Supplementary Online Content

Taquet M, Quoidbach J, Gross JJ, Saunders KEA, Goodwin GM. Mood homeostasis, low mood, and history of depression in 2 large population samples. *JAMA Psychiatry.* Published online April 22, 2020. doi: 0.1001/jamapsychiatry.2020.0588

eTable. Demographic Information for the Complete and Analyzed Data Sets
eMethods 1. List of Activities and Emotions
eMethods 2. Description of the 58sec Project
eMethods 3. Formal Definition of Mood Homeostasis
eMethods 4. Illustration of the Practical Calculation of Mood Homeostasis
eMethods 5. Mathematical Demonstration of the Invariance of Mood Homeostasis
eMethods 6. Covariates
eMethods 7. Details of Statistical Analysis
eMethods 8. Robustness Analysis
eMethods 9. Dynamic Simulation Process
eResults 1. Interpretation of the Effect Sizes
eResults 2. Differences in Mood Homeostasis Between the 2 Data Sets
eFigure 1. Procedure Followed to Calculate Mood Homeostasis From the Input Data and From the 2 Regression Models
eFigure 2. Mood Homeostasis is Invariant to Difference in Mood Level, Difference in Fluctuation Amplitude, and Combinations Thereof
eFigure 3. Different Factors May Drive Group Differences in Mood Homeostasis
eFigure 4. Results of the Robustness Analyses Carried Out in Both Data Sets
eFigure 5. Group Difference in Mood Homeostasis Between High and Low Average Mood in the 58sec Data Set After Excluding Thinking Which Was Found to be Partially Driving the Results
eFigure 6. Same Figure as Figure 2B-C and Figure 3B-C but With All Activity Labels Included
eFigure 7. Effect of Sampling Mood and Activities for Only 1 Day on the Value of Mood Homeostasis and Compounded Effect of Sampling Duration and Different Average Mood on Mood Homeostasis
eFigure 8. Mood Homeostasis Does Not Significantly Differ Between Age Group (A) and Between Income Levels (B)

eReferences.

This supplementary material has been provided by the authors to give readers additional information about their work.

eTable. Demographic Information for the Complete and Analyzed Data Sets

	58 seconds complete data (*N* = 64,133)	58 seconds analyzed data (*N* = 28,212)	Δ	W.H.O. complete data (*N* = 47,443)	W.H.O. analyzed data (*N* = 30,116)	Δ
Gender [% female]	62.5	65.8	Chi^2 = 126.5***	57.1	57.0	$Chi^2 = 0.1^{n.s.}$
Age (SD)	28.0 (9.3)	28.1 (9.0)	*t* = 3.65***	58.3 (13.9)	57.8 (14.7)	*t* = 5.0***
Country [%]						
France	91.3	94.7				
Belgium	4.0	0.6	Chi^2 =			
Switzerland	2.8	3.3	906.2***			
Other	1.9	1.4				
China				31.7	34.4	
Ghana				11.7	11.8	
India				25.7	27.9	Chi^2 =
Mexico				11.5	6.7	726.2***.
Russia				10.4	9.9	
South Africa				8.9	9.2	
History of depression [%]	N/C	N/C	N/C	2.8	2.9	$Chi^2 = 1.0^{n.s.}$

Note. $^{n.s.}$*P* > .05; **P* < .05; ***P* < .01; ****P* < .005; N/C = Not collected.

In the 58sec data, relatively more females than males were included (compared to the whole sample). There was also an effect of country of residence on the likelihood to be included. From the complete dataset, French people were the most likely and Belgian participants were the least likely to be included. This might reflect the fact that French people got more opportunities to be reminded of the app (it was featured on a French TV channel for multiple episodes of a reality TV show and was also featured on the TV show website). Belgian people on the other hand were probably informed of the app when it was featured (only once) in the news on a Belgian TV channel so that they were more likely to use the app only once. In the WHO SAGE dataset, there was no effect of gender on inclusion in the analyzed data and a small effect of age (the participants in the analyzed dataset were on average 6 months younger than the participants in the complete dataset). There was also an effect of country of origin with Mexican people being most likely to be included, and Chinese people being least likely). Importantly, there was no impact of a history of depression on the chance of being included.

eMethods 1. List of Activities and Emotions

In the 58sec application, participants were asked to select their activities from a non-mutually exclusive list of 25 choices which included (in brackets are the labels that we used in figures):
Working, Studying (working); Commuting (commuting); Cooking (cooking); Housework (housework); Shopping (shopping); Waiting, Being in line (waiting); Taking care of children (childcare); Playing (playing); Helping someone (helping); Sleeping, Resting (resting); Thinking, Praying (thinking)*; Showering, Getting Ready (grooming); Chatting (chatting); Eating (eating); Drinking (drinking); Texting, Emailing, Social Networking (texting); Calling someone (calling); Being on the Internet (internet); Watching TV (television); Listening to music (music); Cultural activity (culture); Doing sport (sport); Going out into nature (nature); Leisure (leisure); Other (other). We excluded activities that were represented in fewer than 0.1% of the samples. This resulted in no activity being excluded.

* In the application which was in French, this activity was called "méditer, prier". However, the terms "méditer" and "méditation" in French have a different meaning from the terms "meditate" and "meditation" in English. The definition from the Larousse French dictionary defines "méditation" as (translated to the best of our ability): "The action of reflecting, thinking deeply about something or about the realization of something". To avoid confusion with the English term meditation (defined in the Oxford dictionary primarily as "focussing one's mind for a period of time, in silence or with the aid of chanting, for religious or spiritual purposes or as a method of relaxation."), we decided to label this activity "thinking".

In the WHO SAGE study, participants reported their activities in chronological order (up to a maximum of 10 over one day) chosen from a mutually exclusive list of 23 activities [in square brackets are the labels that we used in figures]: working [working]; subsistence farming [subsistence farming]; preparing food [cooking], doing housework [housework], watching children [childcare], shopping [shopping], walking somewhere [walking], traveling by bicycle [traveling by bike], traveling by car/bus/train [transport], rest (includes tea/coffee break) [resting], chatting with someone [chatting], playing (include cards/games) [playing], reading [reading], listening to radio [radio], watching TV [television], exercising or leisurely walk [exercising], other leisurely activity [leisure], grooming or bathing (self) [grooming], eating [eating], religious activity [religion], providing care to someone [care], intimate relations/sex [sex], went to sleep for the night [going to bed]. We excluded activities that were represented in fewer than 0.1% of the samples within each group (history of depression and no history of depression). This resulted in "intimate relations/sex" being excluded from the analysis.

In the WHO SAGE study, participants were asked to report the positive and negative emotions that they felt from a list including: worried, rushed, irritated or angry, depressed, tense or stressed, calm or relaxed, enjoying. Each emotion was encoded as not at all (=1), a little (=2), or very much (=3). A total mood score was calculated by subtracting the mean of the negative emotions (worried, rushed, irritated or angry, depressed, tense or stressed) from the mean of positive emotions (calm or relaxed, enjoying). The maximum mood score was therefore 2 and the minimum was -2. To remain consistent with the notations in the 58sec dataset, we mapped this range to [0,100] by adding 2 and multiplying by 25 the mood score. This simple mapping has no bearing on the findings and was used for presentation only.

eMethods 2. Description of the 58sec Project

The "58sec" research project is a multi-lab collaborative effort to collect large scale experience-sampling data in the general population. At its launch in 2013, the research project and associated iOS and Android applications were featured in a popular TV show called "J'ai décidé d'être heureux" on the French TV channel M6, and further relayed on the "Journal Télévisé" of the Belgian TV channel RTBF.

eMethods 3. Formal Definition of Mood Homeostasis

We let M_t and M_{t+1} be the mood of an individual at time t and $t+1$ respectively (i.e., recorded by two consecutive questionnaires) and A^j_t and A^j_{t+1} be dichotomous variables denoting whether the participant was engaged in the j-th activity at time t and $t+1$. To assess the relation between current mood (M_t) and subsequent choices of activity (A^j_{t+1}), logistic regressions were used:

$$logit\ (A^j_{t+1}) = \beta^j_0 + \beta^j_c\, M_t + \beta^j_h\, H + \beta^j_d\, D + \beta^j_a\, A^j_t + \beta^j_m\, M_{avg}, \quad \text{for all } j, \qquad [1]$$

where β^j_c is the coefficient of interest which relates current mood to the probability of subsequently engaging in a specific activity, and H, D, A^j_t and M_{avg} are covariates that need to be controlled for (see eMethods 6). This is what we refer to as the logistic regression in the main text. A highly positive value of β^j_c implies that individuals are much more likely to engage in the j-th activity at time $t+1$ if their mood is high at time t. Conversely, a highly negative value of β^j_c implies that individuals are much more likely to engage in the j-th activity at time $t+1$ if their mood is low at time t. These coefficients can also be interpreted in terms of odd ratios as follows. All other things being equal, if two individuals differ in their mood at time t by an amount q, their propensity (expressed as odd ratios) to engage in the j-th activity at time $t+1$ differ by a factor

$$OR_j = e^{\beta jc \times q}$$

There is therefore a one-to-one relationship between the coefficient β^j_c and this probability. This is used for graphical representations (using q=-24.3 for the 58sec dataset, and q=-18.7 in the WHO SAGE dataset which correspond to 1 standard deviation in mood in the respective population).

Similarly, to assess the relation between each activity and the corresponding change in mood compared to the previous record ($\Delta M_{t+1} = M_{t+1} - M_t$), a linear regression was used:

$$\Delta M_{t+1} = \Delta_0 + \Sigma_j \Delta_j\, A^j_{t+1} + \ \Delta_h\, H + \Delta_d\, D, \qquad [2]$$

where the H and D are covariates that need to be controlled for (see eMethods 6). This is what we refer to as the linear regression in the main text. The values Δ_j are the coefficients of interest which correspond to the change in mood that results from an individual engaging in the j-th activity. For instance, if Δ_j=10, then an individual's mood increases on average by 10 points compared to their mood at time t (all other things equal) when they engage in the j-th activity at time $t+1$.

In the main text, we mention that "mood homeostasis is positive if the probability of next engaging in an activity when current mood is low (estimated with one regression) is negatively correlated with the change in mood resulting from this activity (estimated with another regression)". Formally, this implies that mood homeostasis is quantified by the negative correlation between the β^j_c, one per activity, in model [1] (which encode the effect of current mood on the probability of subsequently engaging in an activity) and the Δ_j, one per activity, in model [2] (which encode the resulting change in mood). We therefore define it as:

$$MH = -\ Cor(\boldsymbol{\beta_c}, \Delta),$$

where the bold symbols represent the vector of coefficients for all activities and *Cor* is the Pearson correlation coefficient.

A high level of mood homeostasis means that the larger an increase in mood an activity provides, the more likely individuals are to engage in it when their current mood is low (and vice versa when their current mood is high).

eMethods 4. Illustration of the Practical Calculation of Mood Homeostasis

Imagine two flat mates, Luke and Helen, who are reporting what they do and how they feel using a smartphone application (like the 58sec app). Helen is not an early bird and reports a mood of 60 while she is getting ready for work. While having breakfast, she is listening to a fascinating podcast on ancient Egypt and reports a mood of 66. She then starts her commute and records a mood of 64. Half-way through her commute to work, she starts listening to her favorite playlist and she reports her mood as 69. Upon arriving at work, she checks and answers her emails and starts thinking about the outstanding tasks that need completing. She reports a mood of 61. Thereafter, she takes a coffee break with her friend and colleague with whom she chats about "anything but work". She reports a mood of 70. Helen would have a high mood homeostasis: whenever her mood is relatively low, she engages in an activity that will increase it (e.g. listening to music, chatting with a friend) and she engages in mood-decreasing activities (e.g. thinking, working) when her mood is relatively higher. Let us say that her average mood throughout the experience is 67.

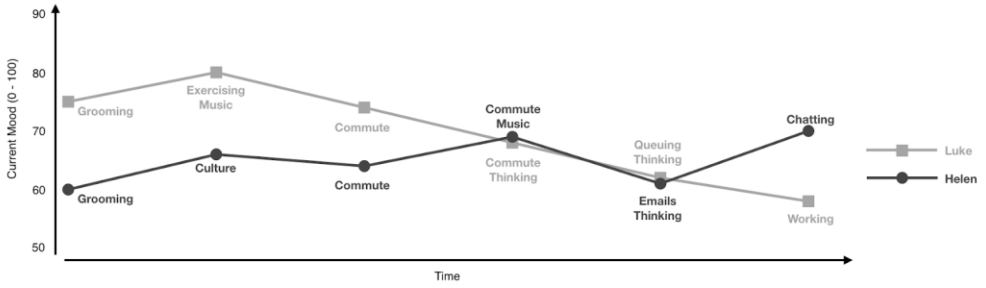
Luke is an early bird and reports a mood of 75 on wakening. He starts his day with a bit of exercise which he does while listening to music. At that point his mood is 80. He then starts his commute to work and reports a mood of 74. Half-way through the commute, he starts thinking about his day and the stress that he is facing at work at the moment. He reports a mood of 68. Upon arriving at work, he waits in the queue to buy coffee while still ruminating. His mood is then 62. He then starts to work which he does not derive much pleasure from and reports his mood as 58. Luke would have a low mood homeostasis: when feeling relatively low, he engages in further mood-decreasing activities (ruminating, waiting, working) and he engages in mood-increasing activities (exercising) when his mood is already relatively high. Let us say that Luke's average mood throughout the experience is 60.

The input data resulting from Luke and Helen's recording of their morning mood and activities are illustrated in eFigure. 1. Two regression models are then estimated using these input data (also illustrated in eFigure 1). A first regression model is used to assess the association between an individual's current mood and the future propensity to engage in a specific activity (so there are 25 regressions, one per activity). The coefficient β (the "slope") of the model for each activity is recorded. Owing to the interaction term in the regression, this coefficient is a function of an individual's average mood so that 25 coefficients β (one per activity) can be recorded for people with high average mood and a different set of 25 coefficients β can be recorded for people with low average mood. A second regression model is used to assess the change in mood (Δ) resulting from engaging in a particular activity. There are as many Δ as there are activities (i.e., 25) and again these coefficients are functions of an individual's average mood owing to the interaction term in the regression. A total of 25 coefficients Δ (one per activity) can therefore be recorded for people with high average mood and a different set of 25 coefficients Δ can be recorded for people with low average mood.

In each group (people with high and people with low average mood), mood homeostasis can finally be calculated as the negative correlation between the 25 coefficients β and the 25 coefficients Δ.

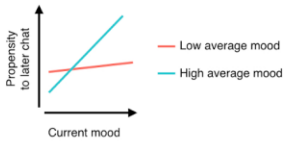
eFigure 1. Procedure Followed to Calculate Mood Homeostasis From the Input Data and From the 2 Regression Models

Input data: Mood and activities recorded alongside covariates (time of day, day of the week, etc)



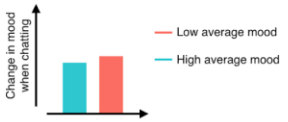
User ID	Average mood	Day	Time	Current mood	Later Chatting	Later Commute	Later Culture	Later Exercising	Later Listening to music	Later Thinking	Later Waiting	Later Working	Change in mood
Helen	67	Monday	06:48AM	60	0	0	1	0	0	0	0	0	+6
Helen	67	Monday	07:20AM	66	0	1	0	0	0	0	0	0	-2
Helen	67	Monday	08:05AM	64	0	1	0	0	1	0	0	0	+5
Helen	67	Monday	08:45AM	69	0	0	0	0	0	1	0	1	-8
Helen	67	Monday	09:32AM	61	1	0	0	0	0	0	0	0	+9
Luke	60	Monday	07:15AM	80	0	0	0	1	0	0	0	0	+5
Luke	60	Monday	08:05AM	74	0	1	0	0	0	0	0	0	-6
Luke	60	Monday	08:45AM	68	0	1	0	0	0	1	0	0	-6
Luke	60	Monday	09:10AM	62	0	0	0	0	0	1	1	0	-6
Luke	60	Monday	09:20AM	58	0	0	0	0	0	0	0	1	-4

First regression model: Association between current mood and the probability of later engaging in a particular activity



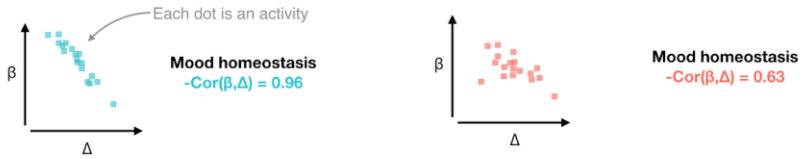
For each activity, we record the **slope (β)** of the regression line for low and high average mood

Second regression model: Association between each activity and the resulting change in mood



For each activity, we record the **change in mood (Δ)** for low and high average mood

Calculation of mood homeostasis: this is the negative correlation between the β and the Δ of all activities



Note: Graphs in Fig. 2C are equivalent to those above except that the β's have been replaced by OR = $e^{-24.3\beta}$ as described in eMethods 3.

The input data presented at the top are those described in the illustrative example of eMethods 4.

eMethods 5. Mathematical Demonstration of the Invariance of Mood Homeostasis

In this section, we demonstrate that mood homeostasis remains the same if an individual's mood level is shifted up or down or if the amplitude of the oscillations in mood are increased or decreased (see eFigure 2). This means that if two individuals or two groups of individuals differ only by their average mood or by the amplitude of their mood fluctuations, they will have the exact same mood homeostasis. This is an important property as it guarantees that any difference in mood homeostasis between two groups is not due to their difference in average mood (this is particularly important in the 58sec dataset where we contrast individuals based on their average mood). Intuitively, this property follows from the fact that mood homeostasis is formally defined as a correlation coefficient between two quantities and correlation coefficients remain unchanged if either or both these quantities are multiplied by a constant or if a constant is added to them. In this section, we provide a formal demonstration that this is indeed the case.

A difference in mood level implies that all mood records across the timeline are shifted up or down by a constant C. In other words, this implies that all M_t and M_{t+1} are replaced by M_t+C and $M_{t+1}+C$ in equations [1] and [2]. A difference in the amplitude of mood fluctuations implies that all mood records across the timeline are multiplied by a constant K, that is M_t and M_{t+1} are replaced by $K \times M_t$ and $K \times M_{t+1}$ in equations [1] and [2] (and that a constant is then added or subtracted from the result to keep the same mean value). More generally, changes in levels and fluctuation amplitudes of mood records therefore amounts to replacing M_t and M_{t+1} in equations [1] and [2] by $M_t' = K \times M_t + C$ and $M_{t+1}' = K \times M_{t+1} + C$. We now demonstrate that such a replacement does not alter the value of mood homeostasis. We proceed by mathematically deriving how the coefficients of interest $\boldsymbol{\beta_c}$ and Δ change as a result of the substitution of M_t and M_{t+1} by $M_t' = K \times M_t + C$ and $M_{t+1}' = K \times M_{t+1} + C$ and then by showing that these changes do not affect the value of the correlation between them. $\boldsymbol{\beta_c}'$ and Δ' denote the values of the coefficients of interest after substitution.

After substitution of M_t and M_{t+1}, equation [1] reads (where, for the sake of generalizability, we write the covariates as $\Sigma_k \beta^j_k' X_k$ which in the present case is equal to $\beta^j_h' H + \beta^j_d' D + \beta^j_a' A^j_t + \beta^j_m' M_{avg}$):

$$\begin{aligned} logit\,(A^j_{t+1}) &= \beta^j_0' + \beta^j_c' (K \times M_t + C) + \Sigma_k \beta^j_k' X_k \\ &= (\beta^j_0' + \beta^j_c' C) + \beta^j_c' \times K \times M_t + \Sigma_k \beta^j_k' X_k \end{aligned}$$

Since the left-side of this equation has not changed, it must remain equal to the right side of equation [1]. Therefore, we have:

$$(\beta^j_0' + \beta^j_c' C) + \beta^j_c' \times K \times M_t + \Sigma_k \beta^j_k' X_k = \beta^j_0 + \beta^j_c M_t + \Sigma_k \beta^j_k X_k .$$

Since this equality must hold for any values of M_t and the covariates X_k, it follows that

$$\begin{aligned} &\beta^j_0 = (\beta^j_0' + \beta^j_c' C), \\ &\beta^j_c' = \beta^j_c / K, \text{ and} \\ &\beta^j_k' = \beta^j_k . \end{aligned}$$

So the coefficients of interest $\boldsymbol{\beta_c}'$ (which is a vector made of all the β^j_c') are equal to the original ones $\boldsymbol{\beta_c}$ divided by the constant K.

Since $\Delta M_{t+1} = M_{t+1} - M_t$, its new value after substitution of M_t and M_{t+1} is:

$$(K \times M_{t+1} + C) - (K \times M_t + C) = K \times (M_{t+1} - M_t) = K \times \Delta M_{t+1}.$$

As a result, equation [2] becomes (where, for the sake of generalizability, we write the covariates as $\Sigma_i \beta_i' Y_i$ which in the present case is equal to $\Delta_h' H + \Delta_d' D$):

$$\begin{aligned} &K \times \Delta M_{t+1} = \Delta_0' + \Sigma_j \Delta_j' A^j_{t+1} + \Sigma_i \beta_i' Y_i , \\ &\Rightarrow \quad \Delta M_{t+1} = (\Delta_0'/K) + \Sigma_j (\Delta_j'/K)\, A^j_{t+1} + \Sigma_i (\beta_i'/K)\, Y_i . \end{aligned}$$

And since the left side of this equation is now equal to the left side of the original equation [2], their right sides must also be equal (for all values of A^{j}_{t+1} and Y_i) and it follows that:

$$\Delta_0' = \Delta_0 \times K,$$
$$\Delta_j' = \Delta_j \times K, \text{ and}$$
$$\beta_i' = \beta_i \times K.$$

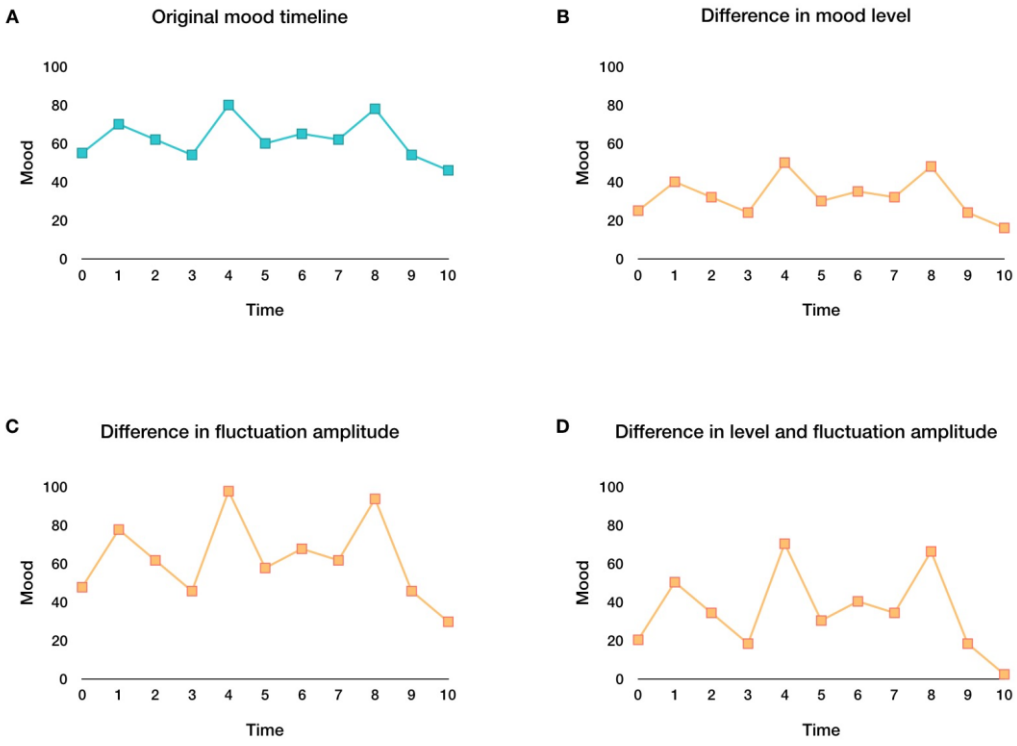
So the coefficients of interest Δ' (which is a vector made of all the Δ_j') are equal to the original ones Δ multiplied by the constant K.

The mood homeostasis after substitution of M_t and M_{t+1} by their new values therefore becomes:

$$\begin{aligned} MH' &= -Cor(\boldsymbol{\beta_c}', \Delta') \\ &= -Cor(\boldsymbol{\beta_c}/K, \Delta \times K) \\ &= -Cor(\boldsymbol{\beta_c}, \Delta) \\ &= MH \end{aligned}$$ □

and is this equal to the original value before substitution. The last step of this demonstration is a simple consequence of the fact that the Pearson correlation coefficient is invariant to the multiplication of its inputs by a constant (independently for both inputs).

eFigure 2. Mood Homeostasis is Invariant to Difference in Mood Level, Difference in Fluctuation Amplitude, and Combinations Thereof



These differences are illustrated here for an original timeline (A), and differences in level (B), amplitudes (C), and combinations thereof (D).

eMethods 6. Covariates

The choice of covariates to include in the regression models followed from previous studies[1,2]. In the latter, different regression models (with different sets of covariates) were estimated and the best set of covariates was selected based on their Akaike Information Criterion (AIC). Covariates included in regression models [1] and [2] were encoded as follows:

1) The time of day was represented as a categorical variable (denoted by *H* in the regression models above) which can take 12 values, by binning the time in 12 periods of 2 hours each (i.e., 00:00:00 AM–1:59:59 AM, 02:00:00 AM - 3:59:59AM, …, 10:00:00 PM–11:59:59 PM).

2) The day of the week was only available for the 58sec dataset as this information was not part of the WHO SAGE data. It was represented as a categorical variable (denoted by *D* in the regression models) with 3 categories (weekday, Saturday, or Sunday).

3) The latency effect is simply represented by the variable A^j_t which is equal to 1 if an individual was already engaged in the *j*-th activity at time *t*, and 0 otherwise.

4) The daily average mood (denoted by M_{avg} in the regression models) was obtained by averaging all other mood reports recorded on the same day as M_t. Inclusion of this covariate is important as our goal is to capture the high frequency dynamics in mood (e.g., hourly changes) while controlling for the low frequency dynamics (e.g., weekly changes). This guarantees that associations between current mood and subsequent choices of activity are not merely reflecting longer-term affective trends.

The time difference between two activities was not included as a covariate for two reasons. First, this variable cannot be meaningfully defined in the WHO SAGE dataset since the participants were asked to name the consecutive activities that they were engaged in under the assumption that there was no gap in time between two consecutive activities. Second, in a previous study[1], we found that this time difference had little to no impact on the main result. To keep the design of the analysis as consistent as possible between the two datasets and to avoid unnecessarily complexifying it, time difference was therefore not included as part of the covariates.

eMethods 7. Details of Statistical Analysis

To assess how a variable *x* affects mood homeostasis, we added it as an interaction term in models [1] and [2]:

$$\beta^j_c(x) = \beta^j_{c0} + \beta^j_{cx}\, x \qquad [3]$$

$$\Delta_j(x) = \Delta_{j0} + \Delta_{jx}\, x, \qquad [4]$$

so that mood homeostasis becomes:

$$MH(x) = -Cor[\boldsymbol{\beta}_c(x), \Delta(x)].$$

In the 58sec dataset, *x* denotes an individual's average mood. *MH(x)* is a curve which can be evaluated at any value of *x*. In the WHO SAGE dataset, *x* is a dichotomous variable encoding an individual's history of depression (*x=1*) or the absence thereof (*x=0*).

Non-parametric bootstrap was used to calculate 95% confidence intervals and p-values[3]. Unlike parametric tests, this approach accounts for the fact that the few data points (one per activity, i.e., 25 in the 58sec dataset and 22 in the WHO SAGE dataset) were estimated from large datasets.

Group differences in mood homeostasis imply that the relationship between the propensity to engage in activities as a function of current mood (β_c) and the resulting change in mood (Δ) differs between the two groups. In a plane where the x-axis represents Δ and the y-axis represents β_c (or the corresponding odd-ratio as in Fig. 7C-D), all activities can be represented by a dot. If all the dots are perfectly aligned, then mood homeostasis is maximum. A lower value of mood homeostasis must therefore follow from a relative misalignment of the dots. As depicted in

eFigure 3, we explore three factors that might drive such misalignments (and thus group differences in mood homeostasis): specific activities (i.e., an activity is so misaligned with the others that it suffices to cause a significant difference in mood homeostasis), a specific dimension (i.e., whether the misalignment is driven by the propensity to engage in activities $\boldsymbol{\beta_c}$ or by the resulting change in mood Δ), or a specific valence (i.e., whether the misalignment is driven by the relation between $\boldsymbol{\beta_c}$ and Δ among activities with a positive Δ or among activities with a negative Δ).

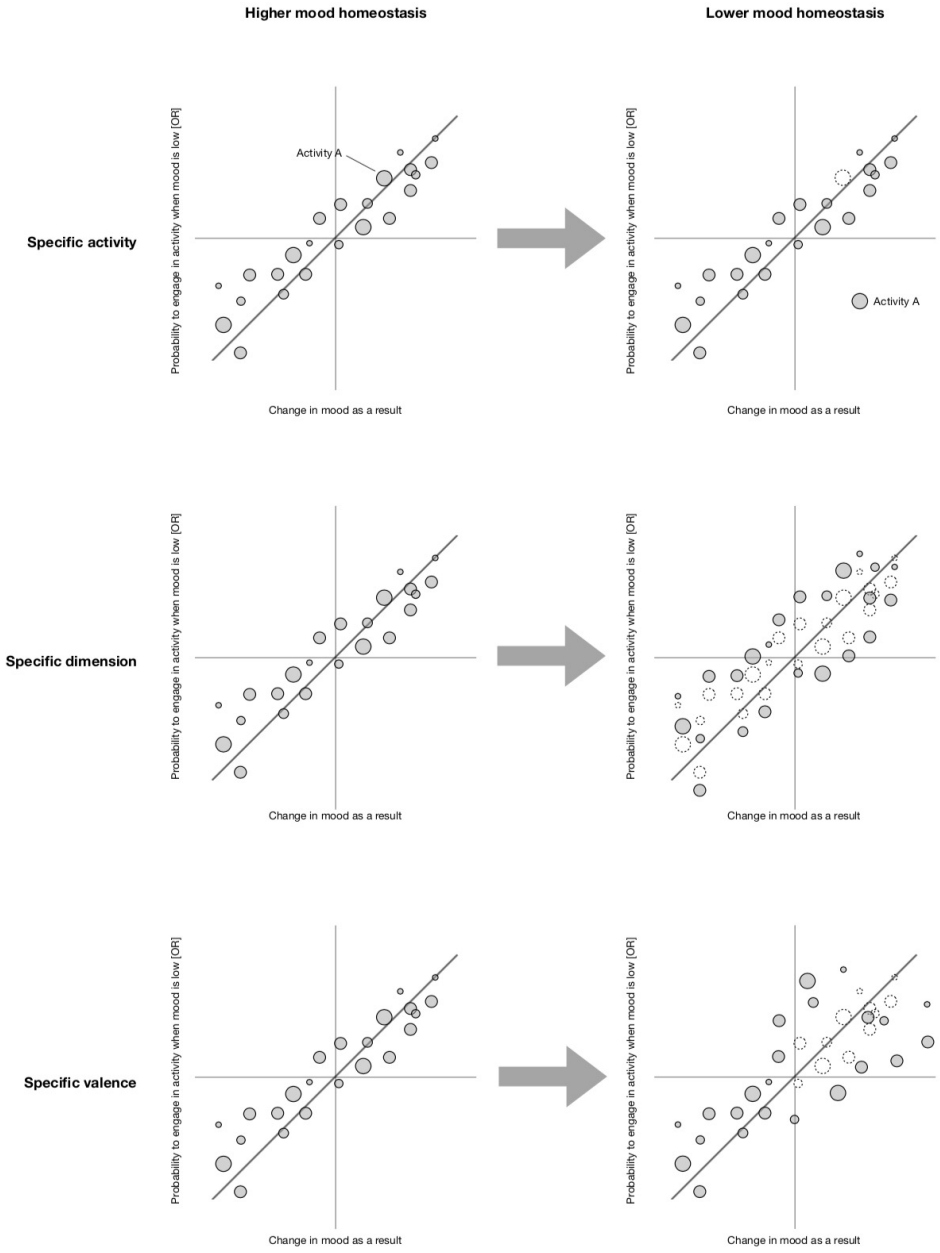
To test whether one or several specific activities drive the group differences in MH, we first computed the line of best fit between $\boldsymbol{\beta_c}$ and Δ within the group with higher MH (eFigure 3). We then calculated, within the group with lower MH the distance between each activity point and this line. We calculated the mean and SD of these distances. An activity is considered to be driving differences in MH if its distance to the line of best fit is larger than 3 SD above the mean. Values of mood homeostasis, their confidence intervals, and their group-differences were then calculated within the remaining activities to assess whether other factors are likely to drive the group differences in mood homeostasis too.

To test whether a specific dimension drives the group differences in MH, we tested whether groups differ most in terms of their propensity to engage in different activities $\boldsymbol{\beta_c}$ or in terms of their resulting change in mood Δ. To do so, we calculated (i) the Pearson correlation coefficient $C_{\boldsymbol{\beta}}$ between the coefficients $\boldsymbol{\beta_c}$ in the group with higher MH and the same coefficients in the group with lower MH, and (ii) the Pearson correlation coefficient C_{Δ} between the coefficients Δ in the group with higher MH and the same coefficients in the group with lower MH. We claim that one dimension drives the difference in MH if its correlation is significantly larger (at $P < .05$) than the correlation of the other dimension (e.g., if $C_{\boldsymbol{\beta}}$ is significantly larger than C_{Δ}).

To test whether a specific valence drives the group differences in MH, we first calculated restricted mood homeostasis MH^+ and MH^- by restricting the coefficients $\boldsymbol{\beta_c}$ and Δ to the activity with positive and negative Δ respectively (referred to as positive and negative valences). We compared the MH^+ and MH^- in the two groups and claim that the positive valence is driving the group differences in MH if the difference in MH^+ between the groups is significantly larger than the difference in MH^- (and vice versa to claim that the negative valence is driving group differences in MH).

Statistical analyses were performed in R 3.4.3.

eFigure 3. Different Factors May Drive Group Differences in Mood Homeostasis



On the left is a hypothetical relationship between the propensity to engage in a particular activity and the resulting change in mood (leading to high mood homeostasis as observed from the good alignment of dots). On the right are different patterns of misalignment of dots which result from the three investigated factors. Dashed lines represent the original locations of dots to help the visualization. (Top) A single outlier activity drives the difference. (Middle) A single dimension (here, the probability to engage in an activity) drives the difference. (Bottom) A single valence (here activities with a *positive* Δ) drives the difference.

eMethods 8. Robustness Analysis

We tested the robustness of the group differences in mood homeostasis by reproducing the analysis in the following four scenarios.

1) *In two mutually exclusive subsets of the data randomly selected from the original datasets*
The 58sec dataset was randomly split in two subsets and the whole analysis was conducted independently in the two subsets. The WHO SAGE dataset was randomly split it two in a way that keeps the same number of observations from people with a history of depression in the two subsets. The whole analysis was then conducted independently in the two subsets.

2) *Using parametric tests rather than bootstrap tests*
Bootstrap tests were used in the primary analysis as they account for the fact that even if mood homeostasis is driven by 25 data points (one per activity), the values (β_i, Δ_i) of these data points are themselves based on data from about 30,000 individuals. Specifically, confidence intervals on the Pearson's correlation coefficient obtained from parametric t-tests assumes that the underlying population-wise values (β_i, Δ_i) come from a bivariate normal distribution whereas the bootstrap test used in the main analysis makes no assumption and leverages individual-level data by bootstrapping individual mood and activity records. As a robustness test however, we repeated the analysis using parametric t-tests with 23 (=25 - 2) degrees of freedom to obtain parametric confidence intervals for each value of mood homeostasis (achieved using cor.test in R version 3.4.3). The difference between these correlations was achieved using a parametric z-test of the difference of the Fisher's z-transformed correlations divided by the standard error of the difference[4] (as implemented in the function r.test of the R-package "psych" version 1.8.4).

3) Using multilevel models with random intercepts
Entries in both the 58sec and the WHO SAGE datasets contain values to all the variables in Equation [1] and [2] (M_t, M_{t+1}, $\boldsymbol{A_t}$, $\boldsymbol{A_{t+1}}$, *H, D, …*) as well as the user identifier. The latter can be integrated in the analysis using a multilevel regression approach wherein a random intercept is estimated for each individual independently. The reason for using a fixed intercept is that the number of data entries per individual can be low. In the WHO SAGE dataset in particular, individuals only reported their activities over a single day. Some of them only engaged in 2 or 3 activities leading to 1 or 2 data entry (as each entry is made of 2 moments to compose the current mood M_t and activities A_t, and the future mood M_{t+1} and activities A_{t+1}). Estimating a random intercept for these individuals may be ill-posed (i.e. no single answer exists) or ill-conditioned (i.e., a single answer exists but algorithms to calculate it are not guaranteed to find it because the optimization is unstable). Estimations of the random intercepts may therefore be unreliable (which was confirmed in both datasets by warnings issued by the R code used to compute them). To try and achieve stable estimates, the same code needs to be executed multiple times (with different random initializations of the values) until convergence is reached. But the accuracy of these estimates remains unknown. Furthermore, because this process has a high computational time, it is not well suited to bootstrap testing (as it would need to be repeated 1000 times). We therefore opted for fixed intercepts in the main analysis. As a robustness test however, we calculated the results with random intercepts to assess how different they were from the main results. Since bootstrap tests with multilevel modeling was computationally intractable, parametric tests were used for the group difference in this robustness analysis (as described in the above robustness test). The impact of the multilevel approach must therefore be compared to the parametric version of the main results.

4) Adjusting for country of origin (for the WHO SAGE dataset only)
Country of origin (a categorical variable taking on 6 values: China, Ghana, India, Mexico, Russia, or South Africa) was included as a covariate in the regression models [1] and [2] and the whole analysis was otherwise repeated in the same way as the primary analysis.

eMethods 9. Dynamic Simulation Process

Using simulations, we can alter one variable (in this case, mood homeostasis) and assess the impact of this variable on future outcome (in this case, depressive episodes). Because all other things are kept equal, this can demonstrate a plausible causal effect between the variable of interest and the outcome. It is only plausible and not actual because the world of simulations might not accurately represent the real world. The existence of a causal link in simulations is therefore a necessary—but not sufficient—condition for demonstrating actual causality.

Dynamic simulation was achieved by initializing a simulated individual's mood (M_0) and activities (A^j_0) randomly (using a uniform distribution between 0 and 100 for M_0 and independent Bernoulli distributions with probability p=0.5 for the A^j_0) and then iterating between Equations [1] and [2] to generate subsequent moods ($M_{t+1} = M_t + \Delta M_{t+1}$) and activities ($A^j_{t+1}$). All covariates of Equations [1] and [2] were kept apart from daily average mood (as it depends on moods simulated only after the one being simulated). The values of the coefficients in Equation [1] and [2] were those estimated with the 58sec dataset including the interaction terms of Equation [3] and [4]. This guarantees that values obtained via the simulation are as realistic as possible. Inclusion of the interaction term enables to modulate mood homeostasis in a realistic way. The value of x (which is the average mood of the individual scaled across the population) was set to 0 (i.e., equal to the population mean) for 100 simulated individuals thereby leading to high mood homeostasis (as seen in Fig. 6) and -2 (i.e., equal to 2 SD below the population mean) for 100 simulated individuals thereby leading to low mood homeostasis.

Twelve time points per day were simulated for 5 years for each of these simulated individuals, so that each simulated individual was represented by two 5-year timelines: one for activities (each timepoint has 25 binary values) and one for mood (each timepoint has one continuous value). From the randomly selected initial state (mood and activities), the timelines first progressively converge to a steady state. During this transient state, inferences based on the pattern of mood and activities are unreliable—a well-known phenomenon of stochastic processes and dynamic simulations[5]. This was dealt with by discarding the first 5,000 timepoints (which were observed to fully include the transient state) before keeping the next 5 years of simulated data. The mood timelines were then standardized for each simulated individual independently so that a mood of 0 represents that individual's overall average mood. This is important as depressive episodes are defined based on a mood level that is abnormal compared to one's baseline. Depressive episodes were defined using a quantitative interpretation of the mood component of ICD-10 criteria[6] (and ignoring additional symptoms, e.g. fatigability, low self-esteem, etc.): A mood lower than 2 standard deviations below the simulated individual's average mood (corresponding to "depressed mood to a degree that is definitely abnormal for the individual" in ICD-10), consistently over the 12 time points of the day ("present for most of the day"), every day ("almost every day") for 2 weeks ("the depressive episode should last for at least 2 weeks"). The difference in the number (Δn) and duration (Δd) of depressive episodes between the two groups were subject to non-parametric permutation tests using 10,000 permutations. At each permutation, simulated individuals were randomly reallocated to either group and the surrogate differences Δn^* and Δd^* were calculated. The p-values for the null hypothesis that the Δn and Δd equal 0 were then given by:

$$p = N_{>}/10{,}000,$$

where $N_{>}$ is the number of permutations which lead to $|\Delta n^*| > |\Delta n|$ and $|\Delta d^*| > |\Delta d|$ respectively.

eResults 1. Interpretation of the Effect Sizes

It can be difficult to interpret the effect size of a change in mood in absolute terms. For instance, what does it mean for thinking to cause a decrease in mood by 3 points among people with low average mood? There are four ways to interpret these effect sizes in clinically more meaningful terms:

1) It is helpful to interpret effect sizes in terms of the range of changes in mood that individuals experience (i.e., the difference between the maximum and minimum mood value reported by an individual). Indeed if an individual reports moods that mostly vary from 45 to 60, a change of 3 points corresponds to 20% of the range in mood that is explored, which is substantial. Conversely, if their mood varies across the range from 0 to 100, then a change of three points might indeed seem negligible. The median range of moods that individuals reported in the 58sec dataset was 33, so that a change in mood of 3 points corresponds to 9% of their normal range, which is far from trivial.

2) The changes in mood reported correspond to those resulting from engaging in a specific activity at a specific time. The effects of engaging in multiple activities are *additive*. For instance, if an individual in the low mood homeostasis group is both thinking and waiting then this will decrease their mood on average by 5 points (i.e., a change of 3 points for thinking and a change of 2 points for waiting). In the 58sec data, the average change in mood (in absolute value) at any one time was 11.9 points which is substantially larger than the average mood change resulting from any single activity (thus demonstrating the additive effect). Expressed in relative terms, the average change in mood at any one point was 25.1% of the individual's mood range. In addition, the effects over time are *cumulative*. For instance, if an individual is commuting (-0.5 points) and then starts thinking during the commute (-3.5 points), continues to think while waiting for the elevator (-5 points) and then gets to work (-0.6 points), then their mood would be on average 9.6 points lower than before they started commuting. This simple (and arguably not uncommon) sequence of activities would therefore cause a decrease in mood corresponding to 29% of the median range of mood reports. This is also reflected in the 58sec data: after only 2 records, the average cumulative mood change was 15.6 points (or 28.1% of the individual mood range) showing that consecutive mood changes tend to accumulate on average (since 15.6 is larger than 11.9). This further supports that the change in mood is far from small.

3) Another helpful benchmark to compare an absolute change in mood to is the difference between weekdays and weekends. It is a robust finding that mood is affected by the day of the week[7]. Comparing the change in mood resulting from a specific activity (e.g., thinking) with the change in mood resulting from the day of the week thus enables a better interpretation of the effect sizes. In the 58sec dataset, people's mood decreased by 1.8 points on average from weekends to weekdays. Therefore, thinking has an impact on mood that is 170% greater than the impact of the day of the week.

4) Finally, a primary motivation for running the simulation analysis was to assess whether effect sizes were clinically (and not just statistically) significant. Indeed those simulations account for the different other factors that affect mood and choices of activities (such as time of day and day of the week). If the effect of an activity on the change in mood was too small in comparison to the variability of mood resulting from the day of the week and the time of the day, then one would not expect to observe clinically significant differences in the incidence and duration of depressive episodes through simulations. The fact that low mood homeostasis leads (in simulation) to clinically significant differences in the incidence and duration of depressive episodes testifies to the clinical significance of the changes in mood that result from the different activities.

eResults 2. Differences in Mood Homeostasis Between the 2 Data Sets

The value of mood homeostasis differs between the group with high average mood in the 58sec dataset (MH = 0.961) and the group without a history of depression in the WHO SAGE dataset (MH = 0.675). As detailed in the discussion, different factors might explain this difference and a complete account of this difference would require a new study wherein the same methodology is used across high- and low-income countries. However, using the data at hand we can start to explore this question by testing four factors that might play a role in the difference in mood homeostasis: sampling duration, average mood level, age, and income.

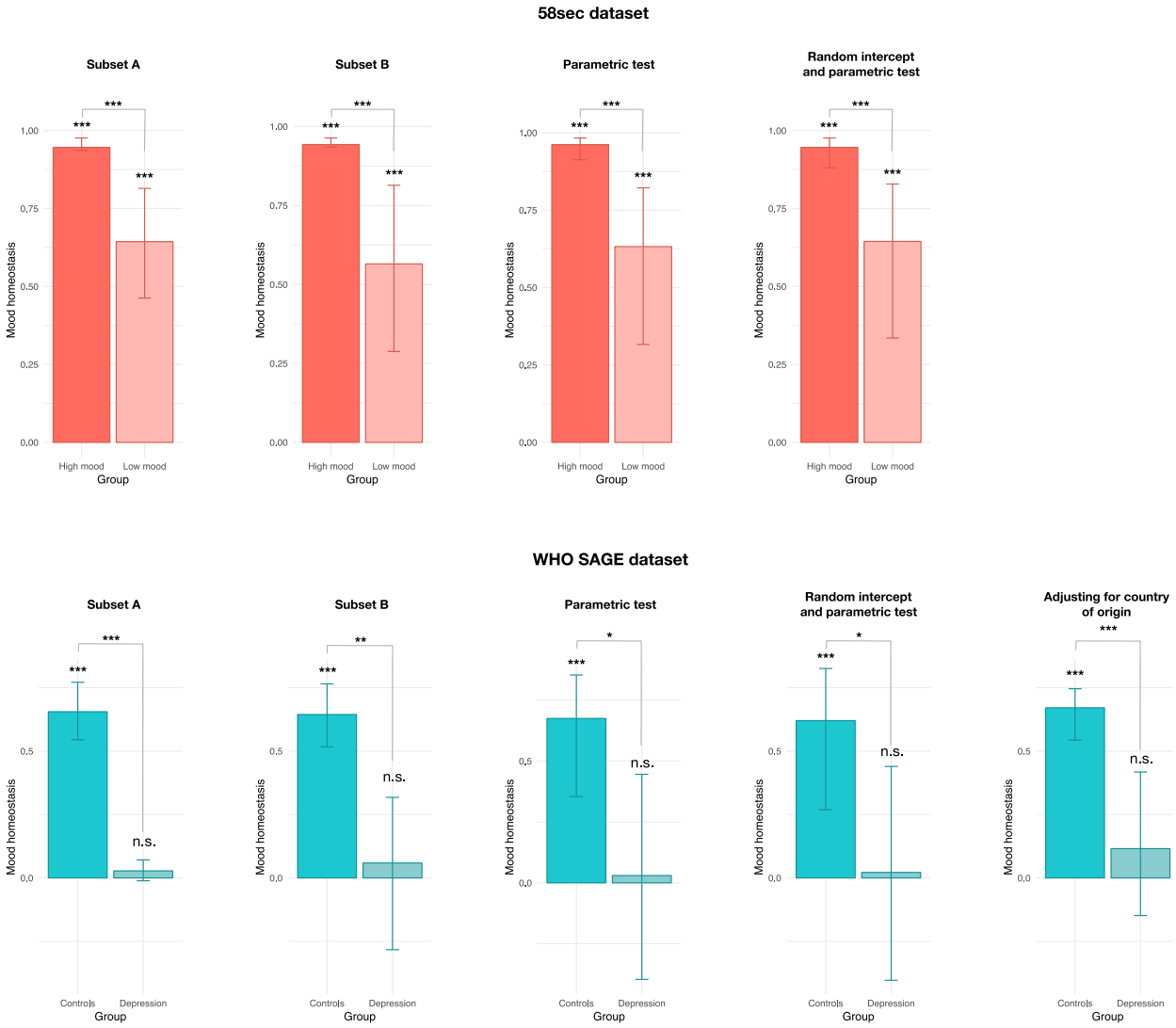
Sampling duration Participants in the 58sec dataset recorded their mood and activities for one or several days whereas participants to the WHO SAGE study provided records for only one day. To test whether this might explain the difference in mood homeostasis observed between the two datasets, the data from the 58sec dataset was restricted to the first day of recording for each participant. The same approach as in the primary analysis was used. This resulted in a mood homeostasis among participants with high average mood of 0.843 (95% C.I. [0.794, 0.932]) which is lower than the value obtained using all the data, but still significantly higher than the value among people without a history of depression in the WHO SAGE data (bootstrap $P < .02$; eFigure 7). This result therefore shows that sampling duration may explain part of (but not all) the difference in mood homeostasis between the datasets.

Average mood level In the 58sec dataset, we observed differences in mood homeostasis between the top half and the bottom half of the population in terms of average mood. Since the mood scales differ between the two datasets, it is impossible to know exactly what the average mood of controls in the WHO SAGE dataset would be on the 58sec mood scale. But there is no reason to believe that their average mood would correspond to that of the *top half* of the 58sec dataset. As the controls in the WHO SAGE dataset make up 97% of the sample, it might be more realistic to assume that their average mood corresponds to that of the top 97% of the 58sec dataset. To assess the compounded effect of differing average moods and sampling duration, we assessed the mood homeostasis for an average mood equal to the mean in the top 97% of the 58sec participants while only using the first day of recording. This resulted in a mood homeostasis of 0.709 (95% C.I. [0.422,0.862]) which is not statistically significantly different from the mood homeostasis observed among controls in the WHO SAGE dataset ($P > .1$; eFigure 7). In other words, the different in mood homeostasis between the two datasets might be entirely accounted for by differences in sampling duration and levels of average mood. But this does not rule out the role of other factors.

Age The influence of age was assessed by estimating mood homeostasis within each age quartile of the WHO SAGE dataset (corresponding to 18-50, 51-58, 59-67, and > 67 years old). This additional analysis found no statistically significant difference in mood homeostasis across age ranges (eFigure 8A).

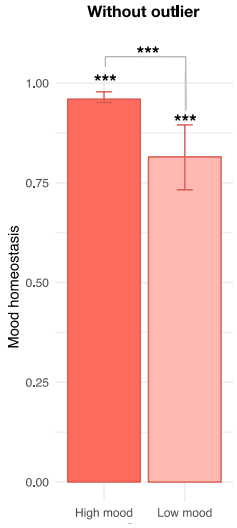
Income The variable most closely related to income within the WHO SAGE dataset was the answer to the question "Do you have enough money to meet your needs?" to which 99.5% of participants answered by choosing between the following options: Completely, Mostly, Moderately, A little, None at all. To keep subsamples of similar sample sizes, we grouped "Completely" and "Mostly" within one category (which we named "high income" and which consisted of 30,602 observations) and "A little" and "None at all" within another (which we named "low income" and which consisted of 37,914 observations). The group of participants who answered "Moderately" was named "medium income" and consisted of 32,986 observations. No significant differences in mood homeostasis was observed between the three subgroups (eFigure 8B).

eFigure 4. Results of the Robustness Analyses Carried Out in Both Data Sets



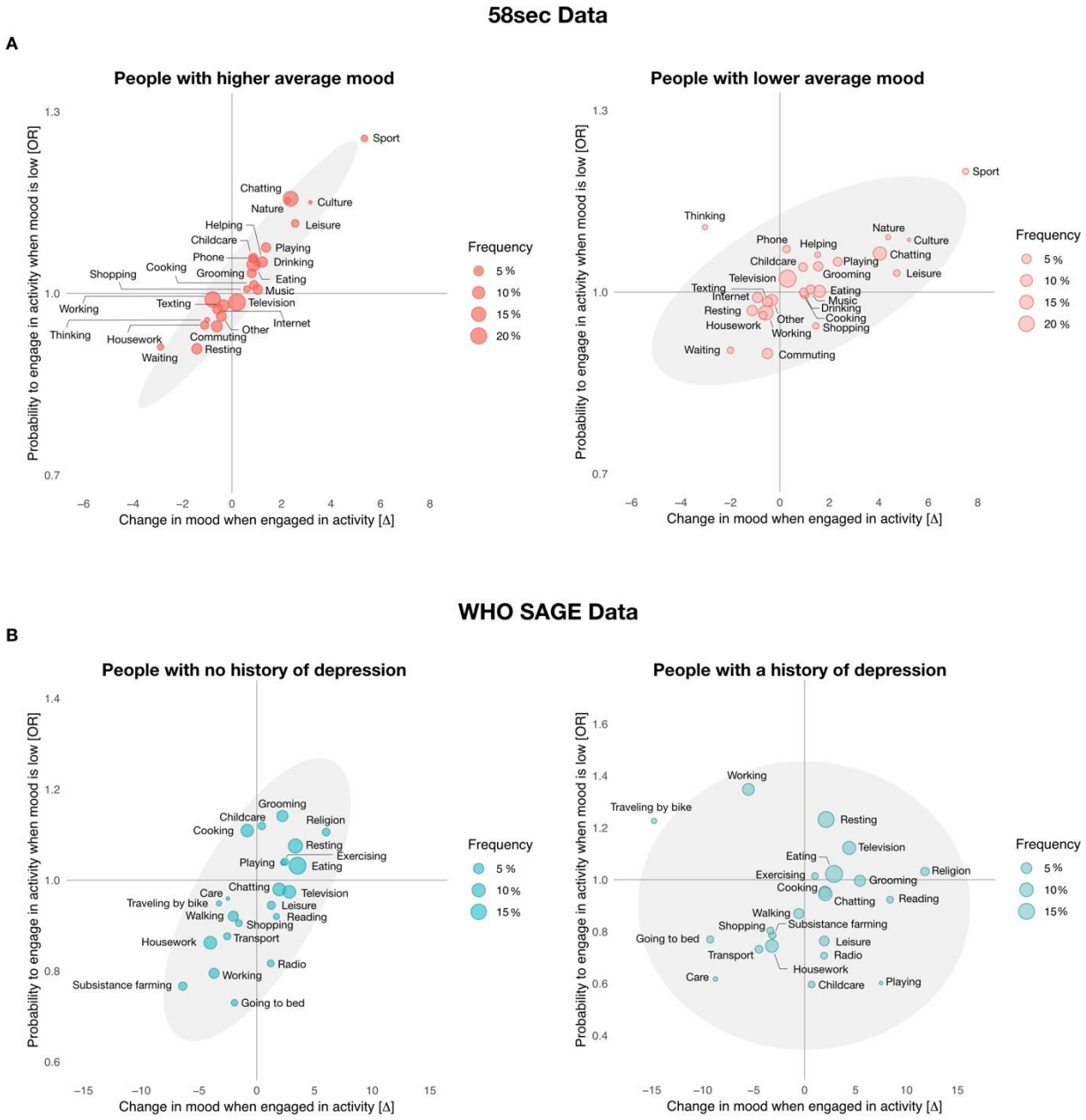
Subset A and B indicate the results obtained after randomly splitting the datasets in two mutually exclusive subsets. Parametric test refers to the use of a z-test performed on the Fisher z-transformed correlation coefficients. Random intercept refers to the use of multilevel regression models with a random intercept. Adjusting for country of origin (in the WHO SAGE dataset) refers to the addition of the country of origin as a covariate in the analysis. Error bars are 95% C.I. *$P < .05$, **$P < .01$, ***$P < .005$

eFigure 5. Group Difference in Mood Homeostasis Between High and Low Average Mood in the 58sec Data Set After Excluding Thinking Which Was Found to be Partially Driving the Results

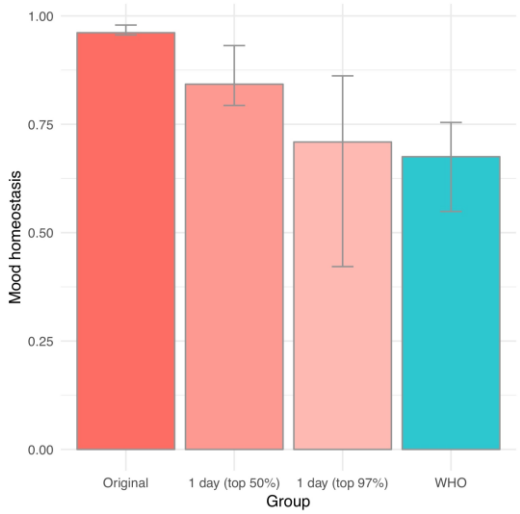


As anticipated, the difference in mood homeostasis is smaller. However it remains statistically significant ($P < .001$).

eFigure 6. Same Figure as Figure 2B-C and Figure 3B-C but With All Activity Labels Included

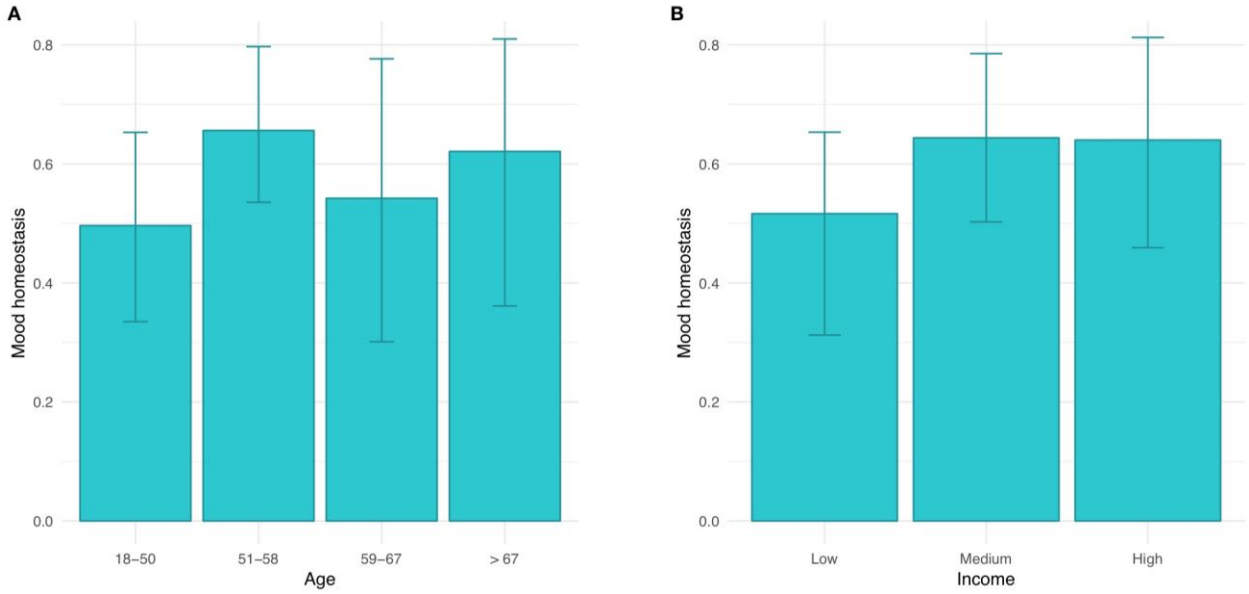


eFigure 7. Effect of Sampling Mood and Activities for Only 1 Day on the Value of Mood Homeostasis and Compounded Effect of Sampling Duration and Different Average Mood on Mood Homeostasis



The first and last bars correspond to the values in Fig. 7 A (high mood) and B (controls) respectively.

eFigure8. Mood Homeostasis Does Not Significantly Differ Between Age Group (A) and Between Income Levels (B)



eReferences

1. Taquet M, Quoidbach J, de Montjoye Y-A, Desseilles M, Gross JJ. Hedonism and the choice of everyday activities. *Proc Natl Acad Sci*. 2016;113(35):9769-9773. doi:10.1073/pnas.1519998113

2. Quoidbach J, Taquet M, Desseilles M, de Montjoye Y-A, Gross JJ. Happiness and Social Behavior. *Psychol Sci*. 2019;30(8):1111-1122. doi:10.1177/0956797619849666

3. DiCiccio TJ, Efron B. Bootstrap Confidence Intervals. *Stat Sci*. 1996;11(3):189-212.

4. Cohen P, West SG, Aiken LS. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Psychology Press; 2014.

5. Law AM, Kelton WD, Kelton WD. *Simulation Modeling and Analysis*. Vol 3. McGraw-Hill New York; 2000.

6. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Vol 2. World Health Organization; 1993.

7. Stone AA, Schneider S, Harter JK. Day-of-week mood patterns in the United States: On the existence of 'Blue Monday', 'Thank God it's Friday' and weekend effects. *J Posit Psychol*. 2012;7(4):306-314. doi:10.1080/17439760.2012.691980