variables were sequentially entered into the submodels using L1 search (`projpred` default for  $> 20$  variables) and the submodels' predictive performance was compared using expected log predictive density (ELPD) information criterion obtained via PSIS-LOO. For the optimal submodel, we chose the smallest submodel for which the difference between its own and the reference model's ELPD was no more than 1 standard error from zero (`projpred` default), that is, we chose the smallest submodel which was expected to perform the same as/outperform the reference model with at least 16% probability (and perform worse with 84% probability).

Finally, we compared the predictive performance of predictive projection feature selection to five other variable selection approaches: the original Bayesian reference multiple regression model, a frequentist multiple regression model with all predictors entered simultaneously, frequentist stepwise selection using p-values, frequentist stepwise selection using Akaike Information criterion (Akaike, 1974), and many simple regression models corrected for multiple comparison. R does not have a default function for stepwise selection using p-values so we used a publicly available R code that implements an SPSS-like stepwise selection with p-values (Meys, 2009). The p-values for the simple regressions were adjusted via Bonferroni correction. Besides the 1 SE rule projected submodel featured above, we also included in the comparison a projected submodel with cross-validated predictive performance matching the reference model, that is, the smallest possible submodel for which there was at least 50% probability that it would perform as well as/better than the reference model. In total, seven models were compared.

In the Bayesian models, we interpreted predictors as "significant" when their 95% credible interval did not overlap zero, and, in the case of predictive projection, when they were selected into the optimal submodel. For the frequentist models, we used the conventional but arbitrary  $\alpha < 0.05$  threshold. To evaluate the

performance of all models, we used root mean squared error (RMSE) and Bayesian  $R^2 = \frac{Var_{\mu}}{Var_{\mu} + Var_{res}}$  (Gelman et al., 2019).

## Results

The reference model converged well, with no divergent transitions and good  $\hat{R}$  values (all  $\approx 1$ ), indicating that the chains mixed well (Gelman, Andrew & Rubin, 1992). All parameters had a good effective sample size (all > 5,000). As a posterior predictive check, we compared the distribution of observed daily flourishing to that of 500 simulated datasets drawn from the posterior predictive distribution, and found that the observed data matched the densities of the simulated data fairly well (see Supplementary Materials). There were few outliers with very low average daily flourishing in the observed data, however, when we evaluated the model via Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO cv; Vehtari, Gelman, & Gabry, 2016), we found no evidence of these observations having a disproportionate influence on the model fit, as indicated by satisfactory Pareto-k values (all “good”,  $k < 0.5$ ; see Supplementary Materials). Visual checks of predicted values vs residuals did not indicate any heteroscedasticity or gross non-linear trends (see Supplementary Materials). On the training data, the reference model had a Bayesian  $R^2$  of 0.359 (0.308 – 0.405, 95% credible interval; CI) and a model standard deviation/RMSE of 0.821 (0.774 – 0.870 95% CI), indicating a moderately good predictive performance.

Using the predictive projection, we found that a submodel with only 3 predictors made predictions that were “similar enough” to those of the reference model with all 28 predictors, and the sensibility of this rule-based approach was confirmed with a visual check of the feature selection trajectory (see Figure 2A). On the training data, the optimal submodel had a Bayesian  $R^2$  of 0.267 (0.213-0.323, 95% CI) and a model standard deviation/RMSE of 0.878 (0.826 – 0.931 95% CI). The optimal submodel’s predictions were strongly correlated with the reference model’s predictions (Pearson  $r = 0.897$ ), and this was again confirmed by visual checks (see Figure 2B).

The optimal submodel contained three predictors (see Figure 3A). In order as they were entered into the submodel, the three predictors were: felt refreshed after waking up today (0.377, 0.306 – 0.453 95% CI), had trouble concentrating today (-0.211, -0.281 – -0.138 95% CI), and servings of fruit today (0.132, 0.065 – 0.200 95% CI).

We tested the submodel’s performance on holdout test data. The submodel performed well on the test data (slightly worse than on the training data but within 95% CI bounds), with an observed Bayesian  $R^2$  of 0.253 and an observed RMSE of 0.884 (see Figure 3B).

The results of the comparison of different variable selection methods are shown in Table 1. The two multiple regression models had the best out-of-sample-predictive accuracy as assessed by test data  $R^2$  of 0.33. After that, the stepwise selection based on AIC had the best out-of-sample predictive accuracy,

followed by projection for which the mean expectation matched the reference model, stepwise selection based on p-values, and projection with the 1 SE rule. One simple regression (felt refreshed after waking up) had a high out-of-sample  $R^2$  of 0.218, however the other simple regressions had much lower predictive accuracy, with the mean  $R^2$  of 0.028 and sd of 0.046. Projection lead to the most parsimonious inference, however, with the 1 SE projection submodel selecting 3 predictors and the matched projection submodel selecting only 6 predictors, fewer than any of the other methods. Overall, larger models tended to have better out-of-sample predictive accuracy, however, the projection methods resulted in more parsimonious models - the multiple, stepwise, and simple regressions variable selection approaches all resulted in greater number of "significant" predictors. Across all models, felt refreshed after waking up and had (less) trouble concentrating were universally selected predictors. Additional predictors that tended to be selected relatively often across models included gender: female, servings of fruit today, (fewer) servings of sweets last night, (fewer) servings of sweets today, (fewer) servings of soft drink last night, felt (less) tired today, and lower BMI.

## Discussion

Using predictive projection feature selection, we found that only three of the 28 candidate well-being predictors were sufficient to approximate the predictions of a large reference model with all 28 predictors. Specifically, participants who reported feeling more refreshed after waking up, having less trouble concentrating, and eating more servings of fruit scored highest in their average daily flourishing. We tested the optimal 3-predictor submodel on test data and found that it predicted the unseen participants' average daily flourishing well, with performance measured by observed RMSE only slightly worse than on the training data. Lastly, we also found that when comparing predictive projection to other variable selection methods, we found that generally, larger models tended to have better predictive accuracy on test data, however, predictive projection tended to select more parsimonious models, and the projection submodel matched to the reference model performed better than stepwise selection based on p-values, with fewer predictors.

The fact that only three predictors were sufficient to predict almost as much variation in average daily well-being as the reference model with all 28 candidate predictors does not suggest that the left out 25 predictors have no relationship with flourishing. In fact, with simple regression, we saw that 10 predictors had a statistically significant association with flourishing, after Bonferroni correction was applied. However, whereas the predictors selected by simple regressions only described the relationship between themselves and flourishing in isolation, the three predictors selected by predictive projection were sufficient to explain most of the variance in average daily flourishing while simultaneously controlling for each other. As such, our results suggest that the three selected predictors are the strongest predictors of average daily flourishing, at least among groups of related predictors. For example, inasmuch as diet is related to flourishing, our results suggest that fruit consumption may be the strongest indicator of good diet and flourishing, as indicated by the fact that it was the earliest from the group of diet-related predictors in the feature selection trajectory, and was the only diet-related predictor that was present in the

optimal submodel. Likewise, while there were several correlated candidate predictors related to fatigue and somatic issues, the fact that having (less) trouble concentrating was the first and only predictor from this group selected into the optimal submodel suggests that it may be the strongest somatic predictor of well-being.

The projected submodel based on the 1 SE rule had worse predictive accuracy on test data than the other variable selection methods, however, we still chose to do inference with this model. The reason why we chose the 1 SE rule submodel over the other projection submodel that was matched on expected predictive accuracy is that the 1 SE rule leads to simpler models, more robust inference, and, thus, greater expectable reproducibility. The choice of rule for the predictive projection submodel size is a trade-off between explanation and robustness (Piironen et al., 2018), similar to the 4-to-1 trade-off between type-I and type-II error that underlies the conventional 0.05 alpha and 0.8 power recommendations (Chase & Tucker, 1976; Cohen, 1965; Lakens, 2013). Additionally, it is important to highlight that while the two stepwise models had higher predictive accuracy on the test data, these models cannot be used for inference as the p-values and standard errors in the selected models are not adjusted for the selection procedure and as such do not adequately control for type-I error (Flom & Cassell, 2007; G. Smith, 2018). The marginal credible intervals from the projection, on the other hand, come from a valid posterior distribution (Piironen et al., 2018) and thus can be used for inference.

The first and therefore strongest predictor selected into the submodel was sleep quality – how refreshed a participant felt after waking up, on their average day. Sleep quality has been consistently shown to be one of the strongest predictors of well-being, especially in young adults, with poor sleep quality being strongly linked to poor mental health outcomes including symptoms of depression (Pilcher et al., 1997; Ridner et al., 2016; Wallace et al., 2017; Wilson et al., 2014). Further, it is important to note that while sleep quality is often shown to be an important predictor of well-being, sleep quantity is not (Pilcher et al., 1997; Wallace et al., 2017). Likewise, in our study, while sleep quality was entered into the submodels early along the feature selection trajectory, reflecting its predictive strength, sleep quantity was only entered long after any improvement in predictive accuracy was shown. Our findings thus strengthen past evidence which shows that sleep quality (but not sleep quantity) is one of the strongest predictors of young adult well-being.

The second predictor selected into the submodel was how much trouble concentrating a participant had on their average day. More trouble concentrating was associated with lower well-being. Having trouble concentrating is one of the key symptoms of the major depressive disorder (MDD), and is often assessed by diagnostic scales, such as the popular CES-D scale (Kohout et al., 1993) and the DSSS scale (Hung et al., 2006). Less research has linked lack of concentration to lower well-being per se. However, there is evidence that “mind wandering” – thinking about things not present in the immediate environment – is associated with less happiness (Killingsworth & Gilbert, 2010; Smallwood & Schooler, 2015). As such, lacking ability to concentrate could be a proxy measure for MDD symptoms and mind wandering, and may potentially be a stronger predictor of well-being than other somatic symptoms measured in our study, such as feeling tired or run down.

The third and last predictor entered into the submodel was daytime fruit consumption. As was discussed previously, the fact that daytime fruit consumption was the first diet-related predictor to be selected in the feature selection trajectory and the only diet-related predictor that made it into the submodel suggests that it may be one of the strongest indicators of diet quality, as it relates to flourishing. Fruit and vegetable consumption has been previously shown to predict flourishing (Blanchflower et al., 2013; Conner et al., 2015, 2017; Hong & Peltzer, 2017; Mujcic & Oswald, 2016; Prendergast, Mackay, et al., 2016). There were other diet related predictors entered into the submodel early along the feature selection trajectory, namely (lower) night-time softdrink consumption and daytime vegetable consumption as the 4th and 5th predictors, respectively. However, based on the 1 SE rule, fruit consumption only was sufficient to approximate the reference model's predictions. The most straightforward interpretation of this fact is that fruit consumption may contain enough information about the quality of one's diet (as it relates to flourishing) to make other diet-related predictors redundant. Additionally, there is evidence that raw fruit and vegetables are a stronger predictors of well-being than cooked fruit and vegetables (Brookie et al., 2018), and since fruit is more often eaten raw, general fruit consumption may be a stronger indicator of a good diet than vegetable consumption. Lastly, given that poor diet quality is associated with ill-being (Jacka et al., 2010), and that our dependent variable (flourishing) may capture more positive aspects of mental health and well-being, it is possible that some of the predictors associated with unhealthy diet would have greater predictive utility if the dependent variable specifically focused on ill-being, such as depressive symptoms. Be it as it may, our findings suggest that fruit consumption may be one of the strongest diet-related predictors of flourishing.

Interestingly, despite the fact that physical exercise has often been linked to well-being in the past literature (Diener et al., 2018; Goodwin, 2003; Hassmén et al., 2000; Prendergast, Schofield, et al., 2016), exercise was not selected into the optimal submodel in our study, and was in fact added very late along the feature selection trajectory as the 20th predictor. Additionally, as a predictor in simple regression, the coefficient for exercise did not survive Bonferroni correction either. One possible explanation for why exercise was not associated with flourishing in our sample may be that the sample was relatively homogenous (all young university students), and so the range of exercise across participants may have been too narrow to predict flourishing in a meaningful way. On a related note, the cited studies on the associations of well-being and exercise investigated older populations, and so it is possible that the association between exercise and well-being becomes more pronounced over the course of one's life, and might not be as strong initially in young adults.

The fact that no demographic predictors were selected into the optimal submodel is somewhat surprising. Gender was the first demographic factor along the selection trajectory, entered the submodels as the sixth predictor, and had a moderately large point estimate for slope in the reference model (with female gender predicting greater flourishing). The other categorical demographic predictor, ethnicity, had one of its four dummy variables (ethnicity: Māori/Pacific Islander) entered as the 16th predictor. One possible explanation for why gender and ethnicity were not included in the optimal submodel is because there was large uncertainty around their estimates in the reference model, as evidenced by wide marginal posterior credible intervals (see Supplementary Materials). The width of the credible intervals likely

reflected the fact that the sample was unbalanced (72.6% female, 27.4% male; 77.6% European/Caucasian, 10.6% Asian, 5.4% Maori/Pacific Islander, 6.2% Other). Alternatively, some of the variance in flourishing explainable by gender and ethnicity may have been already explained by any of the five selected predictors. Lastly, while some studies do show evidence of higher flourishing in women (Piqueras et al., 2011), there is also evidence suggesting that gender differences in flourishing may not be very large (Huppert & So, 2009). However, in our study, gender did seem to play some role as indicated by the moderately large point estimate for gender in the reference model (see Supplementary Materials). Rather, the more likely explanations for why gender and ethnicity were not selected into the optimal submodel is that there was large uncertainty around their estimates and they did not improve the out-of-sample prediction of average daily flourishing above and beyond the preceding predictors.

The three continuous demographic predictors – age, childhood SES, and BMI – were not selected into the optimal submodel either. All three predictors had relatively narrow marginal credible intervals in the reference model (see Supplementary Materials), and while the marginal CIs of age and SES overlapped with zero, the CI of BMI did not. A potential explanation for why BMI was not included in the optimal submodel may be its moderately weak correlation with fruit consumption, that is, the bulk of flourishing variance explainable by BMI may have already been explained by fruit consumption. It is interesting that neither age nor SES were selected into the submodel, however, a potential explanation of that fact may be that our sample was relatively homogenous regarding these two variables, as it consisted of young university students.

Our findings have several practical implications for the study of well-being and policy. First, we employed rather conservative variable selection and model validation methods, such as the PSIS-LOO cross-validation and the use of the hold-out data set. Thus, we believe that the results of our study should be safe from overfitting, and in turn more robust and replicable. Second, using predictive projection, we were able to select a very parsimonious model with few theoretically relevant predictors and show good predictive accuracy on a hold-out sample. Thus, we have demonstrated that predictive projection can be a very powerful tool when dealing with the issue of selecting relevant predictors from a set of many potential candidates. We believe that predictive projection may be especially useful in of the study of health and well-being, where the issue of variable selection arises frequently. Lastly, our study identified potential targets for future interventions. While poor ability to concentrate is more likely to be a symptom rather than a cause of low flourishing and may be difficult to change systematically, sleep quality and day-time fruit consumption are modifiable health habits that could be targeted in future studies.

Our study has several limitations. First, as was mentioned earlier, the fact that only three variables were selected into the optimal submodel does not mean that the other predictors are not related to well-being. Therefore, we cannot conclude that the three predictors are the only predictors of flourishing, only that they are sufficient to predict it with similar degree of accuracy as all 28 predictors. Second, given that our data is observational, we cannot make causal claims about the predictors' influence. Inasmuch as we say that a predictor predicts higher or lower average daily flourishing, we mean that the predictor has a positive or negative association with flourishing and, that the model with the fitted predictors has been

shown to perform well on unseen data. However, correlational data can be used in an exploratory way to determine possible targets of intervention in follow up research. Third, our sample consisted almost entirely of young, college-age adults (age-range 17–25 years), and thus some of the findings may not generalize to an older (or younger) population. For example, young adults tend to be at an increased risk for poor sleep quality (Lund et al., 2010), and thus it is possible that the influence of sleep quality on well-being might be more pronounced in this population. It may be interesting to apply predictive projection on well-being and health habit data from other age cohorts to see whether there are age-related differences in the sufficient predictors of flourishing in future research. While our data comes from a micro-longitudinal daily diary study, we only applied our analyses to the data cross-sectionally. There are several reasons for why we did not analyse the trends within participants. First off, predictive projection feature selection is not yet implemented for mixed effects models (Piironen et al., 2018) and so we do not have the ability to build a parsimonious within-participants model in the same way as we did across participants. Second, there is no guarantee that the three strongest cross-sectional predictors of flourishing across participants will also be the strongest predictors of flourishing within-participants – in fact, it is rather unlikely (Simpson's paradox; see Wagner, 1982). As such, we believe that a proper repeated measures analysis using mixed effects models is outside of the scope of this article and may require more involved, theory-driven manual modelling and model comparison.

## Conclusion

Using predictive projection feature selection, we were able to accurately predict average daily flourishing in young adults using just the information about how refreshed they felt after waking up, how much trouble they had concentrating, and how many servings of fruit they ate on their average day. That is, we were able to approximate the predictive accuracy of a large model with 28 candidate predictors with a much smaller submodel containing only the above three predictors. Our findings reinforce and extend our understanding of modifiable behaviours that might contribute to well-being, in that sleep quality is more strongly associated with better well-being than sleep quantity, that having trouble concentrating is related to lower well-being, and having good dietary habits (as indicated by fruit consumption) is related to higher well-being. Moreover, our findings suggest that having trouble concentrating may be one of the strongest indicators for poor outcomes among somatic-related predictors, and that fruit consumption may be one of the strongest indicators for good outcomes among diet-related predictors. We have used conservative methods to arrive at our results, and as such they should be robust to overfitting and more replicable. Finally, while the search for predictors of well-being is challenging, we believe that it can be greatly helped by employing the new statistical methods used in the present paper.

## Abbreviations

**BMI**

Body mass index

**CI**

Credible interval

## **CES-D**

Center for Epidemiological Studies - Depression scale

## **DSSS**

Depression and somatic symptoms scale

## **ELPD**

Expected log-predictive density

## **PSIS-LOO cv**

Pareto-smoothed importance sampling leave-one-out cross-validation

## **RMSE**

Root-mean-squared error

## **SE**

Standard error

## **SES**

Socioeconomic status

# **Declarations**

## **Ethics approval and consent to participate**

Ethical approval was required for data collection in this study and was reviewed and granted by the University of Otago Human Ethics Committee (approval #10/177). All participants in the study signed written informed consent.

## **Consent for publication**

Not applicable.

## **Availability of data and materials**

The Daily Life Study dataset contains sensitive demographic information that could potentially be used to identify individual participants. De-identified dataset will be made available from the corresponding author upon reasonable request. All R code used for the analysis is available as part of the Supplementary Materials.

## **Competing interests**

The authors do not have any competing interests.

## **Funding**

The Daily Life Study was funded by the Health Research Council of New Zealand (grant 12/709 to TSC), a University of Otago Research Grant to TSC, and the Department of Psychology, University of Otago.

## Author information

Department of Psychology, University of Otago

William James Building, 275 Leith Walk, Dunedin 9016

Adam Bartonicek, Shay-Ruby Wickham, Narun Pat (Pornpattananangkul), Tamlin Conner

## Authors' contributions

AB and TC conceived the study. AB performed data cleaning and analysis, SW contributed to interpretation of the data, and TC and NP provided consultations on theory and statistical aspects of the analysis. All authors (AB, SW, NP, and TC) contributed to drafting, reading, writing, and revising the manuscript.

## Acknowledgements

We would like to thank Dr. Jayde Flett, Kathryn Bees, and Samantha Jones for their outstanding work coordinating the Daily Life Study, members of The Daily Experiences Lab in 2011-2014 for assistance with data collection, and Hadyn Youens for assistance with study programming.

## References

- Adams, T. B., Bezner, J. R., Drabbs, M. E., Zambarano, R. J., & Steinhardt, M. A. (2000). Conceptualization and measurement of the spiritual and psychological dimensions of wellness in a college population. *Journal of the American College Health Association*, *48*(4), 165–173. <https://doi.org/10.1080/07448480009595692>
- Adams, T. B., & Colner, W. (2008). The association of multiple risk factors with fruit and vegetable intake among a nationwide sample of college students. *Journal of American College Health*, *56*(4), 455–461. <https://doi.org/10.3200/JACH.56.44.455-464>
- Afrabandpey, H., Peltola, T., Piironen, J., Vehtari, A., & Kaski, S. (2019). *Making Bayesian Predictive Models Interpretable: A Decision Theoretic Approach*. 1–16. <http://arxiv.org/abs/1910.09358>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.
- Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(3), 370–374. <https://doi.org/10.1002/wics.84>
- Benjamini, Y. (2010). Simultaneous and selective inference: Current successes and future challenges. *Biometrical Journal*, *52*(6), 708–721. <https://doi.org/10.1002/bimj.200900299>

- Blanchflower, D. G., Oswald, A. J., & Stewart-Brown, S. (2013). Is Psychological Well-Being Linked to the Consumption of Fruit and Vegetables? *Social Indicators Research*, *114*(3), 785–801. <https://doi.org/10.1007/s11205-012-0173-y>
- Boyes, A. W., Girgis, A., D'Este, C., & Zucca, A. C. (2011). Flourishing or floundering? Prevalence and correlates of anxiety and depression among a population-based sample of adult cancer survivors 6 months after diagnosis. *Journal of Affective Disorders*, *135*(1–3), 184–192. <https://doi.org/10.1016/j.jad.2011.07.016>
- Brookie, K. L., Best, G. I., & Conner, T. S. (2018). Intake of raw fruits and vegetables is associated with better mental health than intake of processed fruits and vegetables. *Frontiers in Psychology*, *9*(APR), 1–14. <https://doi.org/10.3389/fpsyg.2018.00487>
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1). <https://doi.org/10.18637/jss.v080.i01>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). <https://doi.org/10.18637/jss.v076.i01>
- Chase, L. J., & Tucker, R. A. Y. K. (1976). *Derivation of Statistical Power*. 473–486.
- Cohen, J. (1965). Some statistical issues in psychological research. In *Handbook of clinical psychology* (pp. 95–121).
- Conner, T. S., Brookie, K. L., Carr, A. C., Mainvil, L. A., & Vissers, M. C. M. (2017). Let them eat fruit! the effect of fruit and vegetable consumption on psychological well-being in young adults: A randomized controlled trial. *PLoS ONE*, *12*(2), 1–19. <https://doi.org/10.1371/journal.pone.0171206>
- Conner, T. S., Brookie, K. L., Richardson, A. C., & Polak, M. A. (2015). On carrots and curiosity: Eating fruit and vegetables is associated with greater flourishing in daily life. *British Journal of Health Psychology*, *20*(2), 413–427. <https://doi.org/10.1111/bjhp.12113>
- Daróczi, G., & Roman, T. (2018). *pander: An R "Pandoc" Writer* (0.6.3). <https://cran.r-project.org/package=pander>
- 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.1758-0854.2010.01045.x>
- Diener, E., Seligman, M. E. P., Choi, H., & Oishi, S. (2018). Happiest People Revisited. *Perspectives on Psychological Science*, *13*(2), 176–184. <https://doi.org/10.1177/1745691617697077>

- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*(2), 143–156.
- Etxebarria, I., Etxebarria, I., & Urdaneta, E. (2019). Subjective well-being among the oldest old: The role of personality traits. *Personality and Individual Differences, 146*(April 2018), 209–216.  
<https://doi.org/10.1016/j.paid.2018.04.042>
- Fink, J. E. (2014). Flourishing: Exploring Predictors of Mental Health Within the College Environment. *Journal of American College Health, 62*(6), 380–388. <https://doi.org/10.1080/07448481.2014.917647>
- Flom, P. L., & Cassell, D. L. (2007). Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. *Northeast SAS User Group (NESUG) Inc 20th Annual Conference: 11-14th November 2007; Baltimore, Maryland, 1–7.*
- Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society. Series A: Statistics in Society, 182*(2), 389–402.  
<https://doi.org/10.1111/rssa.12378>
- Gelman, Andrew, & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457–511. <https://doi.org/10.1214/ss/1177013437>
- Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian Regression Models. *American Statistician, 73*(3), 307–309. <https://doi.org/10.1080/00031305.2018.1549100>
- Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking.” *Unpublished Draft.*  
[http://www.stat.columbia.edu/~gelman/research/unpublished/p\\_hacking.pdf](http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf)
- Goodwin, R. D. (2003). Association between physical activity and mental disorders among adults in the United States. *Preventive Medicine, 36*(6), 698–703. [https://doi.org/10.1016/S0091-7435\(03\)00042-2](https://doi.org/10.1016/S0091-7435(03)00042-2)
- Graham, M. H. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology, 84*(11), 2809–2815. <https://doi.org/10.1890/02-3114>
- Griskevicius, V., Delton, A. W., Robertson, T. E., & Tybur, J. M. (2011). Environmental contingency in life history strategies: The influence of mortality and socioeconomic status on reproductive timing. *Journal of Personality and Social Psychology, 100*(2), 241–254.  
<https://doi.org/10.1037/a0021082>.Environmental
- Hassmén, P., Koivula, N., & Uutela, A. (2000). Physical exercise and psychological well-being: A population study in Finland. *Preventive Medicine, 30*(1), 17–25. <https://doi.org/10.1006/pmed.1999.0597>

- Hawkins, D. M. (2004). The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences*, 44(1), 1–12. <https://doi.org/10.1021/ci0342472>
- Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection – A review and recommendations for the practicing statistician. *Biometrical Journal*, 60(3), 431–449. <https://doi.org/10.1002/bimj.201700067>
- Höge, M., Wöhling, T., & Nowak, W. (2018). A Primer for Model Selection: The Decisive Role of Model Complexity. *Water Resources Research*, 54(3), 1688–1715. <https://doi.org/10.1002/2017WR021902>
- Hong, S. A., & Peltzer, K. (2017). Dietary behaviour, psychological well-being and mental distress among adolescents in Korea. *Child and Adolescent Psychiatry and Mental Health*, 11(1), 1–12. <https://doi.org/10.1186/s13034-017-0194-z>
- Howell, A. J. (2009). Flourishing: Achievement-related correlates of students' well-being. *Journal of Positive Psychology*, 4(1), 1–13. <https://doi.org/10.1080/17439760802043459>
- Hung, C. I., Weng, L. J., Su, Y. J., & Liu, C. Y. (2006). Depression and somatic symptoms scale: A new scale with both depression and somatic symptoms emphasized. *Psychiatry and Clinical Neurosciences*, 60(6), 700–708. <https://doi.org/10.1111/j.1440-1819.2006.01585.x>
- Huppert, F. A., & So, T. T. C. (2009). What percentage of people in Europe are flourishing and what characterises them? The importance of subjective indicators of well-being. *OECD/ISQOLS Meeting "Measuring Subjective Well-Being: An Opportunity for NSOs?"* 1–7.
- Huppert, F. A., & So, T. T. C. (2013). Flourishing Across Europe: Application of a New Conceptual Framework for Defining Well-Being. *Social Indicators Research*, 110(3), 837–861. <https://doi.org/10.1007/s11205-011-9966-7>
- Jacka, F. N., Kremer, P. J., Berk, M., de Silva-Sanigorski, A. M., Moodie, M., Leslie, E. R., Pasco, J. A., & Swinburn, B. A. (2011). A prospective study of diet quality and mental health in adolescents. *PLoS ONE*, 6(9), 1–7. <https://doi.org/10.1371/journal.pone.0024805>
- Jacka, F. N., Kremer, P. J., Leslie, E. R., Berk, M., Patton, G. C., Toumbourou, J. W., & Williams, J. W. (2010). Associations between diet quality and depressed mood in adolescents: Results from the Australian Healthy Neighbourhoods Study. *Australian and New Zealand Journal of Psychiatry*, 44(5), 435–442. <https://doi.org/10.3109/00048670903571598>
- Jacka, F. N., O'Neil, A., Opie, R., Itsiopoulos, C., Cotton, S., Mohebbi, M., Castle, D., Dash, S., Mihalopoulos, C., Chatterton, M. Lou, Brazionis, L., Dean, O. M., Hodge, A. M., & Berk, M. (2017). A randomised controlled trial of dietary improvement for adults with major depression (the "SMILES" trial). *BMC Medicine*, 15(1), 1–13. <https://doi.org/10.1186/s12916-017-0791-y>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. In *Springer Texts in Statistics*. <https://doi.org/10.1016/j.peva.2007.06.006>

- Kadane, J. B., & Lazar, N. A. (2004). Methods and criteria for model selection. *Journal of the American Statistical Association*, *99*(465), 279–290. <https://doi.org/10.1198/016214504000000269>
- Kassambara, A. (2019). *ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2' (0.1.3)*. <https://cran.r-project.org/package=ggcorrplot>
- Kawada, T., Kuratomi, Y., & Kanai, T. (2009). Lifestyle determinants of depressive feeling and a feeling of unhappiness among workers: a study in Japan. *Work*, *33*(3), 255–260.
- Kern, M. L., Della Porta, S. S., & Friedman, H. S. (2014). Lifelong pathways to longevity: Personality, relationships, flourishing, and health. *Journal of Personality*, *82*(6), 472–484. <https://doi.org/10.1111/jopy.12062>
- Keyes, C. L. M. (2007). Promoting and Protecting Mental Health as Flourishing: A Complementary Strategy for Improving National Mental Health. *American Psychologist*, *62*(2), 95–108. <https://doi.org/10.1037/0003-066X.62.2.95>
- Keyes, C. L. M., Dhingra, S. S., & Simoes, E. J. (2010). Change in level of positive mental health as a predictor of future risk of mental illness. *American Journal of Public Health*, *100*(12), 2366–2371. <https://doi.org/10.2105/AJPH.2010.192245>
- Keyes, C. L. M., & Simoes, E. J. (2012). To flourish or not: Positive mental health and all-cause mortality. *American Journal of Public Health*, *102*(11), 2164–2172. <https://doi.org/10.2105/AJPH.2012.300918>
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932. <https://doi.org/10.1126/science.1192439>
- Koenig, H. G. (2009). Research on religion, spirituality, and mental health: A review. *Canadian Journal of Psychiatry*, *54*(5), 283–291. <https://doi.org/10.1177/070674370905400502>
- Kohout, F. J., Berkman, L. F., Evans, D. A., & Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D depression symptoms index. *Journal of Aging and Health*, *5*(2), 179–193.
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*(NOV), 1–12. <https://doi.org/10.3389/fpsyg.2013.00863>
- Larmarange, J. (2020). *labelled: Manipulating labelled data (2.2.2)*. <https://cran.r-project.org/package=labelled>

- Li, N. P., & Kanazawa, S. (2016). Country roads, take me home... to my friends: How intelligence, population density, and friendship affect modern happiness. *British Journal of Psychology*, *107*(4), 675–697. [http://ink.library.smu.edu.sg/sooss\\_research](http://ink.library.smu.edu.sg/sooss_research)
- Liao, Y., Schembre, S. M., O'Connor, S. G. O., Belcher, B. R., Maher, J. P., Dzibur, E., & Dunton, G. F. (2018). An electronic ecological momentary assessment study to examine the consumption of high-fat/high-sugar foods, fruits/vegetables and affective states among women. *Journal of Nutrition Education and Behavior*, *50*(6), 626–631. <https://doi.org/doi:10.1016/j.jneb.2018.02.003>. An
- Lund, H. G., Reider, B. D., Whiting, A. B., & Prichard, J. R. (2010). Sleep Patterns and Predictors of Disturbed Sleep in a Large Population of College Students. *Journal of Adolescent Health*, *46*(2), 124–132. <https://doi.org/10.1016/j.jadohealth.2009.06.016>
- Meys, J. (2009). *Automated model selection*. <https://stackoverflow.com/questions/3701170/stepwise-regression-using-p-values-to-drop-variables-with-nonsignificant-p-value>
- Moreno-Agostino, D., Caballero, F. F., Martín-María, N., Tyrovolas, S., López-García, P., Rodríguez-Artalejo, F., Haro, J. M., Ayuso-Mateos, J. L., & Miret, M. (2019). Mediterranean diet and wellbeing: evidence from a nationwide survey. *Psychology and Health*, *34*(3), 321–335. <https://doi.org/10.1080/08870446.2018.1525492>
- Mujcic, R., & Oswald, A. J. (2016). Evolution of well-being and happiness after increases in consumption of fruit and vegetables. *American Journal of Public Health*, *106*(8), 1504–1510. <https://doi.org/10.2105/AJPH.2016.303260>
- Müller, S., Scealy, J. L., & Welsh, A. H. (2013). Model selection in linear mixed models. *Statistical Science*, *28*(2), 135–167. <https://doi.org/10.1214/12-STS410>
- Oishi, S., Kesebir, S., & Diener, E. (2011). Income inequality and happiness. *Psychological Science*, *22*(9), 1095–1100. <https://doi.org/10.1177/0956797611417262>
- Ouweneel, E., Le Blanc, P. M., & Schaufeli, W. B. (2011). Flourishing students: A longitudinal study on positive emotions, personal resources, and study engagement. *Journal of Positive Psychology*, *6*(2), 142–153. <https://doi.org/10.1080/17439760.2011.558847>
- Piironen, J., Paasiniemi, M., & Vehtari, A. (2018). *Projective Inference in High-dimensional Problems: Prediction and Feature Selection*. 2015, 1–42. <http://arxiv.org/abs/1810.02406>
- Pilcher, J. J., Ginter, D. R., & Sadowsky, B. (1997). Sleep quality versus sleep quantity: Relationships between sleep and measures of health, well-being and sleepiness in college students. *Journal of Psychosomatic Research*, *42*(6), 583–596. [https://doi.org/10.1016/S0022-3999\(97\)00004-4](https://doi.org/10.1016/S0022-3999(97)00004-4)
- Pinquart, M., & Sörensen, S. (2000). Influences of socioeconomic status, social network, and competence on subjective well-being in later life: A meta-analysis. *Psychology and Aging*, *15*(2), 187–224.

<https://doi.org/10.1037/0882-7974.15.2.187>

Piqueras, J. A., Kuhne, W., Vera-Villarreal, P., Van Straten, A., & Cuijpers, P. (2011). Happiness and health behaviours in Chilean college students: A cross-sectional survey. *BMC Public Health, 11*.

<https://doi.org/10.1186/1471-2458-11-443>

Prendergast, K. B., Mackay, L. M., & Schofield, G. M. (2016). The Clustering of Lifestyle Behaviours in New Zealand and their Relationship with Optimal Wellbeing. *International Journal of Behavioral Medicine, 23*(5), 571–579. <https://doi.org/10.1007/s12529-016-9552-0>

Prendergast, K. B., Schofield, G. M., & Mackay, L. M. (2016). Associations between lifestyle behaviours and optimal wellbeing in a diverse sample of New Zealand adults. *BMC Public Health, 16*(1), 1–11.

<https://doi.org/10.1186/s12889-016-2755-0>

Pressman, S. D., & Cohen, S. (2005). Does positive affect influence health? *Psychological Bulletin, 131*(6), 925–971. <https://doi.org/10.1037/0033-2909.131.6.925>

Prochaska, J. J., Delucchi, K., & Hall, S. M. (2004). A meta-analysis of smoking cessation interventions with individuals in substance abuse treatment or recovery. *Journal of Consulting and Clinical Psychology, 72*(6), 1144–1156. <https://doi.org/10.1037/0022-006X.72.6.1144>

R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing, Vienna, Austria* (3.6.0). <https://www.r-project.org/>

Ridner, S. L., Newton, K. S., Staten, R. R., Crawford, T. N., & Hall, L. A. (2016). Predictors of well-being among college students. *Journal of American College Health, 64*(2), 116–124.

<https://doi.org/10.1080/07448481.2015.1085057>

Shrout, P. E., Stadler, G., Lane, S. P., Joy McClure, M., Jackson, G. L., Clavél, F. D., Iida, M., Gleason, M. E. J., Xu, J. H., & Bolger, N. (2018). Initial elevation bias in subjective reports. *Proceedings of the National Academy of Sciences of the United States of America, 115*(1), E15–E23.

<https://doi.org/10.1073/pnas.1712277115>

Sletta, C., Tyssen, R., & Løvseth, L. T. (2019). Change in subjective well-being over 20 years at two Norwegian medical schools and factors linked to well-being today: A survey. *BMC Medical Education, 19*(1), 1–12. <https://doi.org/10.1186/s12909-019-1476-3>

Smallwood, J., & Schooler, J. W. (2015). The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annual Review of Psychology, 66*(1), 487–518.

<https://doi.org/10.1146/annurev-psych-010814-015331>

Smith, G. (2018). Step away from stepwise. *Journal of Big Data, 5*(1). <https://doi.org/10.1186/s40537-018-0143-6>

- Smith, T. B., McCullough, M. E., & Poll, J. (2003). Religiousness and Depression: Evidence for a Main Effect and the Moderating Influence of Stressful Life Events. *Psychological Bulletin*, *129*(4), 614–636. <https://doi.org/10.1037/0033-2909.129.4.614>
- Steyerberg, E. W., Eijkemans, M. J. C., Harrell, F. E., & Habbema, J. D. F. (2001). Prognostic modeling with logistic regression analysis: In search of a sensible strategy in small data sets. *Medical Decision Making*, *21*(1), 45–56. <https://doi.org/10.1177/0272989X0102100106>
- Tay, L., & Kuykendall, L. (2013). Promoting happiness: The malleability of individual and societal subjective wellbeing. *International Journal of Psychology*, *48*(3), 159–176. <https://doi.org/10.1080/00207594.2013.779379>
- Taylor, J., & Tibshirani, R. J. (2015). Statistical learning and selective inference. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(25), 7629–7634. <https://doi.org/10.1073/pnas.1507583112>
- Tierney, N., Cook, D., McBain, M., & Fay, C. (2020). *naniar: Data Structures, Summaries, and Visualisations for Missing Data* (0.5.0). <https://cran.r-project.org/package=naniar>
- van Erp, S., Oberski, D. L., & Mulder, J. (2019). Shrinkage priors for Bayesian penalized regression. *Journal of Mathematical Psychology*, *89*, 31–50. <https://doi.org/10.1016/j.jmp.2018.12.004>
- Vehtari, A., Gelman, A., & Gabry, J. (2015). Efficient implementation of leave-one-out cross-validation and WAIC for evaluating fitted Bayesian models. *ArXiv Preprint*.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Visser, M., & Routledge, L. A. (2007). Substance abuse and psychological well-being of South African adolescents. *South African Journal of Psychology*, *37*(3), 595–615. <https://doi.org/10.1177/008124630703700313>
- Wagner, C. H. (1982). Simpson ' s Paradox in Real Life Author. *The American Statistician*, *36*(1), 46–48.
- Wahl, D. R., Villinger, K., König, L. M., Ziesemer, K., Schupp, H. T., & Renner, B. (2017). Healthy food choices are happy food choices: Evidence from a real life sample using smartphone based assessments. *Scientific Reports*, *7*(1), 1–8. <https://doi.org/10.1038/s41598-017-17262-9>
- Wallace, D. D., Boynton, M. H., & Lytle, L. A. (2017). Multilevel analysis exploring the links between stress, depression, and sleep problems among two-year college students. *Journal of American College Health*, *65*(3), 187–196. <https://doi.org/10.1080/07448481.2016.1269111>

White, B. A., Horwath, C. C., & Conner, T. S. (2013). Many apples a day keep the blues away - Daily experiences of negative and positive affect and food consumption in young adults. *British Journal of Health Psychology, 18*(4), 782–798. <https://doi.org/10.1111/bjhp.12021>

Whittingham, M. J., Stephens, P. A., Bradbury, R. B., & Freckleton, R. P. (2006). Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology, 75*(5), 1182–1189. <https://doi.org/10.1111/j.1365-2656.2006.01141.x>

Wickham, H. (2017). *tidyverse: Easily Install and Load the “Tidyverse”* (R package version 1.2.1). <https://cran.r-project.org/web/packages/tidyverse/index.html>

Wilke, C. O. (2019). *cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2”*. (R package version 0.9.4). <https://cran.r-project.org/package=cowplot%0A>

Wilson, K. T., Bohnert, A. E., Ambrose, A., Davis, D. Y., Jones, D. M., & Magee, M. J. (2014). Social, behavioral, and sleep characteristics associated with depression symptoms among undergraduate students at a women’s college: A cross-sectional depression survey, 2012. *BMC Women’s Health, 14*(1). <https://doi.org/10.1186/1472-6874-14-8>

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science, 12*(6), 1100–1122. <https://doi.org/10.1177/1745691617693393>

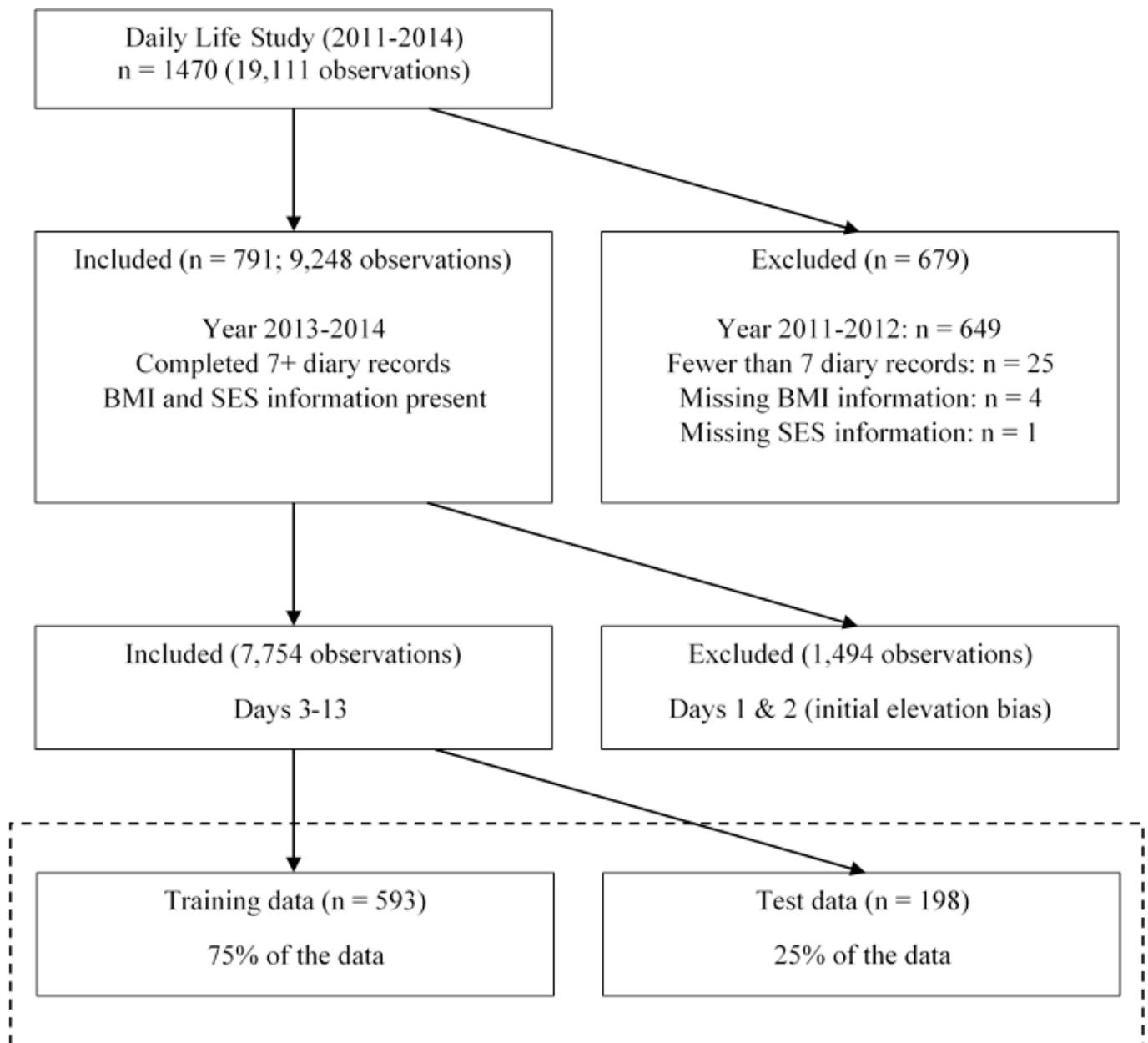
## Tables

**Table 1. Comparison of variable selection methods**

Model	Number of selected predictors	Number of significant predictors	Significant predictors (in order of strength)	
Reference Bayesian multiple regression	0.331	28	8	Felt refreshed after waking up today, had (less) trouble concentrating today, gender: female, (fewer) servings of sweets today, (fewer) servings of soft drink last night, (fewer) servings of sweets last night, felt (less) tired today, lower BMI
Multiple regression	0.332	28	8	Felt refreshed after waking up today, had (less) trouble concentrating today, gender: female, (fewer) servings of sweets today, (fewer) servings of soft drink last night, (fewer) servings of sweets last night, felt (less) tired today, lower BMI
Projected submodel (1 SE)	0.253	3	3	Felt refreshed after waking up today, had (less) trouble concentrating today, servings of fruit today
Projected submodel (matched)	0.284	6	6	Felt refreshed after waking up today, had (less) trouble concentrating today, servings of fruit today, (fewer) servings of soft drink last night, servings of vegetables today, gender: female
Stepwise selection: AIC	0.315	10	10	Felt refreshed after waking up today, ethnicity: Asian (lower), had (less) trouble concentrating today, gender: female, (fewer) servings of soft drink last night, (fewer) servings of sweets today, (fewer) servings of sweets last night, felt (less) tired today, servings of fruit today, lower BMI
Stepwise selection: p-values	0.275	8	8	Felt refreshed after waking up today, had (less) trouble concentrating today, gender: female, (fewer) servings of sweets today, felt (less) tired today, (fewer) servings of sweets last night, servings of fruit today, (fewer) servings of soft drink last night
Simple regressions	Max = 0.218 Mean = 0.028 Sd = 0.046	NA	10	Felt refreshed after waking up today, had (less) trouble concentrating today, felt (less) tired today, servings of fruit today, felt (less) rundown today, felt (less) stressed today, (fewer) servings of soft drink last night, servings of vegetables today, (lesser) stressful event today, (fewer) servings of soft drink today, servings of vegetables last night

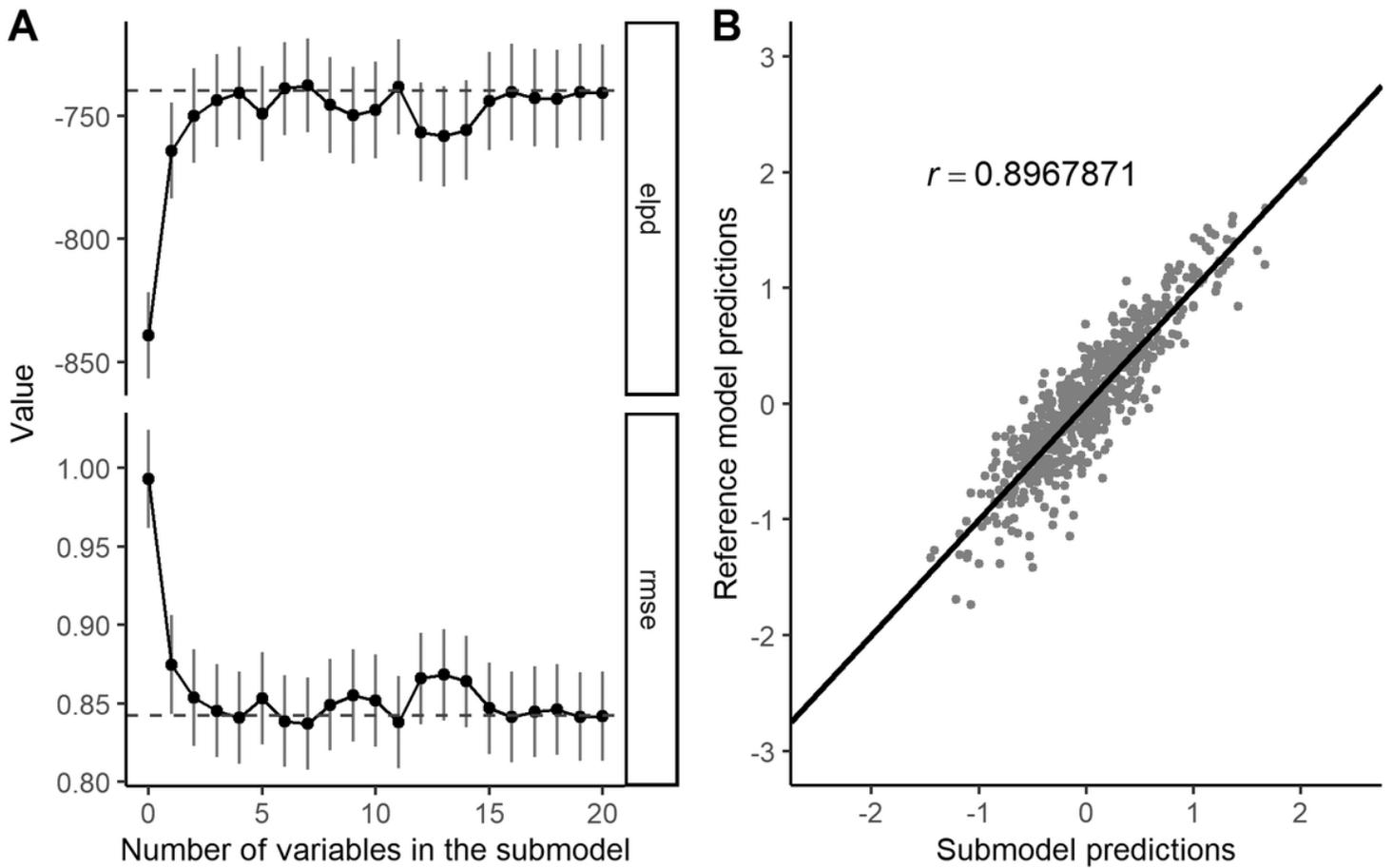
**legend:** Summary statistics of model selection strategies, showing on the test data, number of selected predictors, the number of significant predictors, and the names of the significant predictors (significant predictors ranked by absolute slope size).

## Figures



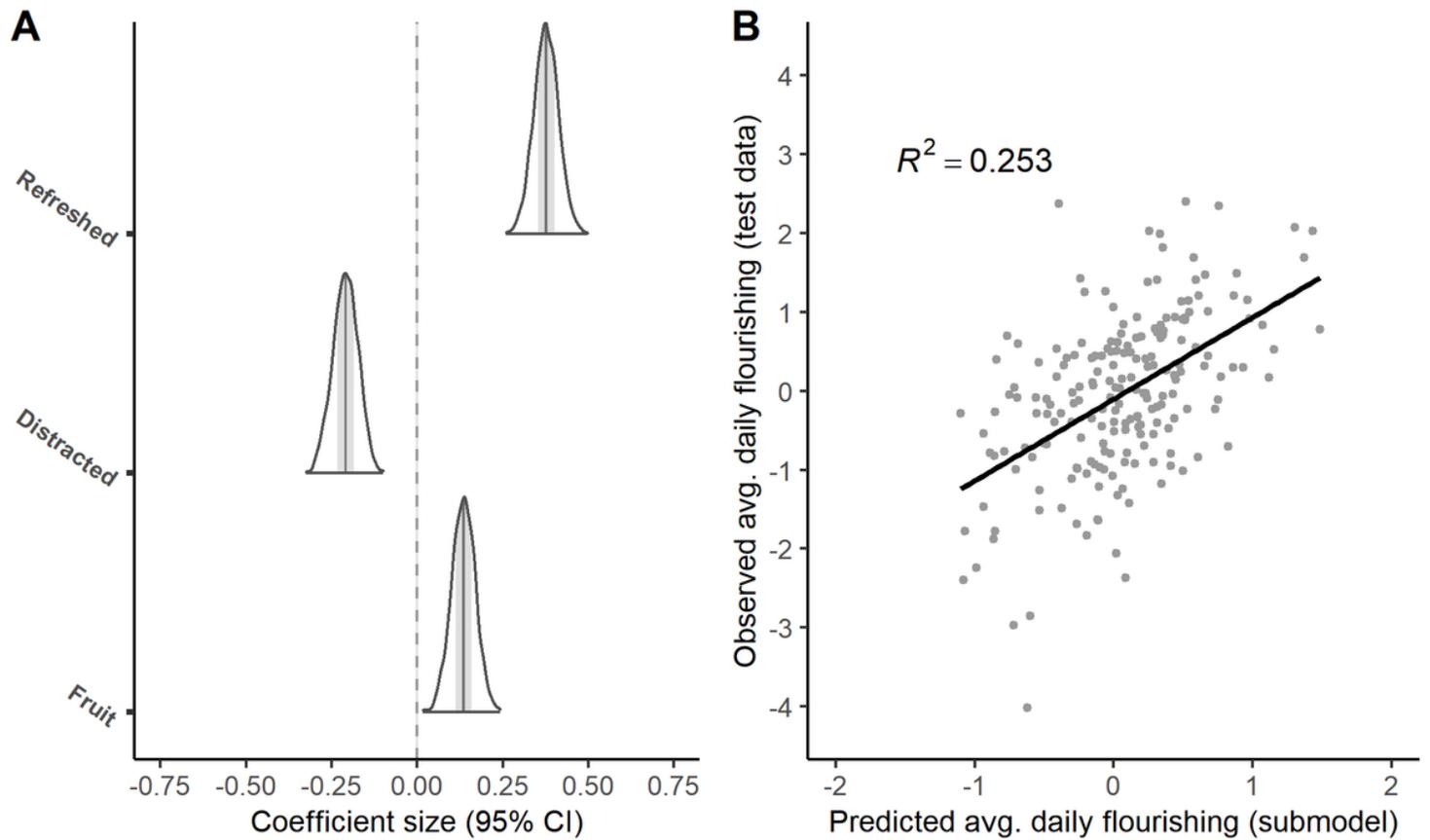
**Figure 1**

Flow diagram of the data selection procedure



**Figure 2**

Predictive projection feature selection trajectory and scatterplot of reference model's vs. submodel's prediction. A) Change in ELPD/decrease in RMSE as more predictors entered the submodel. B) Average daily flourishing predicted by the submodel (3 predictors) vs. the average daily flourishing predicted by the reference model (28 predictors; both predicting training data).



**Figure 3**

Credible intervals for predictors in the submodel and scatterplot of submodel's predictions vs. observed values. Marginal posterior distributions of predictors selected for the submodel (in order: felt refreshed after waking up today, had trouble concentrating today, servings of fruit today). B) Average daily flourishing predicted by the submodel vs. observed daily flourishing (unseen test data), with overlaid least squares fit.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [dailylife5.pdf](#)