Communications Psychology

https://doi.org/10.1038/s44271-026-00498-5

Article in Press

Spontaneous thought dynamics as a signature of positive and negative affectivity

Received: 3 January 2026

Accepted: 25 June 2026

Published online: 09 July 2026

Cite this article as: Lux, B.K., Lee, E., Han, J. *et al.* Spontaneous thought dynamics as a signature of positive and negative affectivity. *Commun Psychol* (2026). https://doi.org/10.1038/s44271-026-00498-5

Byeol Kim Lux, Eunjin Lee, Jihoon Han, Sung-Ha Lee, Suhwan Gim, Incheol Choi & Choong-Wan Woo

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

SPONTANEOUS THOUGHT DYNAMICS AS A SIGNATURE OF POSITIVE AND NEGATIVE AFFECTIVITY

Byeol Kim Lux[1,2,3*], Eunjin Lee[1,2,4*], Jihoon Han[1,2,4], Sung-Ha Lee[5], Suhwan Gim[1,2,4], Incheol Choi[6,7], and Choong-Wan Woo[1,2,4]

[1] Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, South Korea
[2] Department of Biomedical Engineering, Sungkyunkwan University, Suwon, South Korea
[3] Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA
[4] Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University, Suwon, South Korea
[5] Department of Psychology, Yonsei University, Seoul, South Korea
[6] Center for Happiness Studies, Seoul National University, Seoul, South Korea
[7] Department of Psychology, Seoul National University, Seoul, South Korea

[*]These authors contributed equally.

Running Head: SPONTANEOUS THOUGHT DYNAMIC SIGNATURES OF AFFECTIVITY

Please address correspondence to:

Choong-Wan Woo
Department of Biomedical Engineering
Center for Neuroscience Imaging Research
Sungkyunkwan University
Suwon 16419, Republic of Korea

Byeol Kim Lux
Department of Psychological and Brain Sciences
Dartmouth College
Hanover, NH, USA

ARTICLE IN PRESS

ARTICLE IN PRESS

Email: waniwoo@g.skku.edu

Telephone: +82 (31) 299-4363

Email: byeol.kim.gr@dartmouth.edu

Telephone: +1 603-349-0271

ABSTRACT

Spontaneous thought unfolds as a dynamic stream that reflects continuously changing internal states and is closely tied to individual differences in mental health. However, quantifying the temporal structure of spontaneous thought remains challenging due to limited behavioral paradigms and analytic tools that preserve its continuous, multidimensional nature. Here, we introduced Density Map-Based Predictive Modeling, a framework that characterizes spontaneous thought dynamics elicited by the Free Association Semantic Task within valence–self-relevance–time space. Density Map-Based Predictive Modeling summarizes each individual's thought stream using rating and vector density maps that capture both where thoughts concentrate and how they transition over time. Across multiple independent datasets (total N = 392), Density Map-Based Predictive Modeling yields robust, generalizable predictions of positive affectivity and negative affectivity using principal component regression. Feature analysis indicated that affectivity is reflected not only in the content of thoughts, but also in transition dynamics, with self-relevance as a key organizing dimension. Furthermore, higher predicted positive affectivity was associated with lower salivary C-reactive protein, suggesting a link between thought dynamics and inflammation. Together, these findings establish a scalable, interpretable approach for quantifying spontaneous thought dynamics and relating them to transdiagnostic affective traits and potential physiological correlates.

Keywords: spontaneous thought dynamics, positive and negative affectivity, inflammatory markers, free association, mind wandering

ARTICLE IN PRESS

ARTICLE IN PRESS

INTRODUCTION

The human mind is rarely still. Even in quiet moments or while engaged in ongoing tasks, our thoughts often drift, a phenomenon studied as mind-wandering, task-unrelated thought, self-generated thought, and spontaneous thought[1–3]. This stream of inner experience is not merely cognitive noise or a lapse in attention. Rather, it provides a window into personal goals, creativity, executive function, and overall mental health[4–8]. Spontaneous thought can be adaptive or maladaptive depending on its content and context[9–11]. On the one hand, mind-wandering can support memory consolidation, creative ideation, and planning for the future[4,12,13]. On the other hand, certain thought patterns are characteristic of multiple mental disorders and are associated with poorer well-being[2,13,14]. For example, repetitive negative thoughts, such as fixating on humiliating memories or anxious triggers, are characterized in obsessive-compulsive disorder, depressive disorder, and borderline personality disorder[2,6,15–17]. Despite the importance of these thought dynamics for mental health, their spontaneous and elusive nature makes them difficult to study empirically. The lack of robust methods for measuring the flow of thought has limited our ability to understand the links between thought dynamics, stable individual differences, and long-term outcomes. To address this gap, we developed a Density Map-based Predictive Modeling (DMPM)–to sequences of spontaneous thoughts elicited by the free association semantic task (FAST). The first intended contribution of this work is methodological, providing a scalable framework to transform elusive thought streams into mathematical trajectories. Secondly, it also offers a theoretical contribution by shifting the focus from static content to a dynamical systems perspective. With these, we developed generalizable DMPM models for positive affectivity (PA) and negative affectivity (NA) and examined their association with inflammatory markers.

To measure spontaneous thought in a form that can be analyzed quantitatively, we used FAST, a thought-sampling method that captures both the contents and temporal evolution of spontaneous thought[18,19]. In FAST, participants generate a sequence of concepts that come to mind within a time limit. Responses are often initially semantically close to a given seed word, but as participants become more comfortable with the task or when they reach some thought intersection, the stream of thoughts can shift toward personally meaningful themes. In some cases, participants repeatedly return to the same concepts, suggesting the presence of perseverative or “sticky” thoughts that are difficult to disengage from[10,18]. The task also

ARTICLE IN PRESS

ARTICLE IN PRESS

incorporates a self-evaluation component in which participants rate each self-generated concept on multiple affective dimensions with reference to its specific associative context. Accordingly, even repeated concepts can carry different contextual meanings across occurrences, differentiating FAST from a verbal fluency task.

Building on prior FAST work, a remaining methodological challenge is to fully leverage FAST's continuous ratings while representing thought trajectories in a compact, psychologically interpretable multidimensional space. In our prior work, we assessed each self-generated concept on five content dimensions–valence, self-relevance, time, vividness, and safety–and applied Markov chain analysis to characterize transitions between discretized states (e.g., negative to positive)[19]. These transition features successfully predicted individual differences in negative affectivity across independent datasets. However, because the ratings are continuous, binning them into a small number of states for modeling simplicity inevitably discards information and limits the resolution at which thought dynamics can be quantified. Here, we address this challenge by introducing Density Map-based Predictive Modeling (DMPM). Guided by our previous finding that the five content dimensions can be reduced to three core dimensions (valence, self-relevance, and time)[19], DMPM represents each participant's stream of thought as a trajectory through this three-dimensional space. It constructs a rating density map and a vector density map that capture, respectively, where thoughts tend to concentrate and the directions in which they tend to move within this space. This continuous and interpretable representation yields fine-grained features of spontaneous thought dynamics that can be related directly to individual differences in behavioral or physiological outcomes.

Leveraging DMPM's characterization of where thoughts cluster and how they move through a continuous valence–self-relevance–time (VST) space, we tested whether these thought dynamic features predict stable individual differences in affect and link to downstream physiological correlates. First, we focused on PA and NA, two core transdiagnostic dimensions of emotional functioning[20]. NA reflects a propensity toward subjective distress and is a robust predictor of psychopathology across diagnostic categories[21,22]. It is a core feature of major depressive and anxiety disorders and is closely related to neuroticism[23,24]. In contrast, PA reflects pleasurable engagement and approach-related motivation, is a key indicator of well-being, and a buffer against adversity[25–27]. Low PA is linked to anhedonia, the loss of pleasure that is a

ARTICLE IN PRESS

hallmark symptom of depression[24,28]. Crucially, PA and NA are not opposite ends of a single continuum but partially independent dimensions[20,29,30]. Because trait-level PA and NA shape momentary appraisals and emotional experiences[31], they should also influence the valence, self-relevance, and temporal orientation of spontaneous thought as captured by FAST and summarized by DMPM. As transdiagnostic affective traits, PA/NA are tied not only to psychiatric symptoms but also to broader health trajectories. We therefore examined inflammatory biomarkers that provide a mechanistic link between psychological functioning and systemic health.

We extended our analyses beyond self-reported affect to inflammatory biomarkers that index stress-related immune activation and have been repeatedly implicated in depression and anxiety. A substantial literature supports a bidirectional relationship between inflammation and mood disorders[32–34]. Meta-analysis evidence indicates that individuals with major depressive disorder (MDD) show elevated circulating inflammatory markers, including pro-inflammatory cytokines and indices of cellular immune activation[35–40]. Pro-inflammatory cytokines influence brain function and promote sickness behaviors, such as lethargy, social withdrawal, and anhedonia, which overlap with core symptoms of depression and anxiety[41–44]. Converging experimental evidence further supports this link. Cytokine administration or inflammation induction increases depression- and anxiety-like behaviors in animal models[45–47,47,48], whereas anti-inflammation interventions can improve depressive symptoms in some patient groups[49,50].

Among commonly studied mediators, Interleukin-1β (IL-1β) and interleukin-6 (IL-6) increase following laboratory psychosocial stressors[51,52] and are also elevated under chronic stress exposure (e.g., caregiving burden, early-life adversity)[53,54]. C-reactive protein (CRP), produced by the liver in response to inflammatory signaling (including IL-6), is widely used as a marker of systemic low-grade inflammation and is also a well-established risk factor for cardiovascular disease[55,56]. Consistent with these links, elevated IL-1β, IL-6, and CRP have been reported in MDD[34,35,37,38,57], suggesting that affective disturbance is associated with longer-term physical health risk. Accordingly, we tested whether DMPM-based predictions of PA and NA are associated with individual differences in salivary levels of IL-1β, IL-6, and CRP.

ARTICLE IN PRESS

ARTICLE IN PRESS

In the present study, we aim to test whether the dynamic features of spontaneous thought can predict stable individual differences in affective traits and relate to systemic physiological markers. Using the DMPM framework, we investigate how the temporal evaluation of thoughts within a multidimensional space—comprising valence, self-relevance, and time—relates to positive affectivity and negative affectivity. We hypothesize that individuals with higher trait negative affectivity will show thought trajectories characterized by frequent transitions into or persistent within negative and self-relevant states, whereas those with higher positive affectivity will show dynamics associated with pleasurable engagement. Furthermore, we explore whether these computationally derived patterns of thought dynamics are associated with physiological correlates, specifically salivary inflammatory markers such as IL-1β, IL-6, and CRP. By validating these models across multiple independent datasets and controlling for relevant demographic and lifestyle factors, we seek to establish a scalable and interpretable approach for quantifying the link between the stream of thought and transdiagnostic markers of psychological and physical health.

ARTICLE IN PRESS

METHODS

Participants

In Study 1, 117 participants (age = 22.6 ± 2.6 years [mean ± SD], 56 females and 61 males) completed the web version of the FAST (FAST-web) in a laboratory setting. A subset of these participants (N = 49, age = 23.0 ± 3.1 years, 24 females and 25 males) returned for a retest session approximately 8 weeks later (test-retest interval = 54.2 ± 8.5 days). Study 2 involved 213 participants (age = 22.9 ± 3.1 years, 106 females and 107 males) who completed the experiment remotely using their personal computers. In Study 3, 62 participants (age = 23.0 ± 2.5 years, 30 females and 32 males) completed the FAST within a Magnetic Resonance Imaging (MRI) scanner. The sex information was provided by participants, and 100% of the total sample identified as either female or male. Participants were all Korean, recruited from the Suwon area in South Korea. The experiments for Study 1 and Study 3 were conducted at the Center for Neuroscience Imaging Research, Sungkyunkwan University, Suwon, South Korea, while Study 2 was conducted remotely. The institutional review board of Sungkyunkwan University approved all studies. All participants provided written informed consent and received

ARTICLE IN PRESS

compensation for their participation. This study was not preregistered. Studies 1 and 3 were used in the previous study with a different research goal[19,58,59]. In this study, the first session from Study 1 served as the training dataset, while the retest session of Study 1, Study 2, and Study 3 data served as testing datasets.

Self-report questionnaires

Participants in all studies completed a set of self-report questionnaires to evaluate the individual differences in trait affectivity and mental health variables. In Studies 1 and 2, the questionnaire battery included the 20-item Center for Epidemiologic Studies Depression (CES-D)[60], the 20-item Positive and Negative Affect Schedule (PANAS) with subscales for positive and negative affect[61], the 20-item trait version of the State-Trait Anxiety Inventory (STAI-T)[62], 2-items of the Suicidal Ideation Questionnaire (SIQ, “I wished I were dead”, “I thought that life was not worth living”)[63], 9 items from the ‘depressive rumination’ sub-scale of the Rumination Response Scale (RRS)[64], the 3-item Loneliness Scale (LS)[65], the 5-item Satisfaction With Life Scale (SWLS)[66], the 18-item Psychological Well-Being Scale (PWB) encompassing six subscales—autonomy, environmental mastery, personal growth, positive relations with others, purpose in life, and self-acceptance[67]. Additional items related to physical states, such as alcohol consumption, smoking habits, height, and weight, were also included. Participants responded to the questionnaires prior to the web-based FAST experiment. For Study 3, the questionnaire set included the 20-item PANAS, the 20-item CES-D, the 22-item RRS (with brooding and depressive rumination subscales), the 20-item STAI-T, and the 30-item Mood and Anxiety Symptom Questionnaire-D30 (MASQ-D30), which includes subscales for general distress, anhedonic depression, and anxiety arousal[68]. Participants completed these questionnaires before the fMRI experiment. Korean versions of these questionnaires[69–73] were used in all studies.

Web-based Free Association Semantic Task (FAST) experiment (Study 1, 2)

The web-based FAST allowed Study 1 to be conducted in a laboratory setting using a computer, while Study 2 was completed remotely. The web-based FAST consisted of two main tasks: concept generation and subsequent surveys for the self-generated concepts (**Fig. 1A**). During the concept generation task, we asked participants to type a word or phrase that came to mind in response to the previous concept every 10 seconds in Study 1, and every 7 seconds in

ARTICLE IN PRESS

ARTICLE IN PRESS

Study 2, starting from a given seed word. Although we instructed participants to generate responses within the response window and a timer was visually provided on the screen, the trial did not proceed until a response was provided, ensuring a complete sequence of 40 concepts per run. The group-level mean response time was 5.61 ± 1.67 seconds for Study 1 and 5.22 ± 1.80 seconds for Study 2, with the majority of responses (96.8% and 81.3%, respectively) submitted within the given response window. This structured timing was specifically intended to ensure more spontaneous concept generation while minimizing deliberate, goal-directed retrieval or self-monitoring. For each seed word, participants generated 40 consecutive concepts, with responses inputted via typing. We instructed participants to generate any words or phrases, emphasizing that their responses did not need to be related to the seed word or previous concept. They were also told they could repeat their responses at any time. By asking participants to generate brief concepts rather than extended narratives, as in think-aloud paradigms, our task is designed to capture a relatively unconstrained stream of thought while reducing the cognitive demands associated with social self-presentation.

Following the concept generation task, participants rated these 40 concepts on three content dimensions[15,19]: emotional valence (assessing the positive or negative feelings evoked by the concept), self-relevance (measuring the concept's personal relevance), and time (identifying the most relevant time point for the concept, from distant past to distant future). To provide context for associations, the target concept for the rating was displayed with its previous concepts together. Participants clicked on the bar to report their ratings. Valence and time scales ranged from −1 (negative or distant past) to 1 (positive or distant future), with 0 denoting neutral or present. The self-relevance scaled from 0 (not at all self-relevant) to 1 (highly self-relevant). The ratings were recorded up to three decimal places.

In each session, this process consisted of four runs, each starting with a different pre-selected seed word. Participants generated 40 concepts per run, resulting in a total of 160 concepts and dimension ratings per session. The seed words varied between sessions: "family," "tear," "mirror," and "abuse" were used in the first session of Study 1 and Study 2, while a different set of four seed words—"love," "fantasy," "heart," and "pain"—were used only in the re-test session of Study 1. The order of seed words was randomized across participants.

ARTICLE IN PRESS

We instructed participants to rate the concepts while considering the context of their free association process. For example, starting with the seed word "family," a participant might generate a sequence such as "mom" → "love" → "happiness". If they repeated the same words intermittently, they were asked to rate each concept independently based on the specific context in which it came to mind. This approach aimed to distinguish concepts tied to their thought flow from those merely sharing phonetic similarities. Additionally, participants were instructed to rate concepts as 'present' (with a score of 0 in the time dimension) if: (1) the concept was not linked to their specific memories or personal narratives (e.g., transitions based on semantic similarity, such as "tear" → "water"), (2) it emerged spontaneously without a temporal anchor (e.g., immediate sensory experiences like "headache"), or (3) it was not associated with any particular time frame (e.g., timeless abstract concepts like "math" or "sky"). In Study 3, which used a shorter time limit and verbal report, the previous concept was filled in for the trial if participants were unable to generate a word within the allotted time. In the web-based FAST (Studies 1 and 2), participants were given more time to generate each concept, and the trial did not proceed until a concept was provided. We found that the majority of responses were submitted within the given time frame.

fMRI-based Free Association Semantic Task (FAST) experiment (Study 3)

Previous work[19] implemented the FAST alongside functional magnetic resonance imaging (fMRI) to investigate the neural systems associated with self-generated thoughts. The FAST-fMRI experiment also included concept generation and a post-scan survey. Unlike the web experiment, participants verbally reported a word or phrase every 2.5 seconds while inside the MRI scanner, with responses captured via an MR-compatible microphone. The seed words and the number of self-generated concepts from each word were identical to those used in Study 1 and Study 2. Following the fMRI scan, participants completed a post-scan survey in a separate behavioral experimental room, rating their 160 self-generated concepts on five content dimensions. The five content dimensions are valence, self-relevance, time, vividness (the degree of vivid imagery induced by the concept), and safety-threat (the strength of feeling of threat or safety evoked by the concept).

Development of Density Maps in Valence–Self-relevance–Time Multidimensional Space

ARTICLE IN PRESS

To examine the dynamic features of individuals' self-generated thought processes, we created two types of density maps—rating density maps and vector density maps—derived from survey data on 160 self-generated concepts (**Fig. 2**). The three content dimensions form a three-dimensional space called the Valence–Self-relevance–Time (VST) space. Valence ($\boldsymbol{v}$) and time ($\boldsymbol{t}$) ranged from −1 to 1, while the self-relevance ($\boldsymbol{s}$) ranged from 0 to 1.

First, this rating density map provides a cumulative frequency representation of an individual's thoughts across three content dimensions. To generate the rating density map, each concept from a participant's survey data was plotted as a point within this space, resulting in 160 plotted points per participant (**Fig. 2A**). We calculated the number of concepts located at each coordinate. For example, if a participant reports multiple concepts that are very negative ($\boldsymbol{v}$ = −1), highly self-relevant ($\boldsymbol{s}$ = 1), and related to distant memories ($\boldsymbol{t}$ = −1), their rating density map would show increased density in the region corresponding to negative valence, high self-relevance, and distant past orientation, at (−1, 1, −1).

Next, we constructed the vector density map by calculating the density of difference vectors between consecutive concept coordinates. For example, if the n-th concept is at coordinate $[\boldsymbol{v}, \boldsymbol{s}, \boldsymbol{t}]$ = [0, 0, 1] and the n+1-th concept is at coordinate [1, 0.5, 0], the difference vector from the n-th to the n+1-th has a tail at [0, 0, 1] and a head at [1, 0.5, 0]. To analyze the direction their thoughts were heading during free association, we repositioned the difference vectors to start from a common origin. In this example, we reposition the vector so that its tail is at $[\Delta\boldsymbol{v}, \Delta\boldsymbol{s}, \Delta\boldsymbol{t}] =$ [0, 0, 0] and its head at [1, −0.5, −1], retaining its magnitude and direction. Similar to the rating density map, we computed the cumulative frequency of the vectors' heads, increasing the density by 1 at the coordinate [1, −0.5, −1]. The resulting vector density map visualizes the direction of thought flow over time within the VST space, accounting for the cumulative frequency of 156 vectors (39 × 4 runs) from each individual. If a participant generates more negative concepts ($\Delta\boldsymbol{v} = -0.1$), which is more self-relevant ($\Delta\boldsymbol{s} = 0.1$) and shows no change in temporal orientation ($\Delta\boldsymbol{t} = 0$), compared to the previous concepts during the free association, the vector density map would reflect elevated values at coordinates such as $[\Delta\boldsymbol{v}, \Delta\boldsymbol{s}, \Delta\boldsymbol{t}] = [-0.1, 0.1, 0]$.

ARTICLE IN PRESS

The rating density map was generated as a 50 × 50 × 50 three-dimensional matrix, and the vector density map was created as a 25 × 25 × 25 matrix. We applied a convolution to both maps using a 3D spherical kernel (radius = 5) and smoothed the convolved data using a 3D Gaussian kernel (standard deviation = 2). For predictive modeling, these dynamic density maps were vectorized, yielding 140,625 features per participant (125,000 features from the rating density map and 15,625 features from the vector density map).

Density Map-Based Predictive Modeling (DMPM)

We developed Density Map-Based Predictive Modeling (DMPM), using the rating and vector density maps to predict general positive affectivity (PA) and negative affectivity (NA). PA and NA scores were derived from factor analysis of self-report questionnaire data that measured multiple aspects of positive and negative affectivity (**Supplementary Table 1**). The included questionnaires were the Center for Epidemiologic Studies Depression (CES-D), Positive and Negative Affect Schedule (PANAS), State-Trait Anxiety Inventory-Trait version (STAI-T), Suicidal Ideation Questionnaire (SIQ), Rumination Response Scale (RRS), Loneliness Scale (LS), Psychological Well-Being scale (PWB), and Satisfaction With Life Scale (SWLS).

For the factor analysis, we used *z*-scored subscale scores as inputs to consider the different scaling across questionnaires. We applied a target rotation method with a two-factor model to align with our predetermined understanding of positive and negative affectivity. The positive affectivity factor captured constructs such as joy, happiness, subjective well-being, and satisfaction. Conversely, the negative affectivity factor encompassed measures of distress, negative moods, anxiety, and depression.

The PA model and NA model were trained separately on the training dataset, which is the first session data of Study 1 (N = 117), using Principal Component Regression (PCR). We verified that the data met the statistical assumptions for Principal Component Regression, including the linearity of the relationship between predictors and outcomes, and the normality and homoscedasticity of the residuals. The use of Principal Component Regression also addressed potential issues of multicollinearity inherent in high-dimensional density map data. The optimal number of principal components was determined based on the leave-one-subject-out cross-validation (LOSO-CV) performance in the training dataset to minimize performance

ARTICLE IN PRESS

estimation bias. To estimate the within-subject reliability of the maps across the two sessions of Study 1, we conducted a permutation test with 10,000 iterations using data from both sessions. In each iteration, subject identities were randomly shuffled within each session, and the resulting within-subject correlations between maps across sessions were calculated to generate a null distribution.

These models were tested on three datasets. The first was the re-test session data from a subset of the Study 1 participants (N = 49), who responded to a different set of seed words. The second and third test datasets were Studies 2 and 3 (N = 213 and 62, respectively). For Studies 1 and 2, we used consistent normalization and factor analysis parameters as those in the training dataset. However, Study 3 only included negative affect-related questionnaires, necessitating separate parameters for extracting negative affectivity. As a result, the outcome variables in Study 3 were on different scales, and the PA model was not tested in Study 3. To address these scaling differences, we used correlation-based performance estimates, which are scale-invariant and more robust to scale variations than sum-of-squared-based measures[74]. PA and NA predictions were computed as the dot product of the model's predictive weights with the corresponding density maps. Model performance was assessed using robust correlation, which mitigates the impact of outliers for more reliable estimates. The 'partialcor' function from Canlab toolbox (https://github.com/canlab/CanlabCore), which utilizes the 'robustfit' function in Matlab Statistics and Machine Learning Toolbox, was used for the analysis.

To identify important features of the predictive models, we conducted bootstrap tests with 10,000 iterations on our final models. In each iteration, participants were randomly resampled with replacement, and a PCR model was trained on the resampled data. The sampling distribution of the bootstrapped predictive weights was then used to identify which features reliably contributed to model predictions. A false discovery rate (FDR) threshold of $q < 0.05$ was applied to the density maps to determine the significance of these features. To compare the relative influence of the rating and vector density maps in the PA and NA models, we quantified their relative contributions by calculating the proportions of significant predictive features out of the total number of features in each map.

Inflammatory markers

ARTICLE IN PRESS

In Study 1, we collected saliva samples from 68 participants to measure inflammatory markers, including interleukin-1β (IL-1β), interleukin-6 (IL-6), and C-reactive protein (CRP). To minimize confounds in saliva samples, we required participants to abstain from alcohol, smoking, and caffeine for 12 hours, and from eating for one hour before the collection. Following the FAST-web experiment, participants rinsed their mouths with water and collected saliva using the Saliva Collection Aid (SCA, Salimetrics, USA), adhering to the recommended procedure by Salimetrics, which includes tilting the heads forward and drooling to minimize bubble formation. The collected saliva samples were stored at −20°C and subsequently sent to Seoul Clinical Laboratories for analysis. To address extreme values, we applied winsorization to the inflammatory marker response at 3 standard deviations and normalized the data using a square root transformation. The square root transformation and winsorization were applied to ensure the data satisfied the normality assumption required for parametric partial correlation. IL-6 levels in 27 participants were undetectable and thus excluded from further analysis. To investigate the relationship between inflammatory markers and affectivities, we calculated partial correlations between each marker and DMPM-predicted PA/NA or PA/NA from factor analysis while controlling for health-related covariates known to influence inflammation. Covariates include body mass index (BMI), age, gender, daily tobacco use, and alcohol consumption (frequency and quantity per occasion). Partial correlations were computed on the residualized variables, quantifying the association between affectivity and inflammatory markers after removing variance attributable to these covariates.

RESULTS

Probing thought processing using Free Association Semantic Task

The Free Association Semantic Task (FAST) illustrates the distinctive themes and dynamic nature of an individual's spontaneous thought, as seen by a representative participant's free association data (**Fig. 1**). The participant's concept chain was visualized as a directed graph (**Fig. 1B**), where nodes represented the self-generated concepts and directed edges indicated associative links between them. This graph highlighted several central 'hub' concepts with a high degree of centrality, such as "love," "strength," "competence," and "study," each of which was connected to a wide array of related concepts. These hubs appeared to play a critical role in

ARTICLE IN PRESS

constructing the participant's semantic network (for a complete overview of the participant's responses, refer to **Supplementary Figure 1**). Notably, despite the open-ended nature of the task, hub concepts frequently recurred across participants' responses, displaying varying levels of valence (**Fig. 1C**). This suggests that the task effectively engaged participants in spontaneous thought related to personal concerns and ongoing goals during the free association process. In Study 1 (N = 117), "money" was the most frequently occurring concept, followed by "mom," "love," "happiness," and "friend." Across all three studies, core themes such as "mom," "friend," "love," and "sadness" consistently emerged as central conceptual hubs. Despite variations in seed words, experimental settings (remote, onsite, or in the MRI scanner), and response formats (typing vs. verbal), the response patterns and central conceptual hubs remained consistent across datasets (**Supplementary Figures. 2-3**).

Density Map-Based Predictive Modeling (DMPM) for affectivity traits in FAST responses

To develop predictive models for Positive Affectivity (PA) and Negative Affectivity (NA) based on the dynamic features of FAST responses, we introduced the Density Map-Based Predictive Modeling (DMPM) method, which uses the rating and vector density maps as input features for machine learning models (**Fig. 2**). The rating density map and vector density map reveal the distribution of thoughts and transitions between them across the content dimensions during free association. We trained the models on maps obtained from Study 1 (N = 117), with the selected optimal number of principal components using LOSO-CV in the training dataset to minimize performance estimation bias (**Supplementary Figure. 4**). The trained models showed high within-subject stability across two sessions (N = 49), yielding high correlation coefficients (r = 0.693, 95% CI [0.585, 0.777], p < 0.0001 for the rating density map and r = 0.986, 95% CI [0.975, 0.992], p < 0.0001 for the vector density map, permutation tests; **Supplementary Figure. 5**).

Both DMPM models for PA and NA demonstrated strong predictive performances across all datasets (**Fig. 3**). The PA model showed robust correlations between actual and predicted factor scores (**Fig. 3A**): r = 0.597, 95% CI [0.466, 0.702], p < 0.0001 in the training dataset from Study 1 (N = 117; LOSO-CV), r = 0.416, 95% CI [0.153, 0.624], p = 0.0069 in the retest dataset (N = 49, LOSO-CV), and r = 0.596, 95% CI [0.502, 0.676], p < 0.0001 in Study 2

ARTICLE IN PRESS

(N = 213; two-tailed, one-sample t-test). The PA model was not applied to Study 3 due to a limited range of PA questionnaires. For the NA predictive model (**Fig. 3B**), the correlations were r = 0.582, 95% CI [0.448, 0.691], p < 0.0001 in the training dataset (N = 117, LOSO-CV), r = 0.551, 95% CI [0.319, 0.721], p = 0.0002 in the retest dataset (N = 49, LOSO-CV), r = 0.489, 95% CI [0.380, 0.585], p < 0.0001 in Study 2 (N = 213), r = 0.578, 95% CI [0.384, 0.723], p < 0.0001 in Study 3 (N = 62). These results validate the robustness and reliability of our DMPM models in predicting trait affectivity using the density map features across multiple task parameters and experimental conditions.

Exploring key predictive features of DMPM for trait affectivity

Building on the development and validation of our DMPM models, we examined which dynamic aspects of spontaneous thought, as captured in the density maps, predict individual differences in affectivity (**Fig. 4**). Although PA and NA are negatively correlated across datasets (Study 1: r = −0.562, 95% CI [−0.675, −0.424], p < 0.0001; Study 2: r = −0.606, 95% CI [−0.685, −0.513], p < 0.0001; Study 3: r = −0.279, 95% CI [−0.494, −0.031], p = 0.0283), consistent with their characterization as partially independent rather opposite dimensions of affect[20,29,30], a substantial portion of variance in each trait remained unique. We therefore examined whether their predictive features were likewise distinct. Using bootstrap tests with 10,000 iterations and applying a false discovery rate (FDR) threshold of q < 0.05, we identified key predictive features for PA and NA. Analysis of the rating density map (**Figs. 4B-D**) revealed that concepts with positive valence, high self-relevance, and a present-oriented temporal focus were associated with higher PA levels and lower NA. Also, greater densities of neutral, present-oriented, and self-irrelevant thoughts were associated with higher PA and lower NA, although these effects did not survive FDR correction. Overall, the findings suggest that the valence of present-oriented, self-relevant concepts plays a critical role in shaping both forms of affectivity.

Vector density map analysis (**Figs. 4E-G**) showed that thought transitions in which valence and self-relevance shifted in the same direction—becoming more positive and self-relevant or more negative and less self-relevant—contributed to higher PA and lower NA. Conversely, transitions reflecting decreasing valence and self-relevance were linked to higher NA and lower PA. These patterns were most pronounced with minimal changes in the time

ARTICLE IN PRESS

change dimension and were distributed along the time dimension, especially in the PA model (**Figs. 4E-F**). In the PA model, positive associations between valence and self-relevance spanned a broader range of temporal transitions (red areas in **Fig. 4E**) than the negative association between these variables (blue areas in **Fig. 4E**). Additionally, comparisons of feature importance along the time dimension showed that backward (retrospective) transitions were more predictive than forward (prospective) transitions for both PA and NA models (**Supplementary Figure. 6**), highlighting the significance of past-oriented thought dynamics in affectivity.

Comparing the two density maps revealed two major insights. First, the vector density map provided more reliable predictive features than the rating density map, as shown by a higher proportion of significant coordinates in the bootstrap analysis (FDR $q < 0.05$; difference in proportions = 6.46%, 95% CI [6.41, 6.51], $z = 253.136$, $p < 0.0001$ for PA; difference = 0.13, 95% CI [0.10, 0.16], $z = 8.711$, $p < 0.0001$ for NA; two-tailed, binomial z-test; **Supplementary Figure. 7**). This suggests that the dynamics of thought transition carry greater predictive value for trait affectivity than thoughts frequency alone. Second, self-relevance emerged as a consistently critical predictor across both maps. When we divided each map into high and low self-relevance areas (indicated by horizontal dashed gray lines in **Figs. 4B-G**), the high self-relevance areas consistently showed stronger predictive power across all models. In the NA models in particular, all significant features were concentrated in the high self-relevance area (**Figs. 4C** and **4F**), underscoring the strong impact of self-related thoughts on NA.

Correlation between spontaneous thought patterns and inflammatory markers

To examine whether inflammatory responses relate to affectivity as predicted from thought dynamics, we conducted partial correlations between three salivary inflammatory markers and DMPM-predicted affectivity scores, while controlling for BMI, age, gender, tobacco usage, and alcohol consumption. IL-1β and IL-6 were not significantly associated with DMPM-predicted PA and NA (IL-1β, $N = 68$: PA, $r = 0.084$, 95% CI [−0.158, 0.316], $p = 0.498$; NA, $r = -0.100$, 95% CI [−0.331, 0.142], $p = 0.417$; IL-6, $N = 47$: PA, $r = -0.221$, 95% CI [−0.478, 0.071], $p = 0.164$; NA, $r = 0.256$, 95% CI [−0.034, 0.506], $p = 0.106$; **Supplementary Figure. 8**). In contrast, the positive association between DMPM-predicted NA and CRP did not reach significance ($N = 68$, $r = 0.227$, 95% CI [-0.012, 0.442], $p = 0.063$) whereas CRP showed

ARTICLE IN PRESS

a significant negative correlation with DMPM-predicted PA ($r = -0.270$, 95% CI [−0.478, 0.034], $p = 0.026$, **Fig. 5**). Notably, the association between CRP and the PA scores derived directly from self-report questionnaire responses was weaker and did not reach significance ($r = -0.226$, 95% CI [−0.441, 0.013], $p = 0.064$). This relationship suggests that the thought-dynamic features captured by DMPM may reflect aspects of positive affective functioning relevant to systemic inflammation that are less accessible via direct introspective report.

DISCUSSION

DMPM combined with FAST was able to predict positive and negative affectivity from spontaneous thought dynamics represented in the continuous Valence–Self-relevance–Time (VST) space. By summarizing each participant's thought stream with rating and vector density maps, DMPM captures fine-grained dynamics in which thoughts tend to cluster and how they move across these core experiential dimensions, yielding an interpretable set of features associated with affective traits. Whereas valence, self-relevance, and temporal orientation have typically been studied in isolation in studies of autobiographical memory and self-referential cognition, the VST framework allows an integrated examination of how these dimensions interact as thought unfolds over time. Across three datasets that varied in experimental parameters (e.g., seed words, response time limits, experimental environments, and task modalities), the resulting models showed robust performance and generalizability. Finally, the observed association between DMPM-derived affect predictions and salivary inflammatory markers—most notably CRP—suggests that stable patterns in subjective thought dynamics may relate not only to emotional functioning but also to downstream physiological correlates with potential implications for long-term health.

At the level of thought content, the rating density map indicated that greater occupancy in negative, self-relevant, and past-oriented regions of the VST space predicted lower PA and higher NA. This pattern is consistent with previous work linking affective vulnerability to sustained engagement with self-focused negative material, including rumination and brooding, which are well-established risk processes for depression and related disorders[75,6,76,77]. Conversely, higher PA and lower NA were associated with a greater density of positive self-relevant concepts, aligning with evidence that reward-related, positively valenced self-referential

ARTICLE IN PRESS

thought supports well-being. Importantly, these results extend prior findings with FAST by showing that affective traits are reflected not only in average content ratings of self-generated thought but also in the spatial distribution of thoughts and their dynamics within a continuous multidimensional space[19].

Building on these content-level findings, the vector density maps revealed that thought transitions carry additional and more reliable information about trait affectivity than static frequency patterns alone. Specifically, transitions in which valence and self-relevance shifted together (i.e., becoming more positive and more self-relevant, or more negative and less self-relevant states) were associated with higher PA and lower NA, whereas transitions toward more negative and more self-relevant states were associated with higher NA and lower PA. This coupling between valence and self-relevance is consistent with a dynamic analogue of self-positivity bias, the tendency for self-related information to be evaluated more positively than other-related information in non-clinical contexts[78,79]. Furthermore, lower PA/higher NA was characterized not only by a tendency toward negative self-referential thoughts, but also by trajectories that shifted toward positive content that remained relatively self-irrelevant. In other words, their thought flow reflected a more negative "self" alongside more positive "others". This pattern aligns with cognitive accounts of depression in which negative self-schemas are preferentially encoded and elaborated, whereas positive information is less likely to be integrated into the self-concept[80–83]. Converging neuroimaging evidence likewise suggests that, in depression, positive concepts are less likely to engage self-referential systems[84–86] and reward-related circuitry[87,88]. By making these processes observable as temporal transitions in spontaneous thought, the vector density maps provide a mechanistically interpretable dynamic phenotype that would be difficult to recover from summary statistics (e.g., mean ratings) or questionnaire measures alone. This dynamical system approach aligns with recent network science models of psychotherapeutic change, which emphasize the importance of mental state transitions and cognitive fluidity as indicators of psychological health[89].

In line with the valence–self-relevance coupling observed in the dynamic features, self-relevance emerged as a central organizing dimension of affectivity. Significant NA features were concentrated almost exclusively in high self-relevance regions of both the rating and vector density maps, supporting the view that negative affectivity is tightly coupled to self-focused

ARTICLE IN PRESS

processing and that the personal meaning of spontaneous thought shapes its emotional impact[15,90,91]. This interpretation is consistent with neuroimaging evidence that self-related mental states are represented with greater distinctiveness and granularity in neural systems than self-irrelevant thoughts[92], and with prior FAST work highlighting self-relevance as a particularly informative dimension for characterizing affect-linked thought patterns[19].

While the selection of the VST dimensions was motivated by prior research showing that they constitute key components for predicting individual differences in affective traits, the specific findings regarding temporal dynamics in this study emerged in a data-driven manner. Importantly, these findings were made possible by the DMPM framework, which enables a fully continuous, multidimensional characterization of thought dynamics, and thus would have been difficult to detect using more conventional approaches. More specifically, the finding that higher PA is associated with sustaining positive, self-relevant thought across a broader temporal range (e.g., from past through future) could be consistent with the 'broaden-and-build' theory of positive emotions, which suggests that positive affect expands an individual's cognitive and experiential repertoire[26,31]. In contrast, NA-related features were more concentrated in regions of negative self-relevance with comparatively smaller shifts along the time axis, implying a narrower or more persistent temporal focus. This pattern is consistent with clinical accounts of depressive rumination and "sticky" thoughts, which are often characterized by a reduced ability to flexibly disengage from specific memories or to explore a broader range of thought topics[10,18,93].

Moreover, our analysis revealed that retrospective (backward) transitions, compared to prospective (forward) transitions, were more predictive of both PA and NA. The stronger influence of past-oriented thought content and transitions, relative to future-oriented ones, is consistent with accounts proposing that autobiographical memory is more idiosyncratic and affectively heterogeneous than future simulation[94], and supports the constructive episodic simulation hypothesis, which posits that future imagination relies on recombining past experiences[93,95]. Given that negatively valenced, past-oriented thought is associated with depressive symptoms and reduced well-being[15,96], these results suggest that affective traits may be expressed especially strongly in how individuals revisit and shift away from autobiographical memories. By representing these dimensions as a continuous trajectory, the current study extends

ARTICLE IN PRESS

prior FAST work by identifying the temporal flexibility of thought as a crucial dynamic marker of affective health.

Finally, the association between DMPM-derived affectivity predictions and CRP provides preliminary evidence that spontaneous thought dynamics may relate to physiological processes relevant for health. IL-1β and IL-6 did not show reliable associations in this sample, and the CRP effect was modest even after controlling for BMI, age, gender, tobacco use, and alcohol consumption. These results warrant caution in mechanistic interpretation. Nevertheless, CRP related more strongly to DMPM-predicted PA than to self-reported PA, raising the possibility that FAST-derived thought dynamics capture variance in affective functioning that is less accessible to introspective report yet relevant to inflammatory status.

Limitations

The lack of significant correlations for IL-1β and IL-6 (**Supplementary Figure. 8**) may reflect methodological considerations. Although these cytokines have been studied in depression, salivary IL-1β and IL-6 can be strongly influenced by local oral immune activity and periodontal inflammation[97,98], which may obscure associations with affectivity. In contrast, CRP, produced primarily by the liver, is less affected by local oral factors and more closely indexes systemic low-grade inflammation[55,56]. Moreover, IL-1β and IL-6 are often studied as relatively acute stress-responsive markers[51,52], while CRP is commonly interpreted as a more stable indicator of chronic inflammation and disease risk[99]. Because FAST was not designed as a stress-induction paradigm and our analyses targeted stable individual differences in a relatively young, healthy sample, CRP may be better suited than salivary cytokines for detecting trait-like physiological correlates of affect-linked thought dynamics.

Future studies should test the generalizability of the DMPM framework and its predictive models in more diverse populations, including samples drawn from different cultural contexts, community cohorts, and clinical settings. Thought dynamics may also differ qualitatively across disorders. For example, MDD may involve more perseverative and less variable thought trajectories, whereas bipolar-spectrum conditions may involve larger and more rapid shifts in affectively salient content. An important next step is to evaluate whether DMPM can identify disorder-relevant dynamical phenotypes and support differential prediction across

ARTICLE IN PRESS

diagnostic groups. This approach could provide a scalable way to detect certain thought patterns that indicate affective vulnerability, with potential utility for risk stratification and treatment monitoring. Also, the observed CRP association remains correlational and does not establish directionality. Larger biomarker samples, repeated biomarker assessments, and longitudinal or experimental designs will be required to determine whether thought-dynamic signatures prospectively predict inflammatory change, whether inflammatory activity shapes thought dynamics, or whether both reflect shared upstream factors.

CONCLUSION

Together, this work introduces a quantitative approach for linking everyday spontaneous thought dynamics to stable affective traits and the immune biology, with potential translational relevance. By characterizing how individuals move through affectively and self-referentially meaningful states in the continuous VST space, the combination of FAST and DMPM enables tests of whether the unfolding of thought in daily life corresponds to biomarkers implicated in long-term health risk. Such dynamical signatures may reflect early vulnerability before overt clinical impairment and may also be useful for tracking change over time. Specifically, the FAST/DMPM framework could serve as a tool for monitoring risk, with persistently recurring negative or past-oriented patterns potentially serving as behavioral markers of vulnerability to depression. These signatures may also help identify personalized intervention targets by highlighting how thought patterns unfold over time, beyond thought content alone. More broadly, this approach may provide a scalable way to assess changes in thought dynamics during the course of psychotherapy or recovery. The preliminary association with CRP further suggests that thought dynamics may relate to broader biological risk pathways, motivating integrative models that span subjective experience and physiology. Overall, these findings position FAST and DMPM as a scalable tool for identifying maladaptive thought-dynamic phenotypes, potentially supporting longitudinal prediction and mechanistic investigation, and informing personalized prevention and intervention strategies that target not only what people think, but how their thoughts evolve.

DATA AVAILABILITY STATEMENT

ARTICLE IN PRESS

ARTICLE IN PRESS

The predictive models and data to generate the main figures are available at https://doi.org/10.5281/zenodo.10445890. The data not used in the main figures will be shared upon request to the corresponding author (B.K.L.: byeolkimlux@gmail.com).

CODE AVAILABILITY STATEMENT

The code to generate the main figures is available at https://doi.org/10.5281/zenodo.10445890. In-house Matlab codes for behavior data analyses are available at https://github.com/canlab/CanlabCore and https://github.com/cocoanlab/cocoanCORE.

ARTICLE IN PRESS

ARTICLE IN PRESS

References

1. Callard, F., Smallwood, J., Golchert, J. & Margulies, D. S. The era of the wandering mind? Twenty-first century research on self-generated mental activity. *Front. Psychol.* **4**, (2013).
2. Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N. & Andrews-Hanna, J. R. Mind-wandering as spontaneous thought: a dynamic framework. *Nature Reviews Neuroscience* **17**, 718–731 (2016).
3. Schooler, J. W. *et al.* Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences* **15**, 319–326 (2011).
4. D'Argembeau, A. *Mind-Wandering and Self-Referential Thought*. vol. 1 (Oxford University Press, 2018).
5. Girn, M., Mills, C., Roseman, L., Carhart-Harris, R. L. & Christoff, K. Updating the dynamic framework of thought: Creativity and psychedelics. *NeuroImage* **213**, 116726 (2020).
6. Marchetti, I., Koster, E. H. W., Klinger, E. & Alloy, L. B. Spontaneous Thought and Vulnerability to Mood Disorders: The Dark Side of the Wandering Mind. *Clinical Psychological Science* **4**, 835–857 (2016).
7. Mildner, J. N. & Tamir, D. I. Why do we think? The dynamics of spontaneous thought reveal its functions. *PNAS Nexus* **3**, pgae230 (2024).
8. Gray, K. *et al.* "Forward flow": A new measure to quantify free thought and predict creativity. *American Psychologist* **74**, 539–554 (2019).
9. Lapate, R. C., Samaha, J., Rokers, B., Postle, B. R. & Davidson, R. J. Perceptual metacognition of human faces is causally supported by function of the lateral prefrontal cortex. *Commun Biol* **3**, 360 (2020).
10. Nolen-Hoeksema, S., Wisco, B. E. & Lyubomirsky, S. Rethinking Rumination. *Perspectives on Psychological Science* **3**, 400–424 (2008).
11. Smallwood, J., Ruby, F. J. M. & Singer, T. Letting go of the present: Mind-wandering is associated with reduced delay discounting. *Consciousness and Cognition* **22**, 1–7 (2013).
12. Mildner, J. N. & Tamir, D. I. Spontaneous Thought as an Unconstrained Memory Process. *Trends in Neurosciences* **42**, 763–777 (2019).
13. Corbani, F., Mildner, J. N., Bassett, D. S. & Tamir, D. I. Tracking the dynamics of the stream of thought reveals its function. *Technology, Mind, and Behavior* **6**, 380–393 (2025).
14. Killingsworth, M. A. & Gilbert, D. T. A Wandering Mind Is an Unhappy Mind. *Science* 2 (2010).
15. Andrews-Hanna, J. R. *et al.* A penny for your thoughts: dimensions of self-generated thought content and relationships with individual differences in emotional wellbeing. *Frontiers in Psychology* **4**, (2013).
16. Kanske, P. *et al.* The wandering mind in borderline personality disorder: Instability in self- and other-related thoughts. *Psychiatry Research* **242**, 302–310 (2016).
17. Ruby, F. J. M., Smallwood, J., Engen, H. & Singer, T. How Self-Generated Thought Shapes Mood—The Relation between Mind-Wandering and Mood Depends on the Socio-Temporal Content of Thoughts. *PLoS ONE* **8**, e77554 (2013).
18. Andrews-Hanna, J. R. *et al.* The conceptual building blocks of everyday thought: Tracking the emergence and dynamics of ruminative and nonruminative thinking. *Journal of Experimental Psychology: General* https://doi.org/10.1037/xge0001096 (2021) doi:10.1037/xge0001096.

ARTICLE IN PRESS

ARTICLE IN PRESS

19. Lux, B. K., Andrews-Hanna, J. R., Han, J., Lee, E. & Woo, C.-W. When self comes to a wandering mind: Brain representations and dynamics of self-generated concepts in spontaneous thought. *Science Advances* **8**, eabn8616 (2022).
20. Watson, D., Clark, L. A. & Carey, G. Positive and Negative Affectivity and Their Relation to Anxiety and Depressive Disorders. 8 (1988) doi:10.1037/0021-843X.97.3.346.
21. Stanton, K. & Watson, D. Positive and Negative Affective Dysfunction in Psychopathology. *Social and Personality Psychology Compass* **8**, 555–567 (2014).
22. Watson, D. & Clark, L. A. Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin* **96**, 465–490 (1984).
23. Barlow, D. H., Sauer-Zavala, S., Carl, J. R., Bullis, J. R. & Ellard, K. K. The Nature, Diagnosis, and Treatment of Neuroticism: Back to the Future. in *The Neurotic Paradox, Vol 2* (Routledge, 2016).
24. Clark, L. A., Watson, D. & Mineka, S. Temperament, personality, and the mood and anxiety disorders. *Journal of abnormal psychology* **103**, 103 (1994).
25. Carver, C. S. Approach, Avoidance, and the Self-Regulation of Affect and Action. *Motiv Emot* **30**, 105–110 (2006).
26. Fredrickson, B. L. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist* **56**, 218–226 (2001).
27. Pressman, S. D. & Cohen, S. Does positive affect influence health? *Psychological Bulletin* **131**, 925–971 (2005).
28. Pizzagalli, D. A. Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology* **10**, 393–423 (2014).
29. Ahrens, A. H. & Haaga, D. A. F. The specificity of attributional style and expectations to positive and negative affectivity, depression, and anxiety. *Cogn Ther Res* **17**, 83–98 (1993).
30. Warr, P. B., Barter, J. & Brownbridge, G. On the independence of positive and negative affect. *Journal of Personality and Social Psychology* **44**, 644–651 (1983).
31. Fredrickson, B. L. & Losada, M. F. Positive Affect and the Complex Dynamics of Human Flourishing. *American Psychologist* **60**, 678–686 (2005).
32. Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W. & Kelley, K. W. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci* **9**, 46–56 (2008).
33. Miller, A. H. & Raison, C. L. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat Rev Immunol* **16**, 22–34 (2016).
34. Zorrilla, E. P. *et al.* The Relationship of Depression and Stressors to Immunological Assays: A Meta-Analytic Review. *Brain, Behavior, and Immunity* **15**, 199–226 (2001).
35. Dowlati, Y. *et al.* A Meta-Analysis of Cytokines in Major Depression. *Biological Psychiatry* **67**, 446–457 (2010).
36. Haapakoski, R., Mathieu, J., Ebmeier, K. P., Alenius, H. & Kivimäki, M. Cumulative meta-analysis of interleukins 6 and 1β, tumour necrosis factor α and C-reactive protein in patients with major depressive disorder. *Brain, Behavior, and Immunity* **49**, 206–215 (2015).
37. Herbert, T. B. & Cohen, S. Depression and immunity: A meta-analytic review. *Psychological Bulletin* **113**, 472–486 (1993).
38. Howren, M. B., Lamkin, D. M. & Suls, J. Associations of Depression With C-Reactive Protein, IL-1, and IL-6: A Meta-Analysis. *Psychosomatic Medicine* **71**, 171–186 (2009).

ARTICLE IN PRESS

ARTICLE IN PRESS

39. Köhler, C. A. *et al.* Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. *Acta Psychiatrica Scandinavica* **135**, 373–387 (2017).
40. Slavich, G. M. & Irwin, M. R. From stress to inflammation and major depressive disorder: A social signal transduction theory of depression. *Psychological Bulletin* **140**, 774–815 (2014).
41. Dantzer, R. Cytokine-Induced Sickness Behavior: Mechanisms and Implications. *Annals of the New York Academy of Sciences* **933**, 222–234 (2001).
42. Dantzer, R. *et al.* Cytokines and Sickness Behavior. *Annals of the New York Academy of Sciences* **840**, 586–590 (1998).
43. Dantzer, R. & Kelley, K. W. Twenty years of research on cytokine-induced sickness behavior. *Brain, Behavior, and Immunity* **21**, 153–160 (2007).
44. Schiepers, O. J. G., Wichers, M. C. & Maes, M. Cytokines and major depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **29**, 201–217 (2005).
45. Felger, J. C. *et al.* Effects of Interferon-Alpha on Rhesus Monkeys: A Nonhuman Primate Model of Cytokine-Induced Depression. *Biological Psychiatry* **62**, 1324–1333 (2007).
46. Frenois, F. *et al.* Lipopolysaccharide induces delayed FosB/DeltaFosB immunostaining within the mouse extended amygdala, hippocampus and hypothalamus, that parallel the expression of depressive-like behavior. *Psychoneuroendocrinology* **32**, 516–531 (2007).
47. Kaster, M. P., Gadotti, V. M., Calixto, J. B., Santos, A. R. S. & Rodrigues, A. L. S. Depressive-like behavior induced by tumor necrosis factor-α in mice. *Neuropharmacology* **62**, 419–426 (2012).
48. Song, C. & Wang, H. Cytokines mediated inflammation and decreased neurogenesis in animal models of depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **35**, 760–768 (2011).
49. Kappelmann, N., Lewis, G., Dantzer, R., Jones, P. B. & Khandaker, G. M. Antidepressant activity of anti-cytokine treatment: a systematic review and meta-analysis of clinical trials of chronic inflammatory conditions. *Mol Psychiatry* **23**, 335–343 (2018).
50. Raison, C. L. *et al.* A Randomized Controlled Trial of the Tumor Necrosis Factor Antagonist Infliximab for Treatment-Resistant Depression: The Role of Baseline Inflammatory Biomarkers. *JAMA Psychiatry* **70**, 31–41 (2013).
51. Segerstrom, S. C. & Miller, G. E. Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry. *Psychological Bulletin* **130**, 601–630 (2004).
52. Steptoe, A., Hamer, M. & Chida, Y. The effects of acute psychological stress on circulating inflammatory factors in humans: A review and meta-analysis. *Brain, Behavior, and Immunity* **21**, 901–912 (2007).
53. Glaser, R. & Kiecolt-Glaser, J. K. Stress-induced immune dysfunction: implications for health. *Nat Rev Immunol* **5**, 243–251 (2005).
54. Kiecolt-Glaser, J. K. *et al.* Chronic stress and age-related increases in the proinflammatory cytokine IL-6. *Proceedings of the National Academy of Sciences* **100**, 9090–9095 (2003).
55. Pepys, M. B. & Hirschfield, G. M. C-reactive protein: a critical update. *J Clin Invest* **111**, 1805–1812 (2003).
56. Ridker, P. M., Buring, J. E., Cook, N. R. & Rifai, N. C-Reactive Protein, the Metabolic Syndrome, and Risk of Incident Cardiovascular Events. *Circulation* **107**, 391–397 (2003).
57. Hiles, S. A., Baker, A. L., de Malmanche, T. & Attia, J. A meta-analysis of differences in IL-6 and IL-10 between people with and without depression: Exploring the causes of heterogeneity. *Brain, Behavior, and Immunity* **26**, 1180–1188 (2012).

ARTICLE IN PRESS

ARTICLE IN PRESS

58. Kim, H. J., Lux, B. K., Lee, E., Finn, E. S. & Woo, C.-W. Brain decoding of spontaneous thought: Predictive modeling of self-relevance and valence using personal narratives. *Proceedings of the National Academy of Sciences* **121**, e2401959121 (2024).
59. Shin, W.-G. *et al*. Happier individuals generate more spontaneous thoughts about friends and value relationships over money. *Commun Psychol* **3**, 162 (2025).
60. Radloff, L. S. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* **1**, 385–401 (1977).
61. Watson, D., Anna, L. & Tellegen, A. Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. 8 (1988) doi:10.1037/0022-3514.54.6.1063.
62. Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. & Natalicio, D. S. The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology* **5**, (1971).
63. Reynolds, W. M. *Suicidal Ideation Questionnaire*. (Psychological Assessment Resources Odessa, FL, 1987).
64. Lyubomirsky, S., Caldwell, N. D. & Nolen-Hoeksema, S. Effects of Ruminative and Distracting Responses to Depressed Mood on Retrieval of Autobiographical Memories. 12 (1998).
65. Hughes, M. E., Waite, L. J., Hawkley, L. C. & Cacioppo, J. T. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. *Res Aging* **26**, 655–672 (2004).
66. Diener, E., Emmons, R. A., Larsen, R. J. & Griffin, S. The Satisfaction With Life Scale. *Journal of Personality Assessment* **49**, 71–75 (1985).
67. Ryff, C. D. & Keyes, C. L. M. The structure of psychological well-being revisited. *Journal of personality and social psychology* **69**, 719 (1995).
68. Wardenaar, K. J. *et al.* Development and validation of a 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (MASQ). *Psychiatry Research* **179**, 101–106 (2010).
69. Lee, H. H., Kim, E. J. & Lee, M. A validation study of Korea positive and negative affect schedule: The PANAS scales. *Korean Journal of Clinical Psychology* **22**, 935–946 (2003).
70. Lee, H. & Kim, K.-H. Validation of the Korean version of the mood and anxiety symptom questionnaire (K-MASQ). *Korean Journal of Clinical Psychology* **33**, 395–411 (2014).
71. Kim, S. J., Kim, J. H. & Yoon, S. C. Validation of the Korean-ruminative response scale (K-RRS). *Korean Journal of Clinical Psychology* **29**, 1–19 (2010).
72. Kim, S. J. & Chung, M. J. The effect of family therapist's specialty on their marital relationship and family's psychological health through their self-efficacy and psychological well-being. *Korea Journal of Counseling* **11**, 1289–1304 (2010).
73. Cho, H.-J., Chae, J.-H. & Jun, T.-Y. The overview of clinical assessment tools for depression. *Journal of Korean Neuropsychiatric Association* 110–121 (2007).
74. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of Best Practices for Evidence for Prediction: A Review. *JAMA Psychiatry* **77**, 534–540 (2020).
75. Mor, N. & Winquist, J. Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin* **128**, 638–662 (2002).
76. Raffaelli, Q. *et al.* The think aloud paradigm reveals differences in the content, dynamics and conceptual scope of resting state thought in trait brooding. *Sci Rep* **11**, 19362 (2021).

ARTICLE IN PRESS

ARTICLE IN PRESS

77. Kucyi, A., Kam, J. W. Y., Andrews-Hanna, J. R., Christoff, K. & Whitfield-Gabrieli, S. Recent advances in the neuroscience of spontaneous and off-task thought: implications for mental health. *Nat. Mental Health* **1**, 827–840 (2023).
78. Taylor, S. E. & Brown, J. D. Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin* **103**, 193–210 (1988).
79. Watson, L. A., Dritschel, B., Obonsawin, M. C. & Jentzsch, I. Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research* **1152**, 106–110 (2007).
80. Beck, A. T. Cognitive models of depression. *Journal of Cognitive Psychotherapy* **1**, 5–37 (1987).
81. Disner, S. G., Beevers, C. G., Haigh, E. A. P. & Beck, A. T. Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci* **12**, 467–477 (2011).
82. Haaga, D., Dyck, M. & Ernst, D. Empirical status of cognitive theory of depression.Psychological Bulletin, 110, 215-236. *Psychological bulletin* **110**, 215–36 (1991).
83. Segal, Z. V. Appraisal of the self-schema construct in cognitive models of depression. *Psychological Bulletin* **103**, 147–162 (1988).
84. Hamilton, J. P., Farmer, M., Fogelman, P. & Gotlib, I. H. Depressive Rumination, the Default-Mode Network, and the Dark Matter of Clinical Neuroscience. *Biological Psychiatry* **78**, 224–230 (2015).
85. Scalabrini, A. *et al.* All roads lead to the default-mode network—global source of DMN abnormalities in major depressive disorder. *Neuropsychopharmacol.* **45**, 2058–2069 (2020).
86. Sheline, Y. I. *et al.* The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences* **106**, 1942–1947 (2009).
87. Eshel, N. & Roiser, J. P. Reward and Punishment Processing in Depression. *Biological Psychiatry* **68**, 118–124 (2010).
88. Martin-Soelch, C. Is depression associated with dysfunction of the central reward system? *Biochem Soc Trans* **37**, 313–317 (2009).
89. Ganesh, K. & Gabora, L. A Dynamic Autocatalytic Network Model of Therapeutic Change. *Entropy* **24**, 547 (2022).
90. Denny, B. T., Kober, H., Wager, T. D. & Ochsner, K. N. A Meta-analysis of Functional Neuroimaging Studies of Self- and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience* **24**, 1742–1752 (2012).
91. Leary, M. R., Gohar, D. & Gross, J. J. Self-awareness and self-relevant thought in the experience and regulation of emotion. *Handbook of emotion regulation* 376–389 (2014).
92. Thornton, M. A., Weaverdyck, M. E., Mildner, J. N. & Tamir, D. I. People represent their own mental states more distinctly than those of others. *Nat Commun* **10**, 2117 (2019).
93. Schacter, D. L., Addis, D. R. & Buckner, R. L. Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience* **8**, 657–661 (2007).
94. Ross, M. & Newby-Clark, I. R. Construing the Past and Future. *Social Cognition* **16**, 133–150 (1998).
95. Irish, M. & Piguet, O. The Pivotal Role of Semantic Memory in Remembering the Past and Imagining the Future. *Front. Behav. Neurosci.* **7**, (2013).
96. Smallwood, J. & O’Connor, R. C. Imprisoned by the past: Unhappy moods lead to a retrospective bias to mind wandering. *Cognition & Emotion* **25**, 1481–1490 (2011).

ARTICLE IN PRESS

ARTICLE IN PRESS

97. Fitzsimmons, T. R., Sanders, A. E., Bartold, P. M. & Slade, G. D. Local and systemic biomarkers in gingival crevicular fluid increase odds of periodontitis. *Journal of clinical periodontology* **37**, 30–36 (2010).
98. Rhodus, N. L. *et al.* A comparison of the pro-inflammatory, NF-κB-dependent cytokines: TNF-alpha, IL-1-alpha, IL-6, and IL-8 in different oral fluids from oral lichen planus patients. *Clinical Immunology* **114**, 278–283 (2005).
99. Luan, Y. & Yao, Y. The Clinical Significance and Potential Role of C-Reactive Protein in Chronic Inflammatory and Neurodegenerative Diseases. *Front. Immunol.* **9**, (2018).

ARTICLE IN PRESS

ARTICLE IN PRESS

AUTHOR CONTRIBUTIONS

B.K.L. and E.L. shared their first authorship.

B.K.L.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, writing–original draft, and writing–review & editing.

E.L.: data curation, formal analysis, investigation, visualization, and writing–original draft, and writing–review & editing.

J.H.: formal analysis, investigation, and software.

S.-H. L., S.G., and I.C.: formal analysis, and investigation.

C.-W. W.: conceptualization, formal analysis, funding acquisition, methodology, resources, supervision, visualization, writing–original draft, and writing–review & editing.

CONFLICT OF INTEREST STATEMENTS

The authors declare no competing interests.

FUNDING

This work was supported by IBS-R015-D2 (Institute for Basic Science, South Korea; to C.-W.W.). B.K.L. was supported by Fulbright Foreign Student Program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY INFORMATION

Supplementary Figures 1-8

Supplementary Tables 1

ARTICLE IN PRESS

ARTICLE IN PRESS

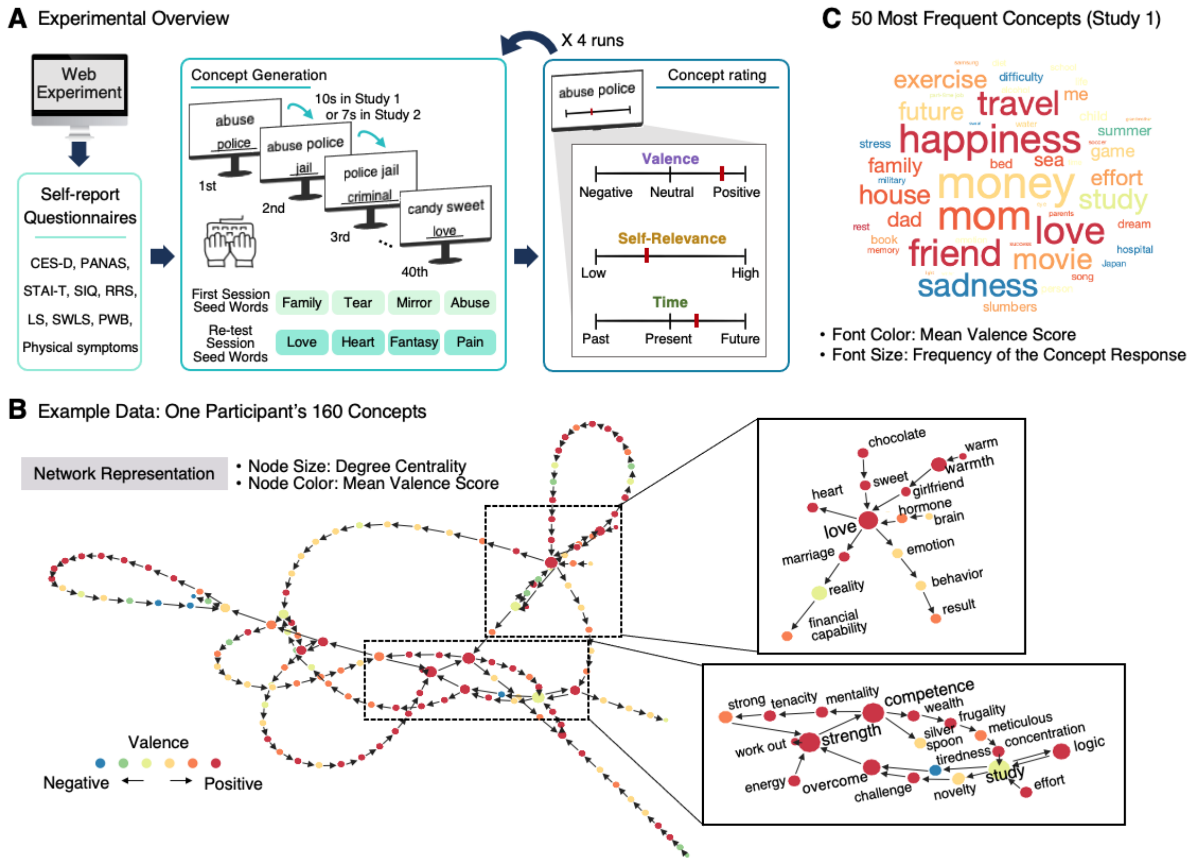

Fig. 1. Overview of experimental design and sample data visualization. (a) Experimental procedure of the web-based Free Association Semantic Task (FAST) involved concept generation and concept rating. During concept generation, participants in Study 1 and Study 2 were instructed to type a word or phrase related to the previous concept every 10 and 7 seconds, respectively, starting from a given seed word. Each participant generated 40 consecutive concepts per seed word. In the subsequent concept rating phase, participants evaluated these self-generated concepts across three content dimensions, including valence, self-relevance, and time. Two consecutive concepts were displayed for context, with participants rating the target concept on the three scales. This dual-phase task was repeated for four different pre-selected seed words. Participants also completed a series of self-report questionnaires prior to the web-based FAST experiment, assessing their affectivity and mental health. **(b)** A network representation of one participant's self-generated concepts. Nodes represent individual concepts, while directed edges (arrows) show the associative connections between concepts. Node size reflects the degree centrality (number of connected edges), and node color indicates the mean valence scores of each concept. **(c)** A word cloud visualizes the 50 most frequently reported concepts aggregated from the first session of Study 1 (N = 117). The size of each word in the cloud corresponds to its frequency of occurrence, and the color indicates the mean valence scores of each concept.

ARTICLE IN PRESS

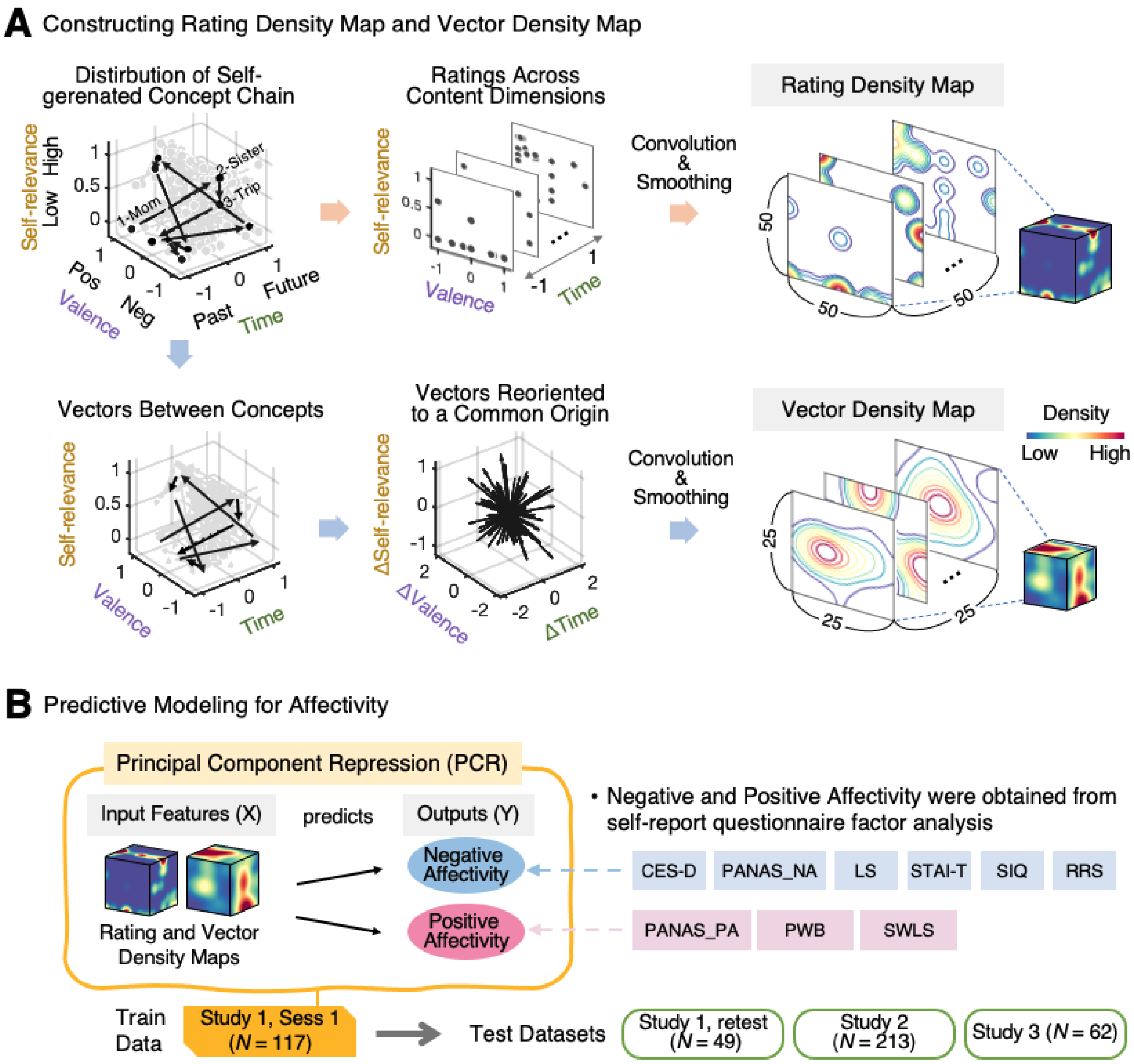

Fig. 2. Implementation of Dynamic Density Map-Based Predictive Modeling (DMPM) (a) The development of DMPM involved creating two types of density maps: rating and vector density maps. The process began by plotting one participant's self-generated concept ratings and the directional vectors between consecutive concepts within a three-dimensional space defined by valence, self-relevance, and time. The rating density map (upper panel) visualizes the frequency of concepts at specific coordinates, creating an accumulative value matrix. The vector density map depicts the directional transitions between concepts. The difference vectors between concepts are shifted to originate from a common origin point, and their heads are mapped to show frequency across coordinates. Both maps undergo refinement through convolution and smoothing. **(b)** Utilizing these density maps as input features, we constructed the Positive Affectivity (PA) model and Negative Affectivity (NA) model through Principal Component

ARTICLE IN PRESS

Regression (PCR) with leave-one-subject-out cross-validation. PA and NA scores, the outcome variables for predictive modeling, were calculated through factor analysis of self-report questionnaires. The models were trained using data from the first session of Study 1 (N = 117) and tested on three test datasets: the re-test session of Study 1 (N = 49), Study 2 (N = 213), and Study 3 data (N = 62).

ARTICLE IN PRESS

ARTICLE IN PRESS

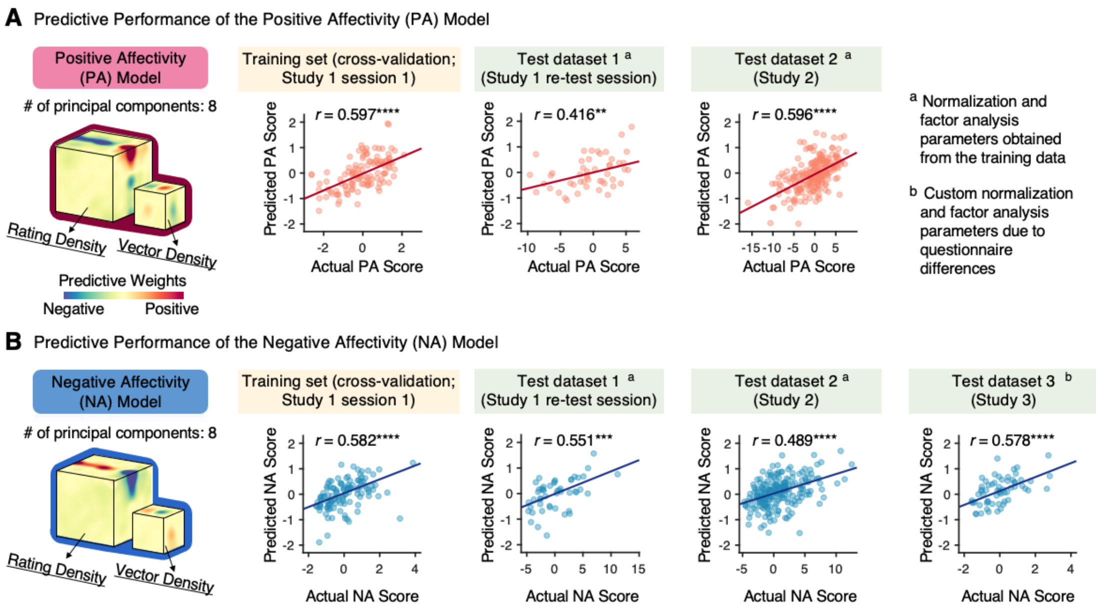

Fig. 3. Evaluation of predictive model performance. (a) Performances of Positive Affectivity (PA) predictive models. The left panel illustrates a schematic representation of the PA predictive model, where the larger cube represents the rating density-based features, and the smaller cube represents the vector density-based features. The color scheme in the diagram corresponds to the predictive weights (details in **Fig. 4**). The right panel displays three sets of results: 1) the leave-one-participant-out cross-validated predictive performance within the training dataset (N = 117, Study 1 session 1), 2) test results on the re-test data of Study 1 with different seed words (N = 49), and 3) test results on Study 2 data (N = 213). The plots represent the correlation between actual and predicted PA scores for each participant. The PA model was not evaluated in Study 3 due to a limited number of PA-related questionnaires. Model performance was assessed using a robust correlation method to provide more accurate estimates by down-weighting the influences of outliers. **(b)** Performances of Negative Affectivity (NA) predictive models. The left panel shows a schematic representation of the NA model. The right panel presents results from 1) cross-validated predictions within the training dataset, followed by test results on 2) the re-test data of Study 1, 3) Study 2 data, and 4) Study 3 data (N = 62). Note that the y-axis scales vary across study panels because Study 3 utilized a distinct set of questionnaires for factor analysis, resulting in outcome variables with different scales. $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$, two-tailed, one-sample t-test.

ARTICLE IN PRESS

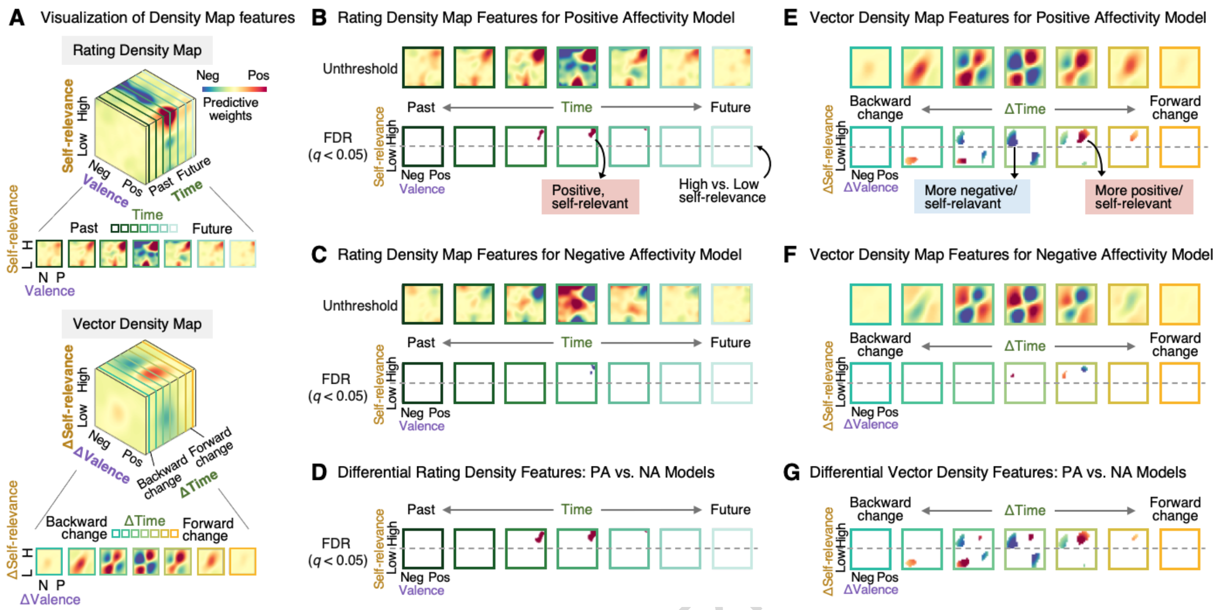

Fig. 4. Significant features of the DMPM predictive models. To identify reliable predictive features, we conducted bootstrap analyses with 10,000 iterations. **(a)** The density maps were segmented along the temporal axis into seven slices, outlined in different colors to represent distinct time coordinates. Each slice is shown as a small square: for the rating density map, axes correspond to valence and self-relevance; for vector density map, axes correspond to Δvalence and Δself-relevance. **(b-c)** Rating density features for Positive Affectivity (PA) and Negative Affectivity (MA) models. Upper panels show unthresholded rating density feature weights; lower panels show significant features after FDR thresholding ($q < 0.05$). The dashed line in the lower panel divides high versus low self-relevance areas. Blue indicates features predicting lower PA or NA, whereas red indicates features predicting higher PA or NA. Overall, positive, present-focused, and highly self-relevant concepts predict higher PA and lower NA. **(d)** Contrast map showing differences in rating density features between the PA and NA models. **(e-f)** Vector density features for the PA and NA models. For PA, predictive transitions were characterized by (1) increased transitions towards more positive/more self-relevance, with forward (prospective) temporal change, (2) reduced transitions toward more negative/self-relevant and more positive/self-irrelevant concepts, with minimal temporal change, and (3) increased transitions toward more negative, less self-relevance, and backward (retrospective) temporal change. In contrast, NA showed an overall opposing pattern, with more temporally restricted predictive features, including (1) increased transitions toward more negative, more self-relevance, without temporal change and (2) reduced transitions toward more positive and more self-relevance, and forward (prospective) temporal change. **(g)** Contrast map showing differences in vector density features between the PA and NA models.

ARTICLE IN PRESS

ARTICLE IN PRESS

ARTICLE IN PRESS

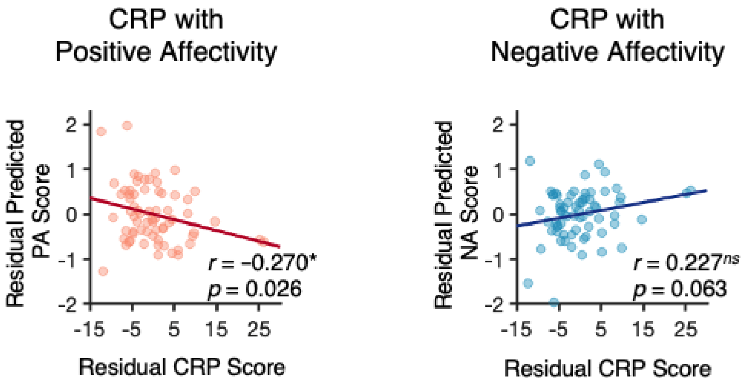

Fig. 5. Association between CRP and the model-predicted affectivity. To investigate the relationship inflammatory markers and model-predicted Positive Affectivity and Negative Affectivity, we conducted partial correlation analyses controlling for Body Mass Index (BMI), age, gender, tobacco use, and alcohol consumption. Each point represents an individual participant (N = 68). Residualized CRP levels were negatively correlation with residualized predicted PA scores, suggesting that lower CRP levels were associated with higher PA scores after accounting for covariates. In contrast, the correlation between residualized CRP and residualized predicted NA scores was not significant. ^{ns}not significant, $^{*}p < 0.05$.

ARTICLE IN PRESS

ARTICLE IN PRESS

ARTICLE IN PRESS