Stress and Health

WILEY

RESEARCH ARTICLE

Combining Life History Calendars and Ecological Momentary Assessment for Evaluating Everyday Stress and Its Impact on Mental Health in Healthy Adults: Longitudinal Study

Pamela Chavarría-Elizondo[1,2,3] | Miquel Tortella-Feliu[4] | Lydia Fortea[5,6] | Víctor De la Peña-Arteaga[7] | Asier Juaneda-Seguí[1] | Ignacio Martínez-Zalacaín[8] | Enric Vilajosana[5] | Joaquim Radua[3,5,6,9] | Alexander J. Shackman[10] | Carles Soriano-Mas[1,3,11] | Miquel A. Fullana[3,5,12]

[1]Psychiatry and Mental Health Group, Neuroscience Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Spain | [2]Faculty of Medicine and Health Sciences, Department of Clinical Sciences, Campus Bellvitge, University of Barcelona, L'Hospitalet de Llobregat, Spain | [3]Network Center for Biomedical Research on Mental Health (CIBERSAM), Carlos III Health Institute (ISCIII), Madrid, Spain | [4]University Research Institute on Health Sciences (IUNICS), University of the Balearic Islands, Mallorca, Spain | [5]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain | [6]Faculty of Medicine and Health Sciences, Department of Medicine, Campus Clínic, University of Barcelona, Barcelona, Spain | [7]Sant Pau Mental Health Research Group, Institut de Recerca Sant Pau (IR SANT PAU), Barcelona, Spain | [8]Department of Radiology, Bellvitge University Hospital, Barcelona, Spain | [9]Bipolar and Depressive Disorders Unit, Hospital Clínic de Barcelona, Barcelona, Spain | [10]Department of Psychology, Neuroscience and Cognitive Science Program, and Maryland Neuroimaging Center, University of Maryland, College Park, Maryland, USA | [11]Department of Social Psychology and Quantitative Psychology, Institute of Neurosciences, University of Barcelona, Barcelona, Spain | [12]Adult Psychiatry and Psychology Department, Institute of Neurosciences, Hospital Clinic, Barcelona, Spain

Correspondence: Carles Soriano-Mas (carles.soriano.mas@ub.edu) | Miquel A. Fullana (mafullana@clinic.cat)

Received: 19 March 2025 | **Revised:** 11 August 2025 | **Accepted:** 21 November 2025

Keywords: ecological momentary assessment | life history calendar | pileup | stress | stress reactivity | stress recovery | stressors

ABSTRACT

Stress is a well-established risk factor for a range of negative health outcomes, making the longitudinal assessment of everyday stress increasingly important. This study tested the feasibility of combining a prospective assessment of stress responses using Ecological Momentary Assessment (EMA) with a retrospective evaluation of stressors via a Life History Calendar (LHC). It also examined how different operationalizations of stress components—reactivity, recovery, and pileup—using various intra-individual baselines influenced stressor count, mental health symptoms, and trait anxiety. A sample of 165 adults (50.9% female; *M* age = 24.91 years, SD = 4.61) was followed over six months. Stress responses were assessed prospectively via EMA, administered three times daily, and stressor exposure was retrospectively captured using the LHC. We computed indices of stress reactivity, recovery, and pileup using three baseline types: local (1-week and 2-week periods prior to a stressor episode) and a cumulative average of all prior stressor-free weeks. Hierarchical multiple regression analyses examined associations between these indices and changes in mental health symptoms and trait anxiety. Combining EMA and LHC methods to assess stress components was feasible. Baseline choice influenced both stressor episode counts and the associations between stress components and mental health outcomes. Stress reactivity and pileup, based on both local and cumulative baselines, were consistently associated with increases in mental health symptoms. Only one operationalisation of stress recovery was linked to changes in symptoms. Trait anxiety was not associated with any stress component, except for pileup when using a 1-week local

Abbreviations: DASS-21, Depression, Anxiety, and Stress Scales; DASS-A, DASS anxiety subscale; DASS-D, DASS depression subscale; DASS-S, DASS stress subscale; EMA, ecological momentary assessment; EMA-A, EMA anxiety scores; LHC, Life History Calendar; STAI-T, State-Trait Anxiety Inventory–Trait.

baseline. These findings support the feasibility of integrating EMA and LHC methods to assess stress dynamics. Different baseline definitions yield distinct results, highlighting the importance of carefully selecting baseline parameters in stress research. This has important implications for designing and refining future studies on stress and health.

1 | Introduction

Stress is a recognized risk factor for a wide range of negative health outcomes. Defined as an individual's perception of a situation as a threat to their homoeostasis and their perceived ability to cope (Lazarus and Folkman 1984). Stress is inherently dynamic and varies both between and within individuals (Sliwinski et al. 2009). However, despite this intra-individual nature, most stress research has focused on static and group-level approaches (Heshmati et al. 2024; Smyth et al. 2023). Recently, there has been increasing interest in using longitudinal methods to study stress, emphasising its dynamic and intra-individual nature by capturing stressors and stress responses over time (Smyth et al. 2023). Ecological momentary assessment (EMA) has become a very valuable tool to capture such within-person variation, allowing for repeated sampling of individuals' experiences in their natural environments, thus enhancing ecological validity (Shiffman et al. 2008). By enabling real-time, contextually grounded measurement of stress and related processes in everyday life—insights often inaccessible through traditional methods—EMA has transformed stress research. Moreover, technological advances, such as the widespread use of smartphones, have enhanced the accessibility and usability of EMA, further accelerating this progress (Zhang et al. 2024; Weber et al. 2022; Lohani et al. 2025).

The classical definition of stress by Lazarus and Folkman's (Lazarus and Folkman 1984) highlights two key components: primary assessment (the initial response to a stressor, i.e. stress reactivity) and secondary assessment (stress recovery) (Lazarus and Folkman 1984). Building on this framework, novel approaches have been proposed for evaluating these components (Kalisch et al. 2021; Smyth et al. 2018). Specifically, Smyth et al. (2023) proposed an EMA-based approach for assessing stress components in daily life, focussing on three key elements: stress reactivity, stress recovery, and a newly introduced component—pileup, which refers to the accumulation of stressor episodes over time. To operationalise these components, they suggested several EMA-based strategies that use different types of individualised baselines, ranging from more proximal ('local', referring to a recent stress-free period) to more distal ('cumulative', referring to the average of past stressor-free periods) (Smyth et al. 2023).

This study builds on the proposal of Smyth et al. (2023), by examining how different operationalizations of stress components -using different baselines-are associated with two critical constructs: declining mental health and trait anxiety. Prior research suggests that increased stress reactivity and pileup are associated with poorer mental health outcomes (Kalisch et al. 2021; Shackman et al. 2016), while individuals with higher trait anxiety (also called neuroticism or dispositional negative affect; Shackman et al. 2016) may experience heightened stress reactivity, reduced stress recovery, and greater pileup (Kalisch et al. 2021; Shackman et al. 2016); although the evidence is not conclusive for trait anxiety/neuroticism (Mey et al. 2020). Thus, some of the EMA-based operationalizations of stress components should demonstrate significant associations with increases in mental health symptoms and with trait anxiety.

One of the challenges in assessing stress involves defining what constitutes a stressor (Cohen et al. 2019). As Kalisch et al. (2021) noted, defining a stressor as 'a situation that elicits a stress response', presents a practical conundrum, as a stressor for one person may not be a stressor for another. Moreover, the magnitude and quality of the stress response play a critical role in defining whether a situation qualifies as a stressor. One approach to this problem is to define a set of potential stressors (events or stimuli) and assess both their occurrence and the extent to which they trigger measurable stress responses (Kalisch et al. 2021). In this study, we used a Life History Calendar (LHC) -a tool widely employed in psychological research (Axinn et al. 1999; Caspi et al. 1996; Freedman et al. 1988)- to assess the presence and number of potential stressors, while EMA was used to measure stress responses. We opted to assess stressors retrospectively using the LHC rather than in real time via EMA for three main reasons. First, repeated in-the-moment questioning about stressor exposure may increase participants' overall psychological reactivity, a well-documented phenomenon in ecological assessment research (Shiffman et al. 2008; König et al. 2022). Second, we aimed to minimise participant burden, which was already substantial given the requirement to complete 11 items per day. Finally, as noted above, defining and operationalising 'a stressor' is complex, and leaving this entirely to participants' subjective interpretation in real time can introduce inconsistency (Kalisch et al. 2021).

To our knowledge, this is the first study to assess the feasibility of combining both strategies (EMA to assess stress responses and LHC to assess stressors) to examine how different operationalizations of stress components using different baselines affect stressor count and the association between stress components and mental health symptoms and trait anxiety. We expected that our findings would inform the design and refinement of future research on stress and health.

2 | Methods

2.1 | Participants

Participants were recruited for a prospective-longitudinal study on predictors of mental health symptoms (e.g., anxiety and depressive). Adults from the university community -including students and staff-residing in Barcelona were invited to participate. Participants were recruited using snowball sampling through the professional and academic networks of the research team. A large number of individuals (n = 840) were initially screened using the Spanish version (Buela-Casal et al. 2016) of

the State-Trait Anxiety Inventory–Trait (STAI-T) subscale (Spielberg et al. 1970) via a secure web system. To ensure a range of trait anxiety levels, participants were stratified into quartiles based on STAI-T scores, including individuals with current anxiety disorders. From these, 361 individuals met preliminary inclusion criteria and were further assessed through a physician-administered telephone interview using the Spanish version (Ferrando et al. 1998) of the Mini International Neuropsychiatric Interview (MINI; Sheehan et al. 1998). Inclusion criteria were: (1) age > 18, (2) fluency in Spanish, (3) owning a smartphone, and (4) willingness to participate in a neuroimaging assessment (part of the larger study). Exclusion criteria included: (1) current or past major medical illness, (2) current or past mental health disorders except for current anxiety disorders (panic disorder/agoraphobia, social anxiety disorder, or generalised anxiety disorder), (3) regular substance use (except tobacco), (4) medication that could interfere with the study, and (5) contraindications to neuroimaging assessment.

A total of 178 participants met the inclusion and exclusion criteria and started EMA data collection. The final sample - consisting of those who completed the 6 months EMA assessment phase-included 165 participants (50.9% females) with a mean age of 24.91 years (SD = 4.61; range: 19–36). In terms of racial and ethnic composition, 119 participants (72.1%) identified as White, 44 (26.7%) as Hispanic/Latino, and 2 (1.2%) as Asian. Eleven participants (6.7%) had a current diagnosis of an anxiety disorder.

2.2 | Ethical Considerations

The study protocol was reviewed and approved by the Institutional Review Board of Bellvitge University Hospital (Barcelona, Spain), under approval number PR144/16. All participants provided informed consent prior to participation, in accordance with the Declaration of Helsinki.

2.3 | Procedures

At baseline, participants underwent a neuroimaging session (as reported elsewhere) and completed self-report questionnaires (see below). They were also instructed to download the EMA app and received both verbal and written instructions detailing the EMA protocol, including plain-language definitions for each survey item. Participants were informed that they would receive €200 for completing 'most' of the EMA notifications throughout the study, without a specific number being defined. After 6 months of EMA data collection, participants attended a follow-up laboratory session where they completed the same self-report questionnaires as at baseline and underwent a LHC assessment to measure stressor exposure (see Figure 1).

2.4 | Assessment of Trait Anxiety and Mental Health Symptoms

The STAI-T completed during screening was used as the baseline measure of trait anxiety (also referred to as 'neuroticism/negative emotionality'; Shackman et al. 2016). The Spanish version of the STAI-T, which has total scores ranging from 0 to 60, exhibits excellent psychometric properties (Guillén-Riquelme and Buela-Casal 2011), with an internal consistency (Cronbach's alpha) of 0.90 in this study.

At both baseline and follow-up, participants also completed the Spanish version (Daza et al. 2002) of the Depression, Anxiety, and Stress Scales (DASS-21; Lovibond and Lovibond 1995), which includes 21 items across three subscales: depression (DASS-D), anxiety (DASS-A), and stress (DASS-S). The total scores for each subscale range from 0 to 21. The internal consistency of the subscales in this study, measured by Cronbach's alpha at both time points, was high: DASS-D (0.85/.85), DASS-A (0.75/.77), and DASS-S (0.84/.88).

2.5 | Assessment of Potential Stressors

Exposure to potential stressors during the 6-month period was assessed at follow-up using a physician-administered LHC. Our LHC enquired about 40 events that 'represented a significant change in the participant's environment'; grouped into 13 domains. The LHC was structured as a matrix, with 24 columns representing weeks and rows listing events (see Supporting Information S1: Figure S1). Participants' responses were recorded by drawing a continuous line in the corresponding row from the week the event began to the week it ended, enabling the assessment of event presence and the number of events per week. Due to the difficulty in estimating the duration of certain events, we decided a priori that for 'Love life', 'Family and friends', 'Home', and 'Education and Work' domains, events would be coded as present for only 1 week following their onset. For other domains, events were coded based on their actual duration. Major traumatic events were included in the LHC, though only a few participants reported such events ($n = 7$, 4.3% of the sample).

2.6 | Assessment of Stress Responses Using EMA

We used a recently developed smartphone EMA app to assess stress responses over the 6-month follow-up period (Fortea et al. 2023). The EMA protocol is described in detail in the referenced report.

We used EMA-Anxiety (EMA-A) scores as our stress response indicator. These scores were calculated by averaging weekly responses to four items that measured anxiety-related symptoms: anxious apprehension, muscle tension, worry, and avoidance. Participants rated these items on a visual analogue scale (VAS), moving a slider along a horizontal line with endpoints labelled 'not at all' and 'very much'. Responses were transformed into scores ranging from 0 to 100, higher scores indicating greater symptom severity.

Participants received three daily notifications at pseudorandomized intervals - morning (9:00 a.m.–2:00 p.m.), afternoon (2:00 p.m.–7:00 p.m.), and evening (7:00 p.m.–12:00 a.m.). They had a two-hour window to respond to the items; any items not

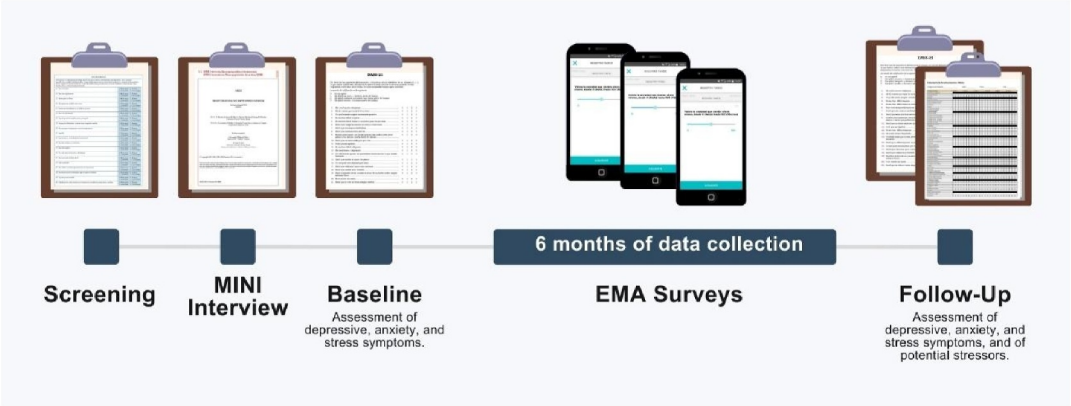

FIGURE 1 | Summary of procedures. EMA = ecological momentary assessment; MINI = MINI international neuropsychiatric interview.

completed within the timeframe were skipped until the next notification. Further details on the app's functionality can be found in (Fortea et al. 2023).

In a previous study involving the first 99 participants from the current sample, we evaluated the app's feasibility (retention and compliance) and psychometric properties (reliability and validity) over a 6-month period. The app demonstrated high feasibility, and EMA-A scores showed excellent within-person ($R_C = 0.99$) and between-person ($R_{KF} = 0.85$) reliability, strong convergent validity, moderate discriminant validity, and significant incremental validity over several self-report questionnaires (Fortea et al. 2023).

2.7 | Operationalisation of Stressors and Stress Reactivity

In longitudinal studies, stressors can occur sequentially or simultaneously, requiring careful consideration of how to define and measure the accumulation of stressors over time (Smyth et al. 2023). To address this, we defined an 'event episode' for each participant as the presence of one or more events lasting one or more weeks, with the episode starting at least 1 week without an event (except for the first event episode, which could begin in the first week). Multiple events occurring during an 'event episode' were considered part of the same episode, allowing participants to experience several event episodes (Figure 2).

As highlighted in the introduction, determining whether an event qualifies as a stressor is challenging, requiring an evaluation of stress responses. These responses must show an increase relative to a baseline for an event to be classified as a stressor. To operationalise this, we calculated stress reactivity for each event episode by measuring changes in EMA-A scores during the episode relative to a 'stressor-free' baseline. We defined 'stressor episodes' as event episodes where stress reactivity was positive (i.e., > 0) compared to the baseline (Figure 2).

Following the approach of Smyth et al. (2023), we computed stress reactivity using three different baselines (see Figure 3). Local baseline 1 was the participant's EMA-A scores in the week preceding the stressor episode. Local baseline 2 was the participant's EMA-A score 2 weeks prior. The cumulative baseline represented the participant's average stressor-free EMA-A scores throughout the study period before the onset of the stressor episode (see Figure 4).

2.8 | Operationalisation of Pileup

After defining a stressor episode for each participant, we operationalised pileup in two ways: (1) the number of stressor episodes experienced by the participant or (2) the number of weeks during which stressor episodes occurred. We then applied the three stress reactivity baselines (local baseline 1, local baseline 2 and cumulative baseline) to each of these definitions. These resulted in six different operationalizations of pileup.

2.9 | Operationalisation of Average Stress Reactivity

We calculated stress reactivity for each stressor episode using the three stress reactivity baselines mentioned above (local baseline 1, local baseline 2 and cumulative baseline). To derive an individual index of average stress reactivity, we averaged the stress reactivity scores within each participant by dividing the stress reactivity in each stressor episode by the number of episodes. This resulted in three operationalizations of average stress reactivity corresponding to each of the three baselines.

2.10 | Operationalisation of Stress Recovery and Average Stress Recovery

Following Smyth et al. (2023), we operationalised stress recovery using two approaches (see Figure 5): (1) immediate recovery, was defined as the difference between mean EMA-A scores during a stressor episode and the mean EMA-A scores during the two stressor-free weeks immediately following that stressor episode. This operationalisation is independent of pre-stressor baselines and focuses on recovery relative to the stressor episode itself. (2) cumulative recovery, was calculated by comparing the mean EMA-A scores during the weeks without a stressor after a stressor episode to the mean EMA-A scores during all the stressor-free weeks prior to the stressor episode, dating back to the start of the study.

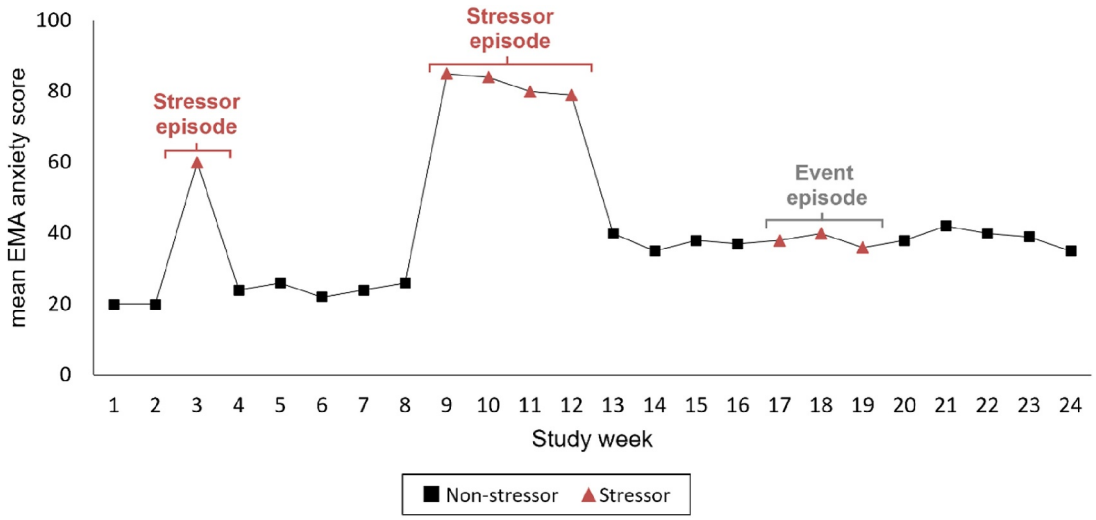

FIGURE 2 | Definition of 'event episode' and 'stressor episode' in the current study. Event episodes with increased stress reactivity relative to a stressor-free baseline (first two episodes depicted here) were defined as stressor episodes (see main text).

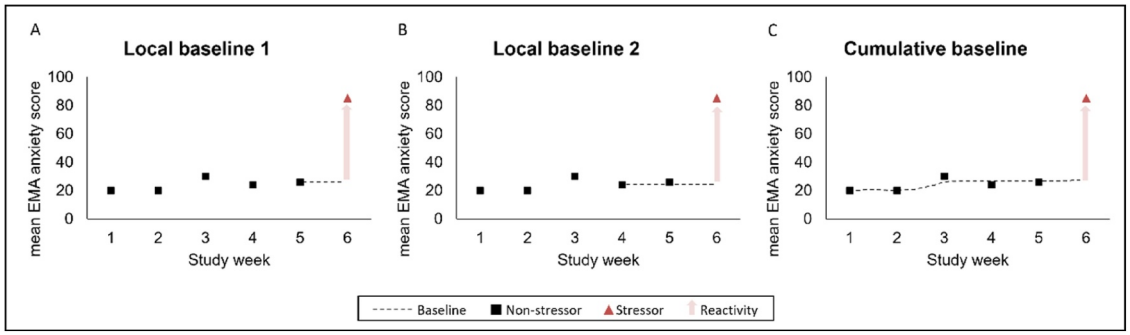

FIGURE 3 | Baselines that were defined in the current study to assess stress reactivity. Each baseline represents within-person mean ecological momentary assessment anxiety (EMA-A) scores. (A) Local baseline 1: computes the mean EMA-A scores in the non-stressor week before the stressor episode. (B) Local baseline 2: computes the mean EMA-A scores of the stressor-free weeks 2 weeks before the stressor episode. (C) Cumulative baseline: computes the mean EMA-A scores of all the stressor-free weeks since the start of the study.

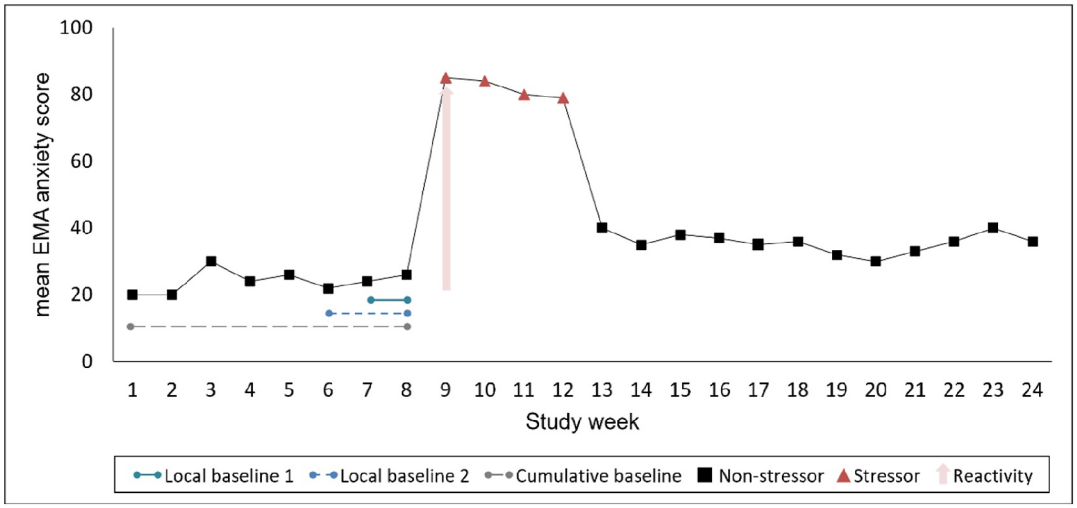

FIGURE 4 | Examples of the three reactivity baselines applied to one stressor episode.

This approach considers the participant's cumulative stress response history and whether stress levels return to that cumulative baseline (see Figure 6).

To define stress recovery, the timing and extent of initial stress reactivity must also be considered, as reactivity baselines influence recovery (Smyth et al. 2023). Therefore, we combined

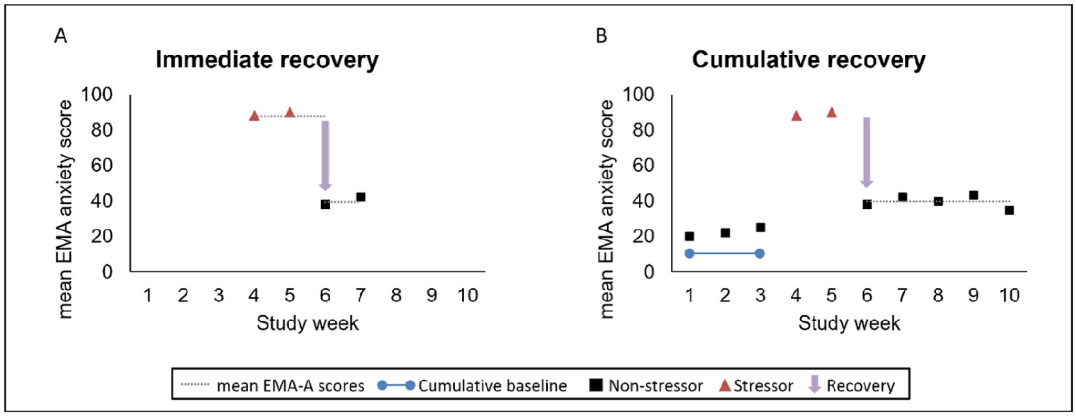

FIGURE 5 | Operationalizations of stress recovery used in the current study. (A) Immediate recovery: computes the change in mean ecological momentary assessment anxiety (EMA-A) scores from the stressor episode to the following two stressor-free weeks. (B) Cumulative recovery: computes the change in mean EMA-A scores during the weeks without a stressor after a stressor episode compared to the mean EMA-A scores during all the stressor-free weeks before the stressor episode since the onset of the study (cumulative baseline).

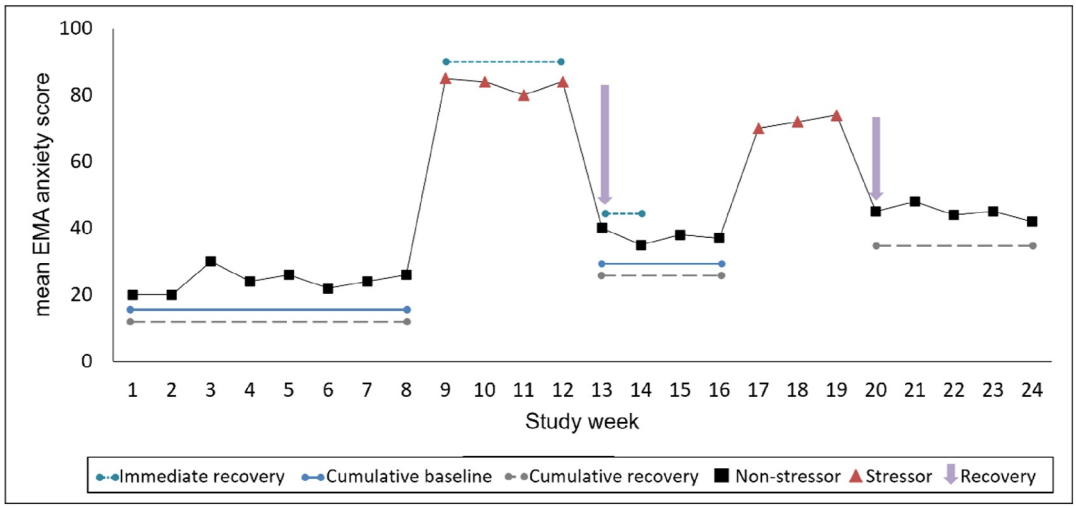

FIGURE 6 | Examples of the recovery operationalizations applied to two stressor episodes. Note that the operationalizations of stress recovery also considered reactivity baselines (see main text).

the two recovery operationalizations (immediate and cumulative) with the three reactivity baselines (local baseline 1, local baseline 2 and cumulative baseline) to create six operationalizations of recovery. We then averaged stress recovery scores across stressor episodes for each participant (stress recovery for each stressor episode divided by the number of episodes) and calculated individual indices of average stress recovery using the resulting six operationalizations.

2.11 | Statistical Analyses

Statistical analyses were conducted using IBM SPSS Statistics v.23. In preliminary analyses, we calculated descriptive statistics for symptom levels (DASS subscale scores), EMA-A scores, and the number of EMA notifications answered or unanswered, as well as changes in symptom levels using repeated-measures Student's *t*-tests.

Descriptive statistics were calculated for the number of events reported in the LHC and the number of stressor episodes according to the proposed operationalizations. Indices of pileup, average stress reactivity, and average stressor recovery were computed for each participant across different operationalizations.

We used hierarchical multiple regression analyses to examine the association of these operationalizations with symptom change, using each operationalisation as a predictor and symptoms (e.g., DASS-A scores at follow-up) as the outcome, while adjusting for age and baseline DASS subscale scores. Age was included as a covariate because preliminary analyses showed significant associations with several operationalizations, while sex was not included as. It showed no significant associations.

Additionally, we assessed the relationship between trait anxiety (STAI-T scores) and the operationalizations through regression analyses, also adjusting for age.

3 | Results

3.1 | Preliminary Analyses

Participants showed low-to-moderate levels of stress ($M = 3.58$, SD = 3.26), depression ($M = 1.93$, SD = 2.22), and anxiety ($M = 1.36$, SD = 1.86) symptoms at baseline. At follow-up, stress ($M = 4.84$, SD = 4.07), depression ($M = 2.59$, SD = 3.13) and anxiety ($M = 1.78$, SD = 2.58) symptoms had significantly increased (DASS-S: $t = -4.11$, $p < 0.001$; DASS-D: $t = -2.87$, $p = 0.005$; DASS-A; $t = -2.17$, $p = 0.032$). The overall mean EMA-A score was 17.67 (SD = 21.70) across all participants.

The data cleaning process involved 89,100 potential notifications (165 retained participants × 180 days × three notifications per day). Of these, participants responded to 68,554 notifications, resulting in a mean of 409 answered notifications per participant (SD = 90.40) and 131 (SD = 90.40) unanswered notifications per participant. The percentage of answered notifications per participant ranged from 15% to 98.7% with the majority of participants (75.7%) responding to more than 50% of the notifications. The overall response rate was 77%. A total of 1239 notifications (1.3%) were lost due to technical issues, such as participants changing time zones or smartphone software updates.

3.1.1 | Stressors and Stress Reactivity

In our sample, all participants reported at least one event from the LHC list during the study period. Supporting Information S1: Table S1 presents the number and percentage of participants who reported each event category included in the LHC. There was considerable variability in the total number of events reported, with a mean of 21.84 (SD = 15.02; range = 1–66; see Supporting Information S1: Figure S2). The choice of stress reactivity baseline significantly influenced both the total number of stressor episodes across the entire sample ($n = 239$ with local baseline 1, $n = 176$ with local baseline 2, and $n = 228$ with cumulative baseline; see Supporting Information S1: Figures S3–S5) and the number of stressor episodes per individual. Specifically, for local baseline 1, the mean was 1.45 (SD = 0.74, range = 1–4), for local baseline 2, the mean was 1.07 (SD = 0.25, range = 1–2), and for the cumulative baseline, the mean was 1.41 (SD = 0.78, range = 1–6).

3.2 | Pileup

Different averages were obtained when pileup was defined as the number of stressor episodes versus the number of weeks with stressor episodes (using number of stressor episodes, $M = 2.47$, SD = 1.47; using number of weeks with stressor episodes, $M = 13.02$, SD = 6.55). Applying the three stress reactivity baselines to these two operationalizations yielded six different measures of pileup. Of these, two operationalizations showed a significant association with symptom change in our multiple regression analyses (Table 1). Specifically, pileup operationalised as the number of stressor episodes with a cumulative baseline was significantly associated with stress scores (DASS-S) at follow-up, explaining an additional 2.4% of variance beyond age and stress symptoms at baseline (F (3, 160) = 16.54, $p < 0.001$). Similarly, pileup operationalised as the number of weeks with stressor episodes and the cumulative baseline was significantly associated with stress scores at follow-up, accounting for 2.4% of additional variance beyond age and stress symptoms at baseline (F (3, 160) = 16.56, $p < 0.001$). Additionally, trait anxiety was significantly associated with pileup operationalised as the number of weeks with stressor episodes and local baseline 2, $r = 0.17$, $p = 0.03$.

In summary, we confirmed the expected association between stressor exposure (pileup) and symptom increase by defining a stressor as an event with increased stress reactivity and operationalising pileup using local (1 week) or cumulative baselines. Other operationalizations of pileup did not support this association. Only one operationalisation of pileup, which used the number of weeks with stressors, supported the expected association of trait anxiety with pileup.

3.3 | Average Stress Reactivity

Both operationalizations of average stress reactivity - using local baseline 1 (F (3, 100) = 9.31, $p < 0.001$) and the cumulative baseline (F (3, 92) = 7.14, $p < 0.001$)- were significantly associated with stress scores (DASS-S) at follow-up, accounting for 11% and 9% of the additional variance beyond age and stress scores at baseline (Table 2). In addition, two operationalizations of average stress reactivity were associated with anxiety scores (DASS-A) at follow-up: the operationalisation using local baseline 1 (F (3, 100) = 15.97, $p < 0.001$), which accounted for 14% of the additional variance, and the operationalisation using the cumulative baseline (F (3, 92) = 12.098, $p < 0.001$), which accounted for 11% of the additional variance. Trait anxiety was not associated with any of our indices of average stress reactivity.

In summary, two operationalizations of average stress reactivity were found to be associated with symptom increase, while none supported its association with trait anxiety.

3.4 | Average Stress Recovery

Among the six operationalizations of average stress recovery, only one was significantly associated with stress (DASS-S) scores (Table 3). The operationalisation using cumulative recovery and the local baseline 2 significantly predicted stress symptoms at follow-up. It explained 9% of additional variance over and above age and stress symptoms at baseline (F (3, 37) = 3.53, $p = 0.02$). This association was negative, indicating that lower average recovery was related to an increase in stress. Trait anxiety was not significantly associated with any of our operationalizations of average stress recovery.

4 | Discussion

In a longitudinal study with a relatively large sample, we tested the feasibility of combining EMA and LHC methods to assess

TABLE 1 | Predicting symptom scores at follow-up from different operationalizations of pileup, controlling for age and symptoms at baseline. Results from multiple regression analyses. Only significant results ($p < 0.05$) are shown.

Variable	*B*	95% CI LL	95% CI UL	SE *B*	*β*	R^2	ΔR^2
DASS-S at follow-up ($n = 164$)							
Step 1						0.21	0.21***
Constant	5.20**	2.06	8.34	1.59			
Age	−0.09	−0.22	0.03	0.06	−0.11		
DASS-S at baseline	0.56***	0.39	0.73	0.09	0.45***		
Step 2						0.24	0.02*
Constant	3.63*	0.24	7.02	1.72			
Age	−0.08	−0.20	0.04	0.06	−0.09		
DASS-S at baseline	0.55***	0.38	0.72	0.09	0.44***		
Pileup episodes using cumulative baseline	0.85*	0.11	1.59	0.38	0.16*		
DASS-S at follow-up ($n = 164$)							
Step 1						0.21	0.21***
Constant	5.20**	2.06	8.34	1.59			
Age	−0.09	−0.22	0.03	0.06	−0.11		
DASS-S at baseline	0.56***	0.39	0.73	0.09	0.45***		
Step 2						0.24	0.02*
Constant	4.34**	1.14	7.53	1.61			
Age	−0.08	−0.20	0.04	0.01	−0.09		
DASS-S at baseline	0.54***	0.37	0.71	0.09	0.43***		
Pileup weeks using cumulative baseline	0.11*	0.01	0.20	0.05	0.16*		

Abbreviations: CI = confidence interval; DASS-S = depression, anxiety and stress scale, stress subscale; LL = lower limit; SE = estimated error; UL = upper limit.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

stress components and examined the effects of multiple operationalizations of these components. Our results showed that different operationalizations affected stressor counts, but only some aligned with expected associations with changes in mental health symptoms or trait anxiety. These findings have methodological, theoretical, and clinical implications.

Stressors were assessed using a LHC, and stress responses were evaluated through EMA. This method effectively distinguishes the assessment of potential stressors from the assessment of stress responses, thereby reducing reporting biases (e.g., individual differences in labelling events as stressors). The classification of events as stressors, based on their potential to trigger stress responses, required establishing a baseline. This, in return, influenced the frequency of stressors and subsequent estimates of stress pileup, underscoring the importance of selecting appropriate baselines for both stressors and stress components.

We operationalised stress components—pileup, stress reactivity, and stress recovery-following the framework proposed by Smyth et al. (2023). However, the expected theoretical assumptions from prior research were only confirmed for certain operationalizations (Kalisch et al. 2021; Shackman et al. 2016) and for specific symptoms - namely, stress and anxiety, but not depression. Notably, only those operationalizations of pileup utilising cumulative baselines were associated with increased mental health symptoms. Both local and cumulative baselines for stress reactivity correlated with increased mental health symptoms. For stress recovery, the most predictive operationalisation combined a cumulative recovery baseline with a local stress reactivity baseline. From a methodological and practical perspective, our findings suggest that capturing stress components in everyday life via EMA may require employing multiple baselines (Smyth et al. 2023), with cumulative (longer-term) baselines proving crucial for measuring pileup and recovery. The significant associations between cumulative baselines and mental health may be due to the relatively minor nature of most stressors. Further research is needed to investigate this hypothesis.

While trait anxiety was positively associated with pileup (stressor exposure), there was no significant association between trait anxiety and average stress reactivity or stress recovery. Previous research has demonstrated an association between trait anxiety and elevated momentary negative affect (see Shackman et al. 2016), but most of these findings come from self-report or inter-individual approaches (Shackman et al. 2016), rather than EMA or intra-individual analyses (with some exceptions, such as Hur et al. 2022). Additionally, prior studies have used different

TABLE 2 | Predicting symptom scores at follow-up from different operationalizations of average stress reactivity, controlling for age and symptoms at baseline. Results from multiple regression analyses. Only significant results ($p < 0.05$) are shown.

		95% CI					
Variable	***B***	**LL**	**UL**	**SE *B***	***β***	***R^2***	***ΔR^2***
DASS-S at follow-up ($n = 104$)							
Step 1						0.11	0.11**
Constant	5.04*	0.92	9.159	2.07			
Age	−0.06	−0.23	0.102	0.08	−0.07		
DASS-S at baseline	0.41**	0.18	0.637	0.11	0.33**		
Step 2						0.22	0.11***
Constant	2.68	−1.41	6.768	2.06			
Age	−0.01	−0.17	0.154	0.08	−0.01		
DASS-S at baseline	0.35**	0.14	0.572	0.11	0.29**		
Stressor reactivity using local baseline 1	0.19***	0.09	0.298	0.05	0.33***		
DASS-S at follow-up ($n = 96$)							
Step 1						0.10	0.10**
Constant	5.36*	0.96	9.763	2.21			
Age	−0.07	−0.25	0.112	0.09	−0.07		
DASS-S at baseline	0.38**	0.14	0.620	0.12	0.31**		
Step 2						0.19	0.09**
Constant	3.20	−1.21	7.612	2.22			
Age	−0.02	−0.19	0.158	0.09	−0.02		
DASS-S at baseline	0.35**	0.12	0.580	0.11	0.29**		
Stressor reactivity using cumulative baseline	0.15**	0.06	0.244	0.05	0.30**		
DASS-A at follow-up ($n = 104$)							
Step 1						0.19	0.19***
Constant	2.49	0.13	4.85	1.19			
Age	−0.06	−0.16	0.03	0.05	−0.11		
DASS-A at baseline	0.61***	0.35	0.87	0.13	0.42***		
Step 2						0.32	0.14***
Constant	0.87	−1.41	3.14	1.19			
Age	−0.02	−0.11	0.06	0.04	−0.05		
DASS-A at baseline	0.57***	0.33	0.81	0.12	0.39***		
Stressor reactivity using local baseline 1	0.13***	0.07	0.19	0.03	0.38***		
DASS-A at follow-up ($n = 96$)							
Step 1						0.17	0.17***
Constant	2.48	−0.08	5.04	1.29			
Age	−0.06	−0.16	0.05	0.05	−0.10		
DASS-A at baseline	0.59***	0.31	0.86	0.14	0.40***		
Step 2						0.28	0.11***
Constant	1.01	−1.50	3.53	1.27			
Age	−0.02	−0.12	0.07	0.05	−0.04		
DASS-A at baseline	0.56***	0.31	0.82	0.13	0.38***		
Stressor reactivity using cumulative baseline	0.10***	0.048	0.15	0.03	0.34***		

Abbreviations: CI = confidence interval; DASS-A = depression, anxiety and stress scale, anxiety subscale; DASS-S = depression, anxiety and stress scale, stress subscale; LL = lower limit; SE = estimated error; UL = upper limit.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

TABLE 3 | Predicting symptom scores at follow-up from different operationalizations of stress recovery, controlling for age and symptoms at baseline. Results from multiple regression analyses. Only significant results ($p < 0.05$) are shown.

Variable	*B*	95% CI		SE *B*	*β*	R^2	ΔR^2
		LL	UL				
DASS-S at follow-up ($n = 41$)							
Step 1						0.13	0.13
Constant	6.88	0.15	13.62	3.33			
Age	−0.12	−0.39	0.15	0.13	−0.14		
DASS-S at baseline	0.31*	0.01	0.61	0.15	0.32*		
Step 2						0.22	0.09*
Constant	5.36	−1.26	11.97	3.26			
Age	−0.07	−0.34	0.19	0.13	−0.08		
DASS-S at baseline	0.32	0.03	0.6	0.14	0.33*		
Cumulative recovery using local baseline 2	−0.09	−0.17	−0.00	0.04	−0.31*		

Abbreviations: CI = confidence interval; DASS-S = depression, anxiety and stress scale, stress subscale; LL = lower limit; SE = estimated error; UL = upper limit.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

definitions of stress reactivity and recovery (e.g., De Calheiros Velozo et al. 2023; Weber et al. 2024).

From a theoretical standpoint, our findings underscore that stress is not a static construct, but rather comprises multiple dynamic components—namely, reactivity, recovery, and pileup—which should be conceptually distinguished and empirically measured as separate processes. This supports process-oriented models of stress (Lazarus and Folkman 1984; Almeida 2024) that emphasise the temporal unfolding and intra-individual variability of stress responses. Although the present study was not conducted in a clinical population, the findings have potential clinical relevance. Specifically, clinicians and health professionals may benefit from incorporating multiple baseline periods (e.g., weekly vs. cumulative) when assessing and monitoring stress in patients, as different baselines can yield distinct insights into an individual's stress dynamics. Furthermore, given the consistent association between stress pileup and increased mental health symptoms, targeted interventions may be warranted to prevent or mitigate the cumulative buildup of stress over time.

This study has several strengths, including its prospective longitudinal design, a relatively large sample size, and a 6-month study period that allowed for extensive data collection on stressor experiences and stress components. Notably, all participants experienced at least one stressor event during the study period, which is essential for detecting variability in stress responses (Smyth et al. 2023). However, the study also has several limitations. First, the sample was relatively young, and most stressors recorded were 'everyday' stressors, which may limit the generalisability of the findings to populations with different sociodemographic characteristics or exposure to other types of stressors, such as traumatic events. Second, self-reported anxiety was the sole measure of stress response, potentially overlooking other negative emotions like sadness and the study did not directly assess subjective stressor appraisal, an important aspect of the stress process. Third, the use of the LHC introduced temporal imprecision due to its weekly assessment of stressor presence and reliance on retrospective recall, which may be particularly susceptible to memory biases, especially for everyday stressors that are harder to recall accurately over extended periods. The LHC approach also excluded some potential definitions of stressors (e.g., 'perceived internal events': Smyth et al. 2023) and made it impossible to assess certain operationalizations of stress components (e.g., using 'proximal' baselines right before the stressor (Smyth et al. 2023)). Finally, we chose to use aggregated means primarily to facilitate integration with the retrospective LHC data and due to the exploratory nature of our analyses. We acknowledge the limitations of this approach, especially regarding the capture of within-person dynamics. Future studies employing statistical methods that account for these dynamics, such as multilevel models (Kim et al. 2018), are warranted.

5 | Conclusions

This study demonstrates the feasibility of integrating Life History Calendar methods with Ecological Momentary Assessment to assess stress and its components, reinforcing the value of this combined approach in capturing the dynamic nature of stress. Future research should further develop and refine operationalizations of stress components that are predictive of mental health outcomes and consider alternative statistical modelling techniques to enhance our understanding of stress dynamics over time.

Author Contributions

Authors C.S.-M., and M.A.F. designed the study and wrote the protocol. Authors V.D.l.P.-A., P.C.-E., A.J.-S., and I.M.-Z. conducted the experiment and collected the data. Authors M.A.F., and J.R. undertook the statistical analysis, and authors M.A.F., M.T.-F., and P.C.-E. wrote the first draft of the manuscript. L.Fo., V.D.l.P.-A., A.J.-S., I.M.-Z., E.Vi.,

J.Ra., A.J.Sh., and C.S.-M. revised the manuscript draft. All authors contributed to and have approved the final manuscript.

Acknowledgements

We thank CERCA Programme/Generalitat de Catalunya for institutional support.

Funding

This work was supported by Instituto de Salud Carlos III (ISCIII) [PI16/00889, PI16/00144, PI19/01171, PI19/00272], FEDER funds/European Regional Development Fund -a way to build Europe-, the Departament de Salut, Generalitat de Catalunya [PERIS SLT006/17/249], and the Agència de Gestió d'Ajuts Universitaris i de Recerca [2021-SGR-01017]. The study has also received funding from the European Union Horizon 2020 research and innovation programme under the Marie Sklowdowska-Curie [Grant 714673] and Fundación Bancaria 'la Caixa'. VDA was supported by 'la Caixa' Foundation (ID 100010434, fellowship code LCF/BQ/IN17/11620071) and by the Spanish Ministry of Science, Innovation and Universities, State Research Agency, and the European Union-NextGenerationEU/PRTR (JDC2022-048445-I). The Institute of Neurosciences of the University of Barcelona is a María de Maeztu Unit of Excellence CEX2021-001159-M of the Ministry of Science and Innovation of Spain. Dr. Shackman was partially supported by the U.S. National Institutes of Health (AA030042, MH131264, and MH121409) and University of Maryland.

Ethics Statement

Written informed consent was obtained from all participants before taking part in the study. The study procedures were carried out following the Declaration of Helsinki and received the approval of the Clinical Research Ethics Committee of Bellvitge University Hospital (reference PR144/16, February 8th, 2018).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study is available from the corresponding authors upon reasonable request.

References

Almeida, D. 2024. "The Time is Now for Integrative Multilayered Stress Research and Mitigation." *Frontiers in Science* 2: 1389481. https://doi.org/10.3389/fsci.2024.1389481.

Axinn, W. G., L. D. Pearce, and D. Ghimire. 1999. "Innovations in Life History Calendar Applications." *Social Science Research* 28, no. 3: 243–264. https://doi.org/10.1006/ssre.1998.0641.

Buela-Casal, G., A. Guillén-Riquelme, and C. N. Seisdedos. 2016. "Cuestionario De Ansiedad Estado-Rasgo." In *Adaptación Española*. 9th ed. TEA Ediciones.

Caspi, A., T. E. Moffitt, A. Thornton, et al. 1996. "The Life History Calendar: A Research and Clinical Assessment Method for Collecting Retrospective Event-History Data." *International Journal of Methods in Psychiatric Research* 6, no. 2: 101–114. https://doi.org/10.1002/(sici)1234-988x(199607)6:2<101::aid-mpr156>3.3.co;2-e.

Cohen, S., M. L. M. Murphy, and A. A. Prather. 2019. "Ten Surprising Facts About Stressful Life Events and Disease Risk." *Annual Review of Psychology* 70, no. 1: 577–597. https://doi.org/10.1146/annurev-psych-010418-102857.

Daza, P., D. M. Novy, M. A. Stanley, and P. Averill. 2002. "The Depression Anxiety Stress Scale-21: Spanish Translation and Validation With a Hispanic Sample." *Journal of Psychopathology and Behavioral Assessment* 24, no. 3: 195–205. https://doi.org/10.1023/a:1016014818163.

De Calheiros Velozo, J., T. Vaessen, G. Lafit, S. Claes, and I. Myin-Germeys. 2023. "Is Daily-Life Stress Reactivity a Measure of Stress Recovery? An Investigation of Laboratory and Daily-Life Stress." *Stress and Health* 39, no. 3: 638–650. https://doi.org/10.1002/smi.3213.

Ferrando, L., A. Franco, M. Soto, et al. 1998. MINI Entrevista Neuropsiquiátrica Internacional.

Fortea, L., M. Tortella-Feliu, A. Juaneda-Seguí, et al. 2023. "Development and Validation of a Smartphone-Based App for the Longitudinal Assessment of Anxiety in Daily Life." *Assessment* 30, no. 4: 959–968. https://doi.org/10.1177/10731911211065166.

Freedman, D., A. Thornton, D. Camburn, D. Alwin, and L. Young-DeMarco. 1988. "The Life History Calendar: A Technique for Collecting Retrospective Data." *Sociological Methodology* 18: 37. https://doi.org/10.2307/271044.

Guillén-Riquelme, A., and G. Buela-Casal. 2011. "[Psychometric Revision and Differential Item Functioning in the State Trait Anxiety Inventory (STAI)]." *Psicothema* 23, no. 3: 510–515. https://www.psicothema.com/pi?pii=3916.

Heshmati, S., C. Muth, R. W. Roeser, J. Smyth, H. Jamalabadi, and Z. Oravecz. 2024. "Conceptualizing Psychological Well-Being as a Dynamic Process: Implications for Research on Mobile Health Interventions." *Social and Personality Psychology Compass* 18, no. 1: e12933. https://doi.org/10.1111/spc3.12933.

Hur, J., M. Kuhn, S. E. Grogans, et al. 2022. "Anxiety-Related Frontocortical Activity is Associated With Dampened Stressor Reactivity in the Real World." *Psychological Science* 33, no. 6 (June): 906–924. https://doi.org/10.1177/09567976211056635.

Kalisch, R., G. Köber, H. Binder, et al. 2021. "The Frequent Stressor and Mental Health Monitoring-Paradigm: A Proposal for the Operationalization and Measurement of Resilience and the Identification of Resilience Processes in Longitudinal Observational Studies." *Frontiers in Psychology* 12: 710493. https://doi.org/10.3389/fpsyg.2021.710493.

Kim, J., D. Marcusson-Clavertz, F. Togo, and H. Park. 2018. "A Practical Guide to Analyzing Time-Varying Associations Between Physical Activity and Affect Using Multilevel Modeling." *Computational and Mathematical Methods in Medicine* 2018: 1–11. https://doi.org/10.1155/2018/8652034.

König, L. M., A. Allmeta, N. Christlein, M. Van Emmenis, and S. Sutton. 2022. "A Systematic Review and Meta-Analysis of Studies of Reactivity to Digital in-the-Moment Measurement of Health Behaviour." *Health Psychology Review* 16, no. 4: 551–575. https://doi.org/10.1080/17437199.2022.2047096.

Lazarus, R. S., and S. Folkman. 1984. *Stress, Appraisal, and Coping*. Springer Publishing Company Inc.

Lohani, M., S. Dutton, Z. E. Imel, and P. L. Hill. 2025. "Real-World Stress and Control: Integrating Ambulatory Physiological and Ecological Momentary Assessment Technologies to Explain Daily Wellbeing." *Frontiers in Psychology* 16: 1438422. https://doi.org/10.3389/fpsyg.2025.1438422.

Lovibond, P. F., and S. H. Lovibond. 1995. "The Structure of Negative Emotional States: Comparison of the Depression Anxiety Stress Scales (DASS) With the Beck Depression and Anxiety Inventories." *Behaviour Research and Therapy* 33, no. 3: 335–343. https://doi.org/10.1016/0005-7967(94)00075-u.

Mey, L. K., A. Chmitorz, K. Kurth, et al. 2020. "Increases of Negative Affect Following Daily Hassles Are Not Moderated by Neuroticism: An Ecological Momentary Assessment Study." *Stress and Health* 36, no. 5: 615–628. https://doi.org/10.1002/smi.2964.

Shackman, A. J., D. P. M. Tromp, M. D. Stockbridge, C. M. Kaplan, R. M. Tillman, and A. S. Fox. 2016. "Dispositional Negativity: An

Integrative Psychological and Neurobiological Perspective." *Psychological Bulletin* 142, no. 12: 1275–1314. https://doi.org/10.1037/bul0000073.

Sheehan, D. V., Y. Lecrubier, K. H. Sheehan, et al. 1998. "The Mini-International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-IV and ICD-10." Supplement, *Journal of Clinical Psychiatry* 59, no. S20: 22–33: quiz 34-57. https://www.psychiatrist.com/jcp/mini-international-neuropsychiatric-interview-mini/.

Shiffman, S., A. A. Stone, and M. R. Hufford. 2008. "Ecological Momentary Assessment." *Annual Review of Clinical Psychology* 4, no. 1: 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415.

Sliwinski, M. J., D. M. Almeida, J. Smyth, and R. S. Stawski. 2009. "Intraindividual Change and Variability in Daily Stress Processes: Findings From Two Measurement-Burst Diary Studies." *Psychology and Aging* 24, no. 4: 828–840. https://doi.org/10.1037/a0017925.

Smyth, J. M., M. J. Sliwinski, M. J. Zawadzki, et al. 2018. "Everyday Stress Response Targets in the Science of Behavior Change." *Behaviour Research and Therapy* 101: 20–29. https://doi.org/10.1016/j.brat.2017.09.009.

Smyth, J. M., M. J. Zawadzki, D. Marcusson-Clavertz, et al. 2023. "Computing Components of Everyday Stress Responses: Exploring Conceptual Challenges and New Opportunities." *Perspectives on Psychological Science* 18, no. 1: 110–124. https://doi.org/10.1177/17456916221082108.

Spielberg, C. D., R. L. Gorsuch, and R. E. Lushene. 1970. *Manual for the State-Trait Inventory*. Consulting Psychological Press.

Weber, J., P. Angerer, and J. Apolinário-Hagen. 2022. "Physiological Reactions to Acute Stressors and Subjective Stress During Daily Life: A Systematic Review on Ecological Momentary Assessment (EMA) Studies." *PLoS One* 17, no. 7: e0271996. https://doi.org/10.1371/journal.pone.0271996.

Weber, J., M. Heming, J. Apolinário-Hagen, S. Liszio, and P. Angerer. 2024. "Comparison of the Perceived Stress Reactivity Scale With Physiological and Self-Reported Stress Responses During Ecological Momentary Assessment and During Participation in a Virtual Reality Version of the Trier Social Stress Test." *Biological Psychology* 186: 108762. https://doi.org/10.1016/j.biopsycho.2024.108762.

Zhang, Y., D. Li, X. Li, X. Zhou, and G. Newman. 2024. "The Integration of Geographic Methods and Ecological Momentary Assessment in Public Health Research: A Systematic Review of Methods and Applications." *Social Science & Medicine* 354: 117075. https://doi.org/10.1016/j.socscimed.2024.117075.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information S1: smi70128-sup-0001-suppl-data.docx.