Check for updates

Original Research Article

Measuring Continuous Affect in Daily Life With Intensity Profile Drawings

Assessment
2025, Vol. 32(5) 689–704
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10731911241266286
journals.sagepub.com/home/asm
Sage

Leonie Cloos[1], Merijn Mestdagh[1], Wolf Vanpaemel[1], Eva Ceulemans[1], and Peter Kuppens[1]

Abstract

We examined continuous affect drawings as innovative measure of affective experiences over time. Intensive longitudinal data often rely on discrete assessments, containing "blind spots" between measurements. With continuous affect drawings participants visually depict their affect fluctuations between assessments. In an experience sampling study, participants (N = 115) rated their momentary positive and negative affect 6 times daily. From the second daily rating on, they additionally drew their positive and negative affect changes and reported affective events between assessments. They received one measurement burst between assessments daily. The strength of the approach is a substantial amount of informational gain (average 7%) over linearly interpolated points between assessments. The additional information was subsequently categorized into positive and negative affect peaks and valleys, each occurring once a day per person on average. The probability of detecting peaks and valleys increased with reported events. The drawings correlated positively with momentary affect scores from the burst. Yet, the drawing predicted the bursts less well suggesting that the momentary ratings may yield different information than the drawings. Although the timing of retrospective drawings is less precise than individual momentary assessments, this method provides a comprehensive understanding of affective experiences between assessments, offering a unique perspective on affect dynamics.

Keywords

affect intensity profile, continuous data, intensive longitudinal research, timescale, temporal process

Essentially every moment of our lives involves some sort of affective component. Most of time affect is an unconscious experience of feeling good or bad, excited or relaxed without knowing a salient cause (Russell, 2003). However, salient events like the anticipation of a deadline can elicit emotional episodes and change our affect (Russell, 2005). Affective experiences are not stable, but inherently dynamic phenomena, resulting from the constant interaction of affect with our internal and external environment (Waugh & Kuppens, 2021). The field of *affect dynamics* focuses on changes in our affective experiences over time and contexts to better understand the mechanisms behind our affective system and its relationship with mental health and well-being (e.g., Brose et al., 2013; Solhan et al., 2009; Wright et al., 2015).

To investigate affect dynamics, researchers collect intensive longitudinal data (ILD), containing multiple datapoints of a single individual over time. One approach to collect ILD in participants' daily life are ecological momentary assessment (Stone & Shiffman, 1994) or experience sampling methods (Csikszentmihalyi & Larson, 1987).[1] These methods are suitable to obtain subjective ratings on affect, by repeatedly administering self-report questionnaires (on a mobile device) multiple times a day for several days or weeks (see Hall et al., 2021 for a review). By now, ILD studies have become well established in contemporary affect research (Fritz et al., 2023). Despite its popularity, there are challenges inherent to the current measurement methods that limit the data quality and quantity. One of the most pressing challenges pertains to the timescale on which affective changes can be measured. While the aim is to capture these changes in real-time, the constraints of discrete repeated measurements may not always align with the true timescale of affective fluctuations (Kuppens et al., 2022). Given that affect is constantly shifting and emotional episodes can unfold at unpredictable rates, critical moments of affective change may be missed between scheduled measurement

[1]Katholieke Universiteit Leuven, Belgium

Corresponding Author:
Leonie Cloos, Faculty of Psychology and Educational Sciences, Katholieke Universiteit Leuven, Tiensestraat 102, Leuven 3000, Belgium.
Email: leonie.cloos@kuleuven.be

intervals (Kuppens et al., 2010; Verduyn, Delvaux, et al., 2009).

The Blind Spot

> "If a tree falls in a lonely forest and no animal is nearby to hear it, does it make a sound?"—Mann and Twiss (1910)

Consider this famous philosophical thought experiment in the context of ILD studies: If a person experiences an affective event in between measurement occasions, did it happen? Certainly it did. We just have no record of it. If the measurements occur before and after the event, the intensity of affect at the time of the event itself will fall into a *blind spot* and the data will show two similar subsequent values with no indication of the perturbation in between. If the peak intensity of someone's affective experience is missed, the subsequent pattern of unfolding cannot be reconstructed from the data (Dejonckheere et al., 2019; Hopwood et al., 2022).

The sampling scheme plays a crucial role in determining the structure and depth of data and potentially mitigating the occurrence of these *blind spots*. It dictates how often assessments are made and at what intervals, impacting the timescale at which affective changes are observed. A common choice is to use *fixed-interval* schedules, where assessments are scheduled at regular intervals (every few hours) or at specific time points. An alternative approach involves *random-interval* schemes, where participants receive a set number of assessments at random times throughout the day or in semi-random intervals, specifying windows in which the assessments are sent randomly. These two methods are so-called signal-contingent methods. In contrast, *event-contingent* schedules rely on participants to initiate assessments upon certain events or experiences (e.g., a social interaction of at least 5 minutes; Himmelstein et al., 2019). This carries the risk that participants are not going to report intense events, because they are too occupied (e.g., during a stressful episode) resulting in fewer reported events and less data over the course of a study. In a similar vein, *context-aware* methods rely on context, like location or physiological changes to trigger the assessments (Hoemann et al., 2020; van Halem et al., 2020).

The most commonly used study design employs fixed or random-interval scheme, with assessments spaced throughout the day, assuming that a (sufficiently large) sample of random assessments provides a valid representation of individuals' affect dynamics (Shiffman et al., 2008). This widely used approach typically has assessment frequencies ranging from 1 to 10 times a day (Hall et al., 2021). For a study with 10 daily assessments within waking hours (e.g., work of our own; Cloos et al., 2023), the average *blind spot* between the assessments would be approximately 75 minutes. To remedy this blind spot problem, there are potential solutions. The most straightforward solution is to measure more frequently. For example, studies on affective instability suggest that short intervals of 15 to 30 minutes can capture this dynamic process more effectively (Ebner-Priemer & Sawitzki, 2007). Despite the technological advancements that allow for real-time assessments, there remains a limit to the number of measurements participants can handle in a day. This creates a trade-off between minimizing intervals between assessments and maintaining a manageable number of assessments to limit participant burden and ensure compliance (Eisele et al., 2022).

Another potential solution to this problem is to limit the assessments throughout the day but enhance them with additional (daily) logs to gain a more complete picture. Daily-diary methods have been widely used in studies involving patients with long-term trajectories, helping maintain study feasibility and participant retention (Schreuder et al., 2020). Of course, the discrete assessments can also be asked in a retrospective format, allowing participants to summarize the period between assessments into a single score. However, a discrete retrospective question, such as "How angry did you feel since the last beep?" suggests that respondents must subjectively summarize their experience, which could lead them to report the average level, the peak intensity, or the most recent instance of anger—leaving it up to the researchers to decide what respondents chose to represent (Leertouwer et al., 2022).

An alternative approach is the implementation of micro-measurement bursts, where short-term follow-up assessments are triggered when there is a significant deviation from a predefined boundary (Mestdagh & Dejonckheere, 2021). Following up on the event provides indication of how the affective episode unfolds (Kuppens & Verduyn, 2015). However, this approach poses a challenge due to the absence of established norms or indicators for determining appropriate boundaries for triggering the bursts and also increasing participant burden by introducing additional measurement occasions.

The goal of this study is to introduce a novel approach to collect uninterrupted, continuous data that provide insight into what happens between measurement occasions. This continuous data should capture how affect changes, and unfolds over the course of a day. By filling in the *blind spot*, we gain a complete picture of affect dynamic processes.

The Current Study

The approach introduced in this study builds upon a method that has been developed by Sonnemans and Frijda (1994), who asked participants to draw the course of their emotional intensity during a particular emotion episode over time. This method captures several components of subjective emotion intensity, such as the number of peaks and valleys, the peak intensity, and the duration, among others. Subsequent studies have shown that these so-called *intensity profile drawings* are a valid method to investigate the unfolding of emotional episodes (see for example Verduyn et al., 2012; Verduyn, Van Mechelen, et al., 2009). Particular studies have investigated the variability in anger intensity profiles and how the profiles relate to specific events and regulation strategies (Heylen et al., 2015). Others have linked differences in emotion intensity profiles to person specific characteristics like rumination and depression (Kalokerinos et al., 2017; Résibois et al., 2018), or to different patterns of neural activity (Résibois et al., 2017). One of the most striking elements in intensity profiles are peaks and valleys. Peaks and valleys reflect particularly high and low levels of affect and may indicate the occurrence of affective events.

So far, intensity profile drawings have only been used as a measure for single emotional episodes. This means that (a) the drawings captured only one single emotion (not affect), (b) the emotion was elicited by a clear stimulus or an emotional event, and (c) the emotion had ended at the time of the drawing. To implement intensity profile drawings as a continuous affect measure, it is crucial to recognize the distinction between measuring emotions and measuring affect (Ekkekakis, 2013). Emotions are characterized by their episodic nature, whereas affect persists continuously—serving as a component of emotions but also existing independently. For instance, experiencing happiness upon receiving a gift represents affect within the context of an emotion, while feeling happy without a specific stimulus illustrates affect without the additional reflective component (Russell & Feldman Barrett, 1999).

The aim of this study is to implement and investigate intensity profile drawings as continuous affect measures for ILD studies. To our knowledge, there has been no research using this method to study positive and negative affect intensity across time. Using m-path (Mestdagh et al., 2023), we implemented the intensity profile drawings as a repeated measure for ILD designs. Unlike the traditional approach, which captures specific emotional episodes the continuous approach aims to capture ongoing affective fluctuations. Because affect is characterized by its constant presence and varying intensity over time (Russell, 2005) there is no start or end point. The current method uses repeated measures to intermittently probe the intensity of affect and define the previous affect intensity and the current affect intensity as anchor points (Supplement Figure 1). Individuals then draw the fluctuations in their affect intensity between these points. With this method, we obtain data on the momentary affect intensity at discrete assessments, and on the dynamics between these assessments. Similar to intensity profile drawings, the data allow to track important information such as intensity peaks and valleys and intermediate course of affect that would otherwise remain hidden.

Research Questions

The first research question we investigated was how much information we gain from the drawings. To measure informational gain, we compared the *continuous drawing* to a *straight line* connecting the discrete datapoints along a timeline. We expect to find visual deviations between the two methods and a quantitative difference. The second research question aimed to determine interesting datapoints between assessments. We expect that the drawing will find the occurrence of additional peaks and valleys that would go unnoticed using discrete methods. The third research question investigated the overlap between the *continuous drawing* and a daily *measurement burst* in between two assessments. We expect that there should be a relationship between the momentary affect item asked every 20 minutes during a burst and the corresponding datapoint in the drawing. The fourth research question we evaluated was whether the previously extracted peaks and valleys in the *continuous drawings* correspond to reported occurrence of positive and negative events. We expect a relation between the number of peaks and valleys detected and the number of events reported within a *continuous drawing* interval. All these research questions are exploratory in nature, intended to investigate the usability and potential insights that can be gained from these methods.

Method

We report how we determined our sample size, all data exclusions, all manipulations, and all measures. The (de-identified) data is available upon request. The code for the formal analyses is available online: osf.io/p6f4q/?view_only=ee68ebb73d6d409888dcebea416763d1. The study was approved by the Social and Societal Ethics Committee—KU Leuven.

Power Analysis

Prior to the study, we conducted a power simulation to determine the number of participants needed to estimate

a small average within-person correlation ($\overline{r_i} = 0.1$) between the *drawing* and a measurement *burst*. In the current context we refer to *measurement bursts* as a method to collect momentary affect ratings more frequently within a condensed timeframe, by reducing the intervals between assessments. We generated 1.000 datasets with 100 participants. The *measurement bursts* were simulated with 10 notifications per day for 6 days, but with 50% missing data. For each person we simulated measurement bursts with 10 notification for 6 days but added 50% missingness to the data. This yielded 30 observations per person on average. For each participant, we sampled the burst data from a standard normal distribution, then generated continuous drawing data based on these bursts. Specifically, burst observations were used to create the *continuous drawing* data. To introduce variability, we applied a small average within-person correlation between the continuous drawing data and the burst data, with an average correlation of 0.1 ($SD = 0.2$), varying across individuals. Results showed that a sample of 100 participants would ensure 96% power ($\alpha = 0.05$). Eventually we adjusted our design to 7 days and six notifications per burst. This yielded of 30 burst observations per person on average, assuming a compliance of 70%.

Participants and Prescreening

We recruited participants via social media channels, mouth-to-mouth advertisements, and flyers that were distributed at public places (e.g., university health center, gym). The study was conducted in English so international and Belgian participants were invited. There were no restrictions for participation, except English language skills. First, interested participants completed a prescreening on the online survey platform Qualtrics, in which they completed a brief demographic survey (about gender, age, nationality, current residence), and the Perceived Stress Scale (PSS; Cohen et al., 1994). From a pool of 869 individuals (69.2% female, 1.4% other; $M_{age} = 23.4$, $SD_{age} = 7.5$) we sampled participants based on their perceived stress score. We stratified our sample in terms of stress levels, to obtain sufficient variation in daily life experiences and affect fluctuations. We computed the quartiles of the stress distribution in the sample and used these to assign participants to one of four groups: low ($0 \leqslant 13.5$), mild ($13.5 \leqslant 19$), moderate ($19 \leqslant 23$), and high ($23 \leqslant 40$) stress levels. From each group, we randomly selected an equal number of participants and contacted them to participate in our study. We repeated this random stratified selection in several waves to keep the balance between the groups equal. Among all participants that filled in the recruitment survey, we drew two participants that received a 25€ voucher for a Belgian online store.

A total of 118 individuals participated in the study. One participant dropped out right after the baseline questionnaires, and two participants were excluded because they provided no drawings. Our final sample consisted of 115 individuals (58.26% female, one other; $M_{age} = 24.24$, $SD_{age} = 6.68$, range = 17–53) with diverse backgrounds. Sixty-one were from Belgium (the country where the study was conducted), six participants were Italian and Turkish (respectively), the other 42 were from other countries. In the final sample, the four categories of PSS scores were about equally represented. Twenty-nine participants (25.21%) scored in the low rank ($PSS = 5 \leqslant 13.5$), 33 (28.70%) in the mild rank ($PSS = 13.5 \leqslant 19$), 25 (21.74%)in the moderate rank ($PSS = 19 \leqslant 23$), and 28 (24.35%) in the high rank ($PSS = 23 \leqslant 34$).

Measures

Screening Questionnaire

Perceived Stress Scale (PSS). The PSS (Cohen et al., 1994) comprises 10 items designed to assess individuals' perceived stress levels over the past month. For instance, one item queried, "In the last month, how often have you found that you could not cope with all the things that you had to do?." The items were rated on a 5-point Likert-type scale, with answers ranging from 0 (*Never*) to 4 (*Very Often*). Four of the items were reverse-worded and were recoded accordingly. The PSS was scored by summing the responses provided for all of the items, so that a higher total score indicated higher perceived stress levels.

Baseline Measures. Participants filled in several baseline questionnaires that are not relevant for the current paper. The full range of materials is in the Supplement (Questionnaires).

ILD Measures

Momentary Affect. At the start of each day, participants received the first notification including only the two items asking about their *momentary positive* and *negative affect*: "To what extent do you experience positive [negative] emotions right now?," based on the validation study by Cloos and colleagues (2023). The items were answered on a slider scale, ranging from 0 (*I don't experience positive [negative] emotions right now*) to 100 (*I experience very strong positive [negative] emotions*). At the remaining notifications the momentary affect items were the first items asked in the questionnaire.

Continuous Affect Drawing. From the second notification onwards, the participants were asked to provide

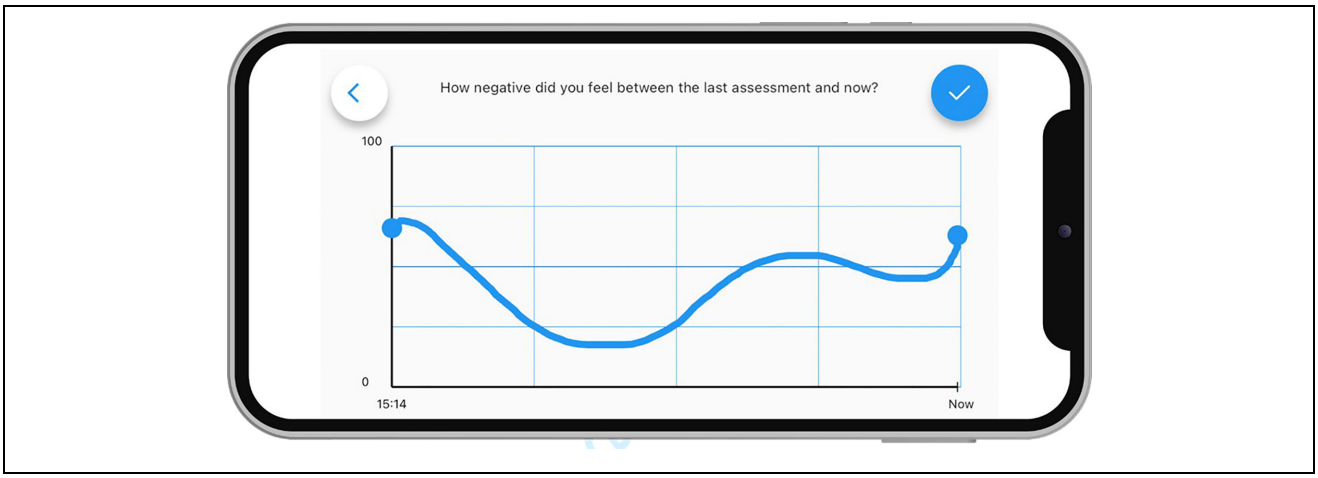


Figure 1. The Continuous Item as Currently Implemented in m-Path.
Note. A sample screenshot of a continuous affect drawing on the participants smartphone. The drawing was made by the researcher and is not taken from the participant data. The momentary affect rating from the previous assessment at 15:14 is indicated with the blue dot on the *y*-axis, and the momentary affect rating from the current assessment is indicated with the blue dot on the opposite end of the graph. The connecting line between the dots represents the fluctuation of negative affect between the previous assessment and the current one.

two continuous affect drawings for positive and negative affect, after rating their *momentary affect*. The *continuous drawing* items are presented as a two-dimensional figure with time of day on the *x*-axis and affect intensity on the *y*-axis, ranging from 0 to 100 (Figure 1). Above the figure participants read the question "*How positive [negative] did you feel between the last assessment and now*?." On the *x*-axis, a timeline was displayed starting at the time (displayed in hours and minutes of the day) the previous notification was answered and ending at the current timepoint (*now*). To answer the items, participants were asked to draw how their affect changed over this timespan. The *momentary affect* score from the previous notification was used as the starting point of the drawing and the current *momentary affect* score as the end point. Responses were only considered complete when both points were connected by the participant's drawing. A demonstration of the item can also be found directly on the m-path platform (www.m-path.io).

Each drawing was saved as a single concatenated string of comma-separated *x*—and *y*-values. We split and organized the values into a column for the *x*-values (*time*) and one for the associated *y*-values (*affect intensity*), ordering the values by *time*. To obtain a (*x* and *y*) pair per minute, we interpolated the affect values for minutes without (*x* and *y*) pair. For a detailed description of the preprocessing see Supplement code (OSF Analysis).

Events. Following the *continuous affect drawing item*, participants reported on the occurrence (*No* [0] *Yes* (1)) of positive and negative *events*: "*Did a positive[negative] event occur since the last assessment?*" If participants responded with *Yes*, they were directed to rate the intensity of the event on a scale ranging from 0 (*Neutral*) to 100 (*Very positive[negative]*). In addition, participants were asked to estimate the event's starting time using a categorical response scale with the options *just now* (1), *about 15 minutes ago* (2), *about 30 minutes ago* (3), *about 45 minutes ago* (4), *about 60 minutes ago* (5), *about 90 minutes ago* (6), or *about 120 minutes ago* (7).

Rumination. Three additional items to measure rumination were assessed but are not relevant for this study. The measure is described in the Supplement (OSF Questionnaires).

Procedure

Participants attended an online or an in-person introduction session where they received an explanation of the study. During this session, participants downloaded m-Path and were instructed what steps they needed to follow to complete their participation. After the session they received a personal code that anonymized their identity and was used to connect them to the m-Path account of this study. Once connected to m-Path, participants filled in the informed consent and baseline questionnaires. The experience sampling period of 7 days began the following day.

During each day of the sampling period, participants received six "fixed" notifications. With the exception of the first notification of the day, notifications did not have an expiration. The first notification of the day only comprised two items for *momentary affect*.

Table 1. Comparison of Distributional Properties Across All Datapoints.

	Momentary affect		Burst		Drawing	
Descriptive Statistic	Negative	Positive	Negative	Positive	Negative	Positive
Datapoints	4,430		3,527		533,526	
Mean	23.78	53.75	23.06	54.21	25.06	53.27
SD	20.91	25.67	20.35	26.24	20.6	25.1
Min	0	0	0	0	0	0
Median	19	57	18	58	20.29	55.09
Max	100	100	100	100	100	100

Note. These estimates are not directly comparable since the amount of datapoints differs. In the Supplement Table S1 we provide average means and standard deviations.

The first notification needed to be completed before noon, as it established the starting point for the continuous affect drawing. Failure to respond to the first notification on time resulted in not receiving subsequent notifications for the remainder of the day. Due to this requirement, a total of 32 days (3.98%) were missing, suggesting that the procedure was well-followed. This occurred across several participants: 24 participants missed 1 day of assessments, while four participants missed 2 days of assessments. Despite these missed assessments, the overall impact on data collection was minimal.

The five notifications presented the two *momentary affect* items along with the *continuous affect drawing* item and the two *event items*. These notifications were scheduled using a time contingent sampling scheme, with notifications scheduled every 140 minutes. The length of the interval between notifications determined the timeline on the *x*-axis of the drawing. The average timeline of a drawing was 146.21 minutes (SD = 37.42; Supplement Figure 2). During one randomly selected interval each day, participants received a so-called *measurement burst*, asking them to rate their *momentary affect* every 20-minutes. This means that the first burst beep was sent 20 minutes after the last fixed notification, the second 40 minutes after, and so forth. The maximum number of burst notifications was set to six.

At the end of the 7 days, participants were debriefed. Their compensation for participating was calculated based on how many notifications they answered. Participants who answered at least 75% of the notifications received the full reward of 40€. With every 5% decrease in responding, the reward decreased by 1.50€. For every day in which they missed the first notification, 4.50€ was deducted from the reward.

Data Analysis and Results

All analyses were conducted using R version 4.2.2 (2022). The dataset was processed and handled using functionalities of *dplyr*, *tidyverse*, *lubridate*, *hms*. The figures and tables were generated using *ggplot*, *ggpubr*, *summarytools*. The appropriate functions for the data processing and the analyses are available in the code online (OSF Analysis).

Descriptive Statistics

We investigated the basic properties of the different measurement techniques for affect applied in this study (Table 1). Across all timepoints the means, standard deviations (*SDs*), and ranges were similar across the assessment methods. The *drawing* yielded minute-to-minute data for a total of 9,024 hours and 51 minutes. The overall compliance rate among participants for the fixed questionnaire, which included the momentary question before noon and the questionnaire notification with the drawing-based component, was 91.72% (SD = 10.10, range = 52.38%–100%). The average compliance to the burst assessments was 70.31% (SD = 20.52; *range* = 9.52%–100%) with participants typically completing 4.80 out of 6 daily burst assessments (SD = 1.37; *range* = 1–6). These rates reflect a generally high level of participant engagement and consistent compliance across the various components of the study.

Research Question 1: Informational Gain of the Continuous Drawing Compared With a Straight Line

The first research question addresses the informational gain the drawing provides over the discrete assessments. Most studies implicitly assume that discrete datapoints are linearly related to each other and visualize them by connecting them with line graphs. We evaluate how much information the *continuous drawing* adds to this interpolated straight line. First, we created a new variable with the datapoints of the *straight line* by linearly interpolating the values between two fixed notifications, again using a 1-minute scale. Then, we computed the

absolute difference between the affect values of the two methods per minute, to quantify the informational gain of the *continuous drawing* over the *straight line*.

To illustrate the informational gain, Figure 2 shows the *continuous drawings* (curved lines) and the interpolated *straight line* for eight individuals with high and low levels of informational gain. The gray area between the *straight line* and *drawings* represents the informational gain per drawing. The drawings in Panels A and C contain large gray areas, due to large deviations from the straight lines, indicating a substantial gain in information during these periods. For participants with low informational gain, the gray surface is very small. In each of the drawings, there is an average informational gain noted in the labels of Figure 2 to obtain an understanding of what an average value of 6.88, or values higher and lower, look like.

Panel A of Figure 3 shows the distribution of the average information that is gained per minute in all of the *continuous drawings* (total = 3,657). The average informational gain per minute was 6.62 points (on a 0–100 scale) for negative affect intensity and 6.88 points for positive affect. The maximum average informational gain recorded in a drawing was 100 points, suggesting that the drawing depicted the opposite information of the straight line. In positive affect drawings, the maximum difference was 57.23 points. Panel B shows the distribution of the average informational gain per minute across the 115 participants showing considerable variation across participants in terms of the additional information provided.

Research Question 2: Identification of Peaks and Valleys Between Assessments

Peaks and valleys in the *continuous drawing* may provide important information about otherwise missed affective episodes. We investigated the occurrence of peak and valleys in each drawing using the *peaks* function in R (package *splus2*; Constantine & Hesterberg, 2021). The function automatically detects local maxima in a time-series, by searching for elements in a sequence that are greater than all other elements within a window. To determine input parameters for this function, we asked three raters to visually identify the peaks and valleys in 20 individuals' drawings. Based on the raters' evaluations, we established the optimal span for the moving window that the local maxima function processes. We set the *span* parameter to 31 minutes, assuring that a peak or valley is at least 15 minutes away from another peak or valley (respectively). A smaller window would lead to overidentification of local maxima, potentially treating a single peak and a nearby secondary maximum as separate peaks. The use of a 31-minute span ensured that such closely spaced maxima are not misclassified as separate peaks. Sensitivity analyses showed that using a span of 21 or 41 minutes did not significantly alter the result (Supplement Table S3). We set the *endbehavior* parameter to 0, ensuring that a peak or valley occurred at least 15 minutes before or after the start of the drawing. This reduces the likelihood of false peaks or valleys due to misplacements or accidental screen touches at the beginning of the drawing process, offering a more accurate analysis of the data. A wrapper function that included additional criteria so that the obtained peaks and valleys achieved maximum correspondence with the visual inspection. To refine the general function, we added additional parameters to increase correspondence with the visual ratings. This included a parameter to detect plateaus and restricted the plateau to last shorter than the length of the drawing minus 30 minutes. It further included parameters to verify the peaks and valleys. These compared the peak's or valley's intensity to the intensity of to all points up to 15 minutes before and 15 minutes after, and verified that the intensity difference between all points around the peak or valley had was least 10% of the individual's intensity range (e.g., if an individual's minimum intensity was 18, maximum intensity was 76, the range would be 58, and the difference would be 58/10 = 5.8). These criteria were eventually used to identify peaks and valleys in all continuous drawing data, which were coded into four binary variables for negative peak, positive peak, negative valley and positive valley (*presence* = 1, absence = 0).

On average we identified 7.25 peaks and valleys per person. The most prominent feature observed in the data was the occurrence of negative affect peaks (red bars above). The feature detection function identified on average 1.38 negative affect peaks (SD = 1.36; *range* = 0 – 8), 1.03 positive affect peaks (SD = 1.23; *range* = 0–7), 0.80 negative affect valleys (SD = 1.16; *range* = 0–7) and 1.10 positive affect valleys (SD = 1.23; *range* = 0–7) a day per person.

However, as shown in Figure 4, the sum of peaks and valleys within-persons varied across persons. This indicates that for some individuals, the drawing did not reveal any detectable peaks or valleys, while for others, there was substantial added value in utilizing the drawing. The distribution of negative valleys (light red bars below the horizontal axis) is skewed toward 0, indicating fewer occurrences, while positive valleys (light blue bars below the horizontal axis) are more evenly distributed at lower counts, except for one person who drew 27 positive affect valleys.

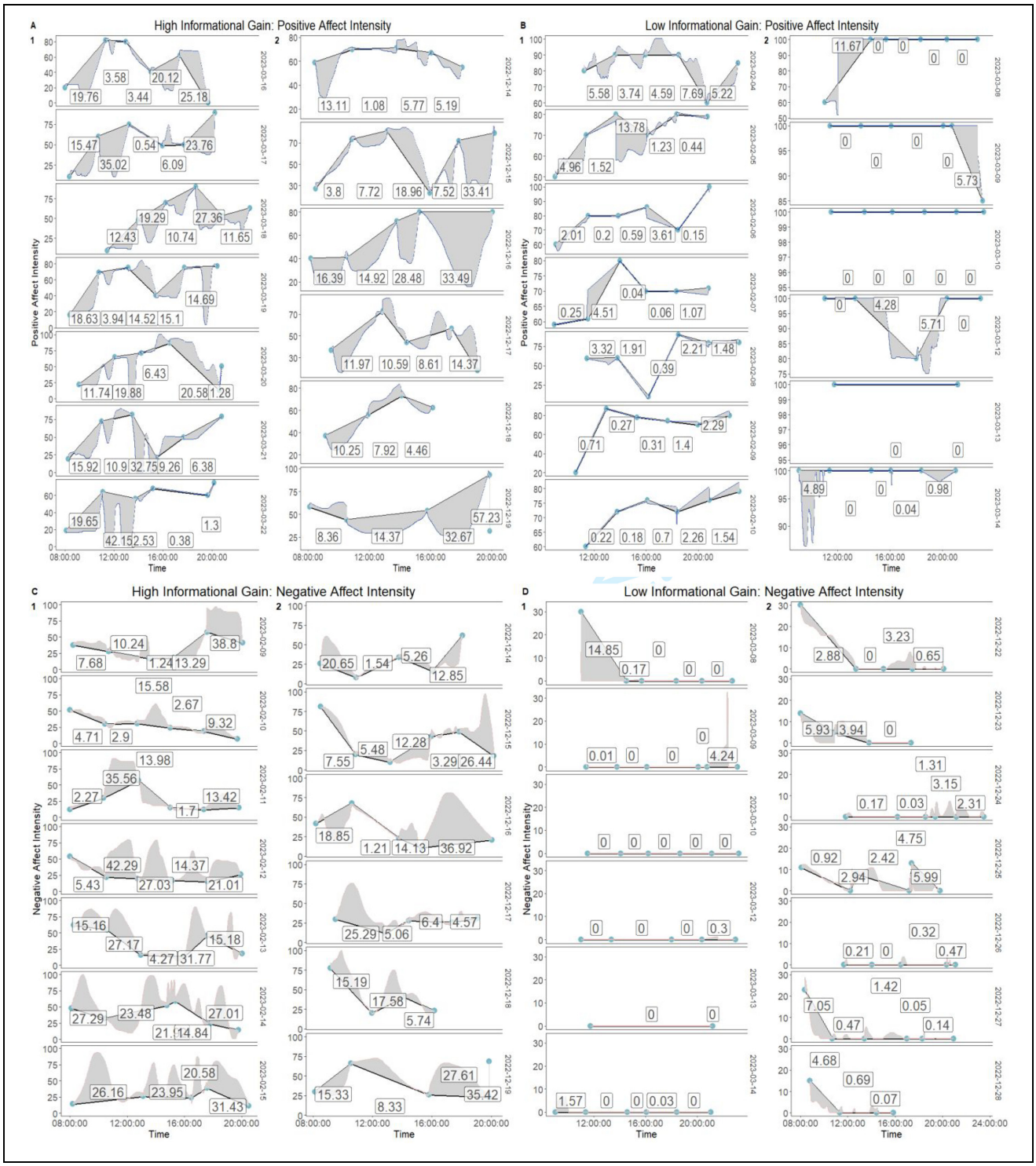


Figure 2. The Curved Lines Depict the Participants' Drawn Patterns, and the Straight Lines Show Interpolated Points Between Momentary Affect Ratings, Indicated as Blue Dots.

Note. The shaded area show the informational gain, measured as the absolute distance between the curved and straight lines. The labels contain the average informational gain for each drawing interval. The top row (A, B) shows positive affect drawings of four participants during the study week, while the bottom row (C, D) displays negative affect drawings of four participants over a week. The left column (A, C) shows participants drawings deviating from a straight line. The right column (B, D) shows participants aligning the drawings with a straight line.

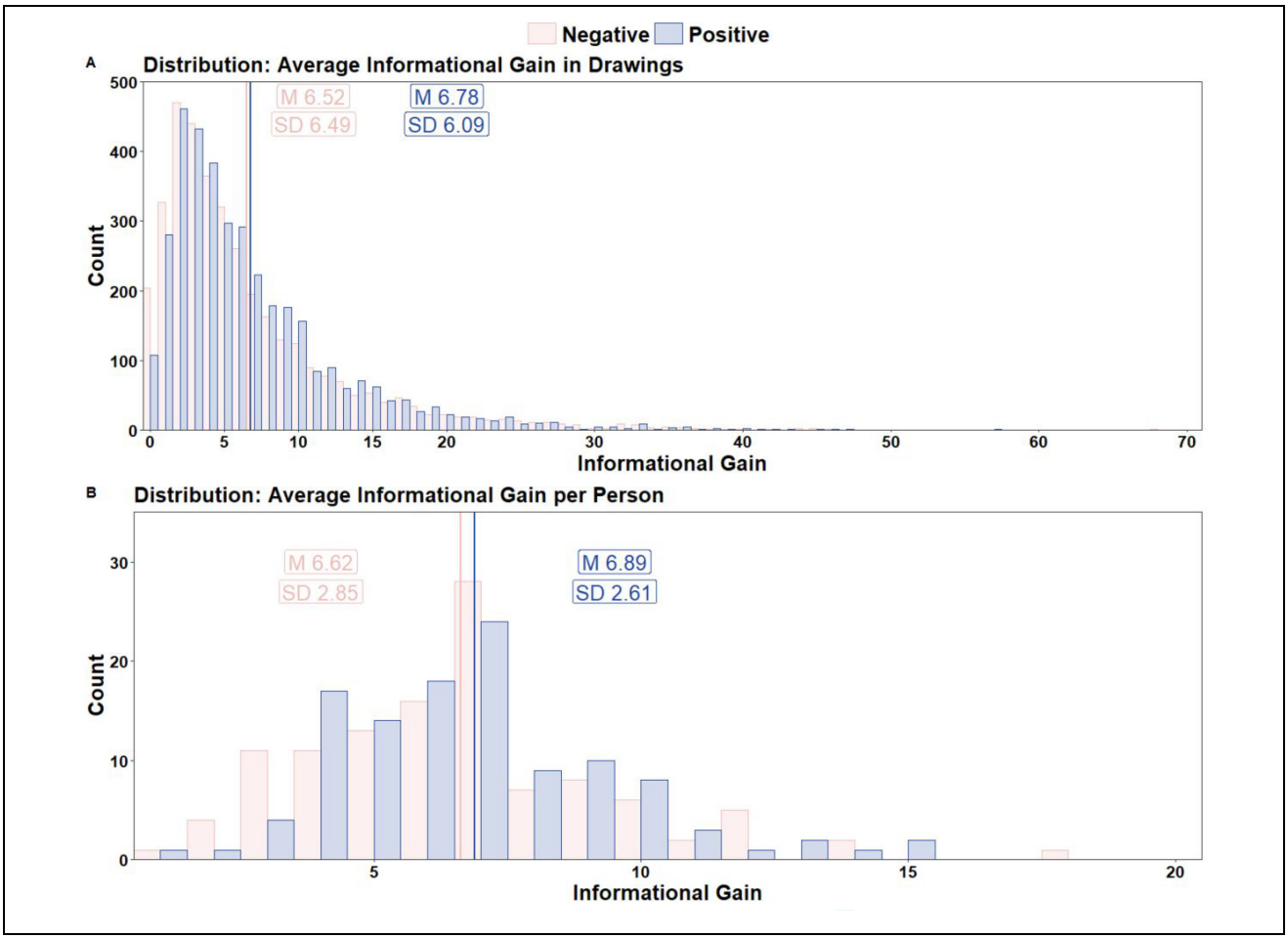


Figure 3. Distributions of Average Informational Gain Per Minute.

Research Question 3: Agreement of Measurement Burst With Continuous Drawing Versus the Straight Line

For this question, we compared the data from the *measurement bursts* to the data of the *continuous drawing* and the interpolated *straight line*. On average, individuals answered 29.53 out of 42 *measurement burst* assessments ($SD = 8.62$; *range* = 4–42) across the whole study.

Since there were only up to six measurement *burst* assessments a day, we selected the corresponding (i.e., at the same minute) affect scores of the *continuous drawing* or *straight line* (see Table 1). Next, we day-mean centered the resulting data for each person. The processing and analyses were conducted using the R packages *dplyr* (Wickham et al., 2022), *nlme* (Pinheiro et al., 2023), and *correlation* (Makowski et al., 2022).

Using these day-mean centered data, we first investigated whether the *drawing* and the *line* differed in their distance to the *bursts*. We computed the absolute difference between the *burst* and the *drawing* scores and between the *burst* and the *line* scores. In Table 2, we report the means of the absolute difference values, showing that the average absolute difference between the burst and the drawing is larger than the average absolute difference between the burst and the line. We tested whether this effect is significant by using the difference of both absolute differences as the dependent variable in an intercept only multilevel model, where timepoints are nested within individuals. We allowed the intercept to vary across individuals and accounted for the serial dependence of the data by including an autocorrelated (*AR1*) within-person error process. The results showed a significant effect for the intercept indicating that the absolute differences significantly differed from each other by $-.80$ points, $t(3281) = -6.346$; $p < .001$, on average, for negative affect and by -0.83 points, $t(3281) = -6.937$; $p < .001$, for positive affect. This indicates that the drawing is further away from the burst than the line is from the burst.

Second, to confirm this conclusion, we estimated the multilevel correlation between the *burst* and the *drawing*

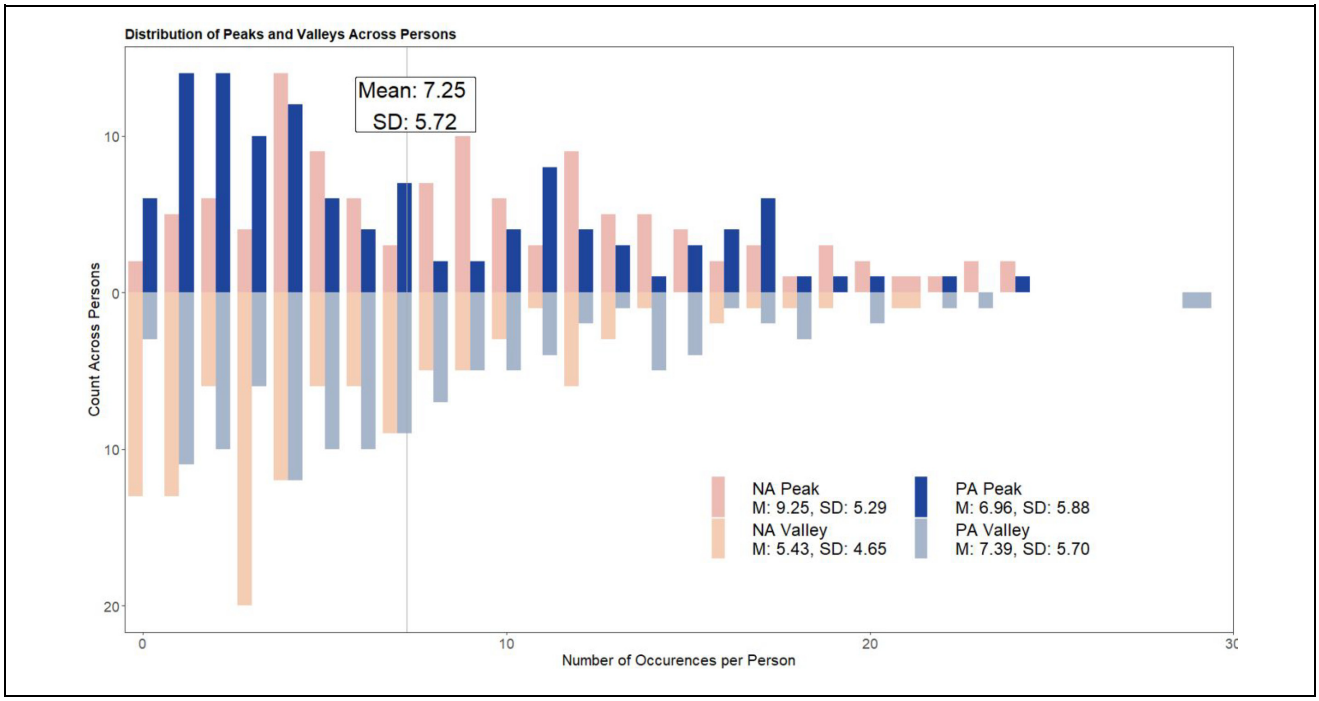


Figure 4. Distribution of the Number of Peaks (Above Horizontal Line) and Valleys (Below Horizontal Line) Detected Within Individuals.

Table 2. Absolute Differences, Correlations, Regression Coefficients.

	Negative affect		Positive affect	
Estimate	Burst-drawing	Burst-line	Burst-drawing	Burst-line
Mean Absolute Difference	7.97	7.16	8.49	7.64
SD	8.88	8.27	8.66	8.11
Within-Person Correlation	0.20	0.27	0.20	0.29
95% CI	[0.17, 0.24]	[0.23, 0.30]	[0.17, 0.24]	[0.26, 0.32]
	$\beta_{Drawing}$	β_{Line}	$\beta_{Drawing}$	β_{Line}
Regression Coefficient	.136	. 457	.117	.531
95% CI	[.079, .192]	[.381, .533]	[.058, .176]	[.455, .607]

Note. M: Mean Absolute Difference, SD: Standard Deviation of Mean Absolute Difference.

as well as the *burst* and the *straight line* datapoints respectively, shedding light on the average within-person correlation. From Table 2, we conclude that these correlations are positive and statistically significant at the $\alpha < .001$ level, ranging between .20 and .29. For both positive and negative affects, the burst-line correlation is larger than the burst-drawing correlation, although in the case of negative affect their 95% confidence intervals (CIs) overlap.

Finally, we estimated multilevel model with the *burst* scores as the outcome and the *drawing* and the *line* as predictors, using the same nesting structure and autocorrelated error process. Our effect of interest was the fixed effect of the *drawing* on the *burst* accounting for the effect of *line*, to uncover whether the *drawing* predicted the *burst* data on top of *linear* interpolated points. The results in Table 2 show that both for positive and negative affect, the fixed effect of the *drawing* was significant. However, the drawing did not seem to contain too much unique information about the burst scores that is not contained in the line scores: the *drawing* only added very little explained variance to the model: 0.6% for negative and 0.3% for positive affect.

Research Question 4: Co-Occurrence of Peaks, Valleys and Events

Along with the continuous drawings participants reported whether a positive and negative *event* occurred

Table 3. Contingency Table for Negative Events and Negative Episodes Within Assessments.

		Negative Event		
		No	Yes	Total
Negative Episode	No	2,051 56.21%	218 5.97%	2,269
	Yes	1,138 31.19%	242 6.63%	1,380
Total		3,189	460	3,649

Note. Negative Episode = a negative peak or a positive valley.

Table 4. Fixed Effects of Negative Event on Negative Episodes.

Estimates	Coefficient (log odds)	*P*	95% CI
β_0	−0.694	<.001	[−0.878, −0.510]
$\beta_{Negative\ Event}$	0.928	<.001	[0.709, 1.150]

Note. β_0: Intercept, representing log-odds of the outcome for no Negative Events. $\beta_{Negative\ Event}$: Regression, representing change in log-odds for presence of Negative Events.

since the last assessment. We validated the *peak* and *valley* detection within a drawing by relating them to the reported *positive* and *negative events* in the same assessment, coded in binary variables. Participants reported 0.60 negative events ($SD = 0.86$: *range* = 0–4) and 1.15 positive events ($SD = 1.17$; *range* = 0–5) per day on average (Supplement Table S5). The previously obtained binary variables for negative peaks and positive valleys were combined by the "or" operator to indicate *negative episodes*, likewise the positive peaks and negative valleys were combined into one variable for *positive episodes*. To summarize the relationship between the occurrence of *positive events* and *positive episodes* on one hand and between *negative events* and *negative episodes* on the other hand, we constructed a 2×2 contingency table, displaying the frequencies of all combinations of levels within the variables. The contingency table (Table 3) shows that the occurrence of negative events and negative episodes aligned for 6.63% of the assessments. While this might initially seem low, it is important to consider that a large portion of intervals (3,184) have no reported events or episodes. In turn, the 64.31% of the assessments where no event was reported, also no peaks or valleys were present. However, among the 460 intervals where a negative event *was* reported, 242 (52.61%) correspond to a negative episode detected in the drawings.

Second, we tested the strength of these associations between *episodes* and *events*, by conducting a mixed effect logistic regression in R using the *MASS* package. The models took the nested nature (assessments nested into individuals) into account and included an *AR1* error structure. The intercept shows that when no negative event was reported the probability of detecting a negative episode was about one third. The coefficient $\beta_{Negative\ Event}$ indicates that the log-odds for detecting a negative episode in the drawings significantly increases when a negative event was reported (Table 4). Specifically, negative events increased the probability of detecting negative episodes to approximately 71.69%.

For the positive event association, we found that in 8.20% of the assessments, both a positive event occurred and a positive episode was detected (Table 5). Positive episodes were only detected in 33.71% of assessments where a positive event was reported. However, in assessments with no event, 71.80% of those drawings did not have a positive episode. For positive episodes, we found that the baseline log-odds of finding one when there is no *positive event* is significantly below zero (Table 6). The intercept β_0 indicates that when no events are reported the probability of *positive episodes* is 24.78%. The coefficient ($\beta_{Positive\ Event}$) means that *positive events* only slightly increase the probability of detecting *negative episodes* to 59.41%.

Discussion

Affect intensity profile drawings offer as a novel means of capturing the dynamics of affect over time. The resulting data potentially contain rich information on affective changes in daily life, but need to be evaluated in terms of how much and what type of information they convey. Our study indicates that while the method shows promise, capturing unique information, there areas that still need further improvement.

Table 5. Contingency Table for Positive Events and Positive Episodes Within Assessments.

		Positive Event		
		No	Yes	Total
Positive Episode	No	1,983 54.34%	588 16.11%	2,571
	Yes	779 21.35%	299 8.20%	1,078
	Total	2,762	887	3,649

Note. Positive Episode = a positive peak or a negative valley.

Table 6. Fixed Effects of Positive Event on Positive Episodes.

Estimates	Coefficient (log odds)	*P*	95% CI
β_0	−1.110	<.001	[−1.310, −.912]
$\beta_{Positive\ Event}$	.381	<.001	[.203, .558]

Note. β_0: Intercept, representing log-odds of the outcome for no Positive Events. $\beta_{Positive\ Event}$: Regression, representing change in log-odds for presence of Positive Events.

First, we showed that the drawings allow to collect additional and unique information that cannot be captured with discretely spaced momentary assessment. On average, the drawings yield an information gain of 7 points on a scale of 100, which is a meaningful difference, considering that most of our daily experiences are rather mundane and uneventful (Schwarz, 1994). Second, we demonstrated how to detect affective episodes with an automatic function that aimed to identify peaks and valleys that are visually prominent. Our findings showed that on average, one positive and one negative episode occur per day occur that may have gone unnoticed with the discrete affect measure. Third, we found evidence that the affective episodes found in the drawings coincided with reported positive or negative events at above chance level. However, the overlap between the self-reported negative events and the detection of negative episodes presents a particularly noteworthy improvement, considering that negative affect is typically very low in general population samples. With the continuous drawing we can identify important negative affect episodes that can clarify our understanding of affect dynamics. However, there are still some observed discrepancies between the affect episodes and the self-reported events. For negative affect there were barely any events reported, while a notable number of negative episodes were found. This may indicate that episodes captured by continuous drawing reflect a broader or different set of experiences than those explicitly reported with a self-report item. Alternatively, it may indicate that the peak detection method might be overly sensitive, identifying too many additional episodes.

For positive affect the mapping of the detected episodes and the reported events was lower. The number of reported events did not increase the number of episodes, yet the relation between events and the episodes was still significant. A hypothesis to explain this is that there were less positive episodes than negative episodes detected across people, while positive events were more frequently reported than negative events. Thus, there is an imbalance in episodes detected for positive and negative affect, where positive changes are better detected by an event measure while the drawing method is better to investigate the presence of negative affective episodes that are not captured by the event item.

The attempt to validate the drawings with measurement bursts that were administered between assessments once a day showed mixed evidence. Primarily, we found a relatively low positive correlation between the momentary bursts data and the corresponding scores of the of affect drawing at the corresponding timepoints. This suggest that the overlap between drawings and bursts may pertain to global trends (increases/decreases) in affect but not to more subtle changes. However, the findings on the absolute differences of the drawing to the burst data compared with the difference of the line to burst data warrant caution. First, the findings seem to suggest that people draw trajectories that are more extreme and more outspoken compared with the discrete momentary burst data. The average mean and standard deviation within drawings is larger than within bursts (taking into account that the estimates of the burst may be inflated because of the lower number of datapoints). Second, while the drawings do pick up the presence and intensity of affect changes, they are certainly not able to

pinpoint the timing of when this occurs as precisely as a momentary measure. Third, when looking at the individual drawings there is considerable variation between persons as to how much extra information is added. There are individuals that consistently draw a lot of added information versus individuals that rarely have added information. Finally, even for these individuals with generally limited informational gain, certain drawings stand out with additional insight.

Limitations

It is important to be aware of the strength and weaknesses of our current approach. Especially a newly developed method has to be evaluated in face of its limitations. The primary limitation uncovered in the current study is that the minute-to-minute data points of the continuous drawing item do not align precisely with the timing of a momentary burst assessment. Realistically, it would be absurd to assume that a retrospective measure asking about the evolution of affect since the last assessment can be precise to the minute. Nonetheless, this is important to keep in mind when analyzing the data and drawing expectations about its relation to other measures: for instance, relating minute-to-minute data from this approach to continuous physiological data does not seem warranted at this point.

A second limitation is the retrospective nature of the item. The continuous item is essentially in the format of an "Ecological Retrospective Assessment" (Leertouwer et al., 2021, p. 7). This means that there is a short time period in which participants have to use their short-term memory to evaluate the drawing. This introduces the potential for retrospective biases, such as peak-end effects, where participants might disproportionately focus on the most intense or final moments (Fredrickson, 2000; Kahneman et al., 1993). Moreover, the varying timeframes in which participants complete their drawings may introduce additional uncertainty. While most drawings are approximately 140 minutes long, we cannot definitively state how different durations, such as a 45-minute drawing versus a 2-hour drawing, might impact the results. Further research is needed to determine the extent of these biases and explore their influence on the collected data.

Third, there is inherent noise due to the constraints of smartphone-based data collection. This noise can arise from various factors, such as varying touch screen sensitivity, input range, orientation of the slider and so forth (for a detailed review see Van Berkel et al., 2022). These irregularities such movements of the finger of difference in pressure can threaten the reliability of the recorded data (Dejonckheere et al., 2022). Concerning the rating scale of affect intensity, there is a fixed boundary from 0 to 100, there may be instances where the gain in information is limited by these scale bounds. For example when the momentary ratings are at the upper or lower scale bounds, where the drawing cannot provide additional information above or below these points. This limitation of bounded scales is evident in Figure 2 (Panels B2 and D1), where ratings reach extremes and limit the drawing's ability to convey further information.

Similarly, the scale of the x-axis is limited in space by the screen. However, as mentioned above, the time interval on the x-axis differs across drawings, creating inconsistencies in how participants phone's represents the timescale. This noninvariance in the x-axis timescale may contribute to some difficulties with the timing accuracy of the data mentioned above.

Future Research

In this study, our emphasis has been on validating peaks in the data. It is important to note that the literature on intensity profile drawings has found other components, such as the shape, amplitude, and duration of affective episodes (Heylen et al., 2015; Verduyn, Van Mechelen, et al., 2009). These profiles need to be validated with additional methods to determine their shape and their changes in the shape over time (Heylen, Van Mechelen, Fried, & Ceulemans, 2016; Heylen, Van Mechelen, Verduyn, & Ceulemans, 2016).

In addition, our method presents a unique opportunity to study the dynamic relationship between positive and negative affect. This could be especially revealing in the context of psychological disorders, where examining trends in correlations between positive and negative emotions over time could yield critical insights (Dejonckheere et al., 2018).

Another intriguing area for future research are the observed individual differences. For example, we may use continuous intensity profile drawings to investigate affect dynamics and their relation to state and trait levels of other affective constructs, such as rumination tendencies, emotion regulation skills, and reactivity to events. The extent to which this is possible with the current approach is warrants further examination. Similarly, personality traits like neuroticism and extraversion have been linked to different patterns in emotion duration and intensity (Verduyn, Delvaux, et al., 2009). Another factor that changes affect dynamic states, worth exploring are response to stressors (Vogelsmeier et al., 2023). A recent study found that the reaction to happy and distressing events recorded with a drawing relate to emotional reactivity at baseline (Veilleux et al., 2024).

Finally, examining the duration, intensity, and frequency of affect has implications for understanding

psychopathological conditions (Gross & Jazaieri, 2014). For example, mood disorders are characterized by extended periods of negative affect (Kuppens et al., 2007). This method could be instrumental in uncovering affect duration and contribute to a deeper understanding of psychological disorders.

Conclusion

The goal of this study was to implement a newly developed continuous affect item for daily life research and evaluate it empirically. The measure reveals a more comprehensive understanding of affective experiences between assessments without increasing participant burden. We found that there is a lot of added information regarding affect change and affective episodes that have been hidden so far. At this point the data can be used to detect affect episodes between the normal assessment points. Furthermore, the fluctuations provide insight into the precursors and outcomes of affective changes, that may help the investigation of affect appraisal and regulation. Yet, the timing of the retrospective assessments may be less accurate than a time-point specific momentary assessment. It is necessary to clarify these limitations, before using the method to investigate affect in different contexts of daily life. This study underscores the importance of expanding our data collection methods to capture the full temporal spectrum of affective experiences.

Acknowledgments

Special thanks to Koen Niemeijer for the data curation and annotate the code for the preprocessing of the research data. For the help with the recruitment, prescreening, and conducting the online introductory sessions for the data collection, we thank the student assistants Rianne Wijnants, Figen Kirkgöz, and Miguel Ángel Segura Vargas.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported in this paper was supported by a Belgian Research Council grant. The second author is funded by a Postdoctoral Fellowship by the Flemish fund for Scientific Research (FWO).

ORCID iD

Leonie Cloos https://orcid.org/0000-0003-1550-7405

Data Availability

Materials and analysis code for this study are available at the OSF and can be accessed at osf.io/p6f4q/?view_only=ee68ebb73d6d409888dcebea416763d1. The data are available upon request.

Supplemental Material

Supplemental material for this article is available online at https://osf.io/p6f4q/.

Note

1. Although these terms are used interchangeably in practice and in the literature, we refer to ILD throughout this report, to emphasize the "intensive" nature of these procedures.

References

Brose, A., Lindenberger, U., & Schmiedek, F. (2013). Affective states contribute to trait reports of affective well-being. *Emotion*, *13*(5), 940–948. https://doi.org/10.1037/a0032401

Cloos, L., Ceulemans, E., & Kuppens, P. (2023). Development, validation, and comparison of self-report measures for positive and negative affect in intensive longitudinal research. *Psychological Assessment*, *35*(3), 189–204. https://doi.org/10.1037/pas0001200

Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. *Measuring Stress: A Guide for Health and Social Scientists*, *10*(2), 1–2.

Constantine, W., & Hesterberg, T. (2021). *splus2R: Supplemental S-PLUS functionality in R*. https://CRAN.R-project.org/package=splus2R

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004

Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., & Mestdagh, M. (2022). Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological Assessment*, *34*(12), 1138–1154. https://doi.org/10.1037/pas0001178

Dejonckheere, E., Mestdagh, M., Houben, M., Erbas, Y., Pe, M., Leuven, K. U., Koval, P., Brose, A., Bastian, B., & Kuppens, P. (2018). The bipolarity of affect and depressive symptoms. *Journal of Personality and Social Psychology*, *114*(2), 323–341. https://doi.org/10.1037/pspp0000186.supp

Dejonckheere, E., Mestdagh, M., Houben, M., Rutten, I., Sels, L., Kuppens, P., & Tuerlinckx, F. (2019). Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*, *3*(5), 478–491. https://doi.org/10.1038/s41562-019-0555-0

Ebner-Priemer, U. W., & Sawitzki, G. (2007). Ambulatory assessment of affective instability in borderline personality disorder. *European Journal of Psychological Assessment*, *23*(4), 238–247. https://doi.org/10.1027/1015-5759.23.4.238

Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The effects of

sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*, *29*(2), 136–151. https://doi.org/10.1177/1073191120957102

Ekkekakis, P. (Ed.). (2013). Untangling the terminological Gordian knot. In *The Measurement of Affect, Mood, and Emotion: A Guide for Health-Behavioral Research* (pp. 33–51). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511820724

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, *14*(4), 577–606. https://doi.org/10.1080/026999300402808

Fritz, J., Piccirillo, M., Cohen, Z. D., Frumkin, M., Kirtley, O., Moeller, J., Neubauer, A., Norris, L., Schuurman, N. K., Snippe, E., & Bringmann, L. (2023). *So you want to do ESM? Ten essential topics for implementing the Experience Sampling Method (ESM)*. Open Science Framework. https://doi.org/10.31219/osf.io/fverx

Gross, J. J., & Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: An affective science perspective. *Clinical Psychological Science*, *2*(4), 387–401. https://doi.org/10.1177/2167702614536164

Hall, M., Scherner, P. V., Kreidel, Y., & Rubel, J. A. (2021). A systematic review of momentary assessment designs for mood and anxiety symptoms. *Frontiers in Psychology*, *12*, Article 642044. https://doi.org/10.3389/fpsyg.2021.642044

Heylen, J., Van Mechelen, I., Fried, E. I., & Ceulemans, E. (2016). Two-mode K-spectral centroid analysis for studying multivariate longitudinal profiles. *Chemometrics and Intelligent Laboratory Systems*, *154*, 194–206. https://doi.org/10.1016/j.chemolab.2016.03.009

Heylen, J., Van Mechelen, I., Verduyn, P., & Ceulemans, E. (2016). KSC-N: Clustering of hierarchical time profile data. *Psychometrika*, *81*(2), 411–433. https://doi.org/10.1007/s11336-014-9433-x

Heylen, J., Verduyn, P., Van Mechelen, I., & Ceulemans, E. (2015). Variability in anger intensity profiles: Structure and predictive basis. *Cognition and Emotion*, *29*(1), 168–177. https://doi.org/10.1080/02699931.2014.896783

Himmelstein, P. H., Woods, W. C., & Wright, A. G. C. (2019). A comparison of signal- and event-contingent ambulatory assessment of interpersonal behavior and affect in social situations. *Psychological Assessment*, *31*(7), 952–960. https://doi.org/10.1037/pas0000718

Hoemann, K., Khan, Z., Feldman, M. J., Nielson, C., Devlin, M., Dy, J., Barrett, L. F., Wormwood, J. B., & Quigley, K. S. (2020). Context-aware experience sampling reveals the scale of variation in affective experience. *Scientific Reports*, *10*(1), 12459. https://doi.org/10.1038/s41598-020-69180-y

Hopwood, C. J., Bleidorn, W., & Wright, A. G. C. (2022). Connecting theory to methods in longitudinal research. *Perspectives on Psychological Science*, *17*(3), 884–894. https://doi.org/10.1177/17456916211008407

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*(6), 401–405.

Kalokerinos, E. K., Résibois, M., Verduyn, P., & Kuppens, P. (2017). The temporal deployment of emotion regulation strategies during negative emotional episodes. *Emotion*, *17*(3), 450–458. https://doi.org/10.1037/emo0000248

Kuppens, P., Dejonckheere, E., Kalokerinos, E. K., & Koval, P. (2022). Some recommendations on the use of daily life methods in affective science. *Affective Science*, *3*(2), 505–515. https://doi.org/10.1007/s42761-022-00101-0

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, *99*(6), 1042–1060. https://doi.org/10.1037/a0020962

Kuppens, P., Van Mechelen, I., Nezlek, J. B., Dossche, D., & Timmermans, T. (2007). Individual differences in core affect variability and their relationship to personality and psychological adjustment. *Emotion*, *7*(2), 262–274. https://doi.org/10.1037/1528-3542.7.2.262

Kuppens, P., & Verduyn, P. (2015). Looking at emotion regulation through the window of emotion dynamics. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory*, *26*(1), 72–79. https://doi.org/10.1080/1047840X.2015.960505

Leertouwer, I., Cramer, A. O. J., Vermunt, J. K., & Schuurman, N. K. (2021). A review of explicit and implicit assumptions when providing personalized feedback based on self-report EMA data. *Frontiers in Psychology*, *12*, Article 764526. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.764526

Leertouwer, I., Schuurman, N. K., & Vermunt, J. K. (2022). Are retrospective assessments means of people's experiences? *Journal for Person-Oriented Research*, *8*(2), 52–70. https://doi.org/10.17505/jpor.2022.24855

Makowski, D., Wiernik, B. M., Patil, I., Lüdecke, D., & Ben-Shachar, M. S. (2022). *correlation: Methods for correlation analysis*. https://CRAN.R-project.org/package = correlation

Mann, C. R., & Twiss, G. R. (1910). *Physics*. Scott, Foresman and Company. https://books.google.be/books?id = 8FERAAAAYAAJ

Mestdagh, M., & Dejonckheere, E. (2021). Ambulatory assessment in psychopathology research: Current achievements and future ambitions. *Current Opinion in Psychology*, *41*, 1–8. https://doi.org/10.1016/j.copsyc.2021.01.004

Mestdagh, M., Verdonck, S., Piot, M., Niemeijer, K., Kilani, G., Tuerlinckx, F., Kuppens, P., & Dejonckheere, E. (2023). m-Path: An easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. *Frontiers in Digital Health*, *5*, Article 1182175. https://www.frontiersin.org/articles/10.3389/fdgth.2023.1182175

Pinheiro, J., & Bates, D., & R Core Team. (2023). *nlme: Linear and nonlinear mixed effects models*. https://CRAN.R-project.org/package = nlme

Résibois, M., Kuppens, P., Van Mechelen, I., Fossati, P., & Verduyn, P. (2018). Depression severity moderates the relation between self-distancing and features of emotion unfolding. *Personality and Individual Differences*, *123*, 119–124. https://doi.org/10.1016/j.paid.2017.11.018

Résibois, M., Verduyn, P., Delaveau, P., Rotgé, J.-Y., Kuppens, P., Van Mechelen, I., & Fossati, P. (2017). The neural basis of emotions varies over time: Different regions go with onset- and offset-bound processes underlying emotion intensity. *Social Cognitive and Affective Neuroscience*, *12*(8), 1261–1271. https://doi.org/10.1093/scan/nsx051

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russell, J. A. (2005). Emotion in human consciousness is built on core affect. *Journal of Consciousness Studies*, *12*(8–10), 26–42.

Russell, J. A., & Feldman Barrett, L. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Schreuder, M. J., Groen, R. N., Wigman, J. T. W., Hartman, C. A., & Wichers, M. (2020). Measuring psychopathology as it unfolds in daily life: Addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*, *20*(1), Article 351. https://doi.org/10.1186/s12888-020-02674-1

N. Schwarz (Ed.). (1994). *Autobiographical memory and the validity of retrospective reports*. Springer.

Solhan, M. B., Trull, T. J., Jahng, S., & Wood, P. K. (2009). Clinical assessment of affective instability: Comparing EMA indices, questionnaire reports, and retrospective recall. *Psychological Assessment*, *21*(3), 425–436. https://doi.org/10.1037/a0016869

Sonnemans, J., & Frijda, N. H. (1994). The structure of subjective emotional intensity. *Cognition and Emotion*, *8*(4), 329–350. https://doi.org/10.1080/02699939408408945

Stone, A. A., & Shiffman, S. (1994). Ecological Momentary Assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, *16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Van Berkel, N., Merritt, T., Bruun, A., & Skov, M. B. (2022). Tangible self-report devices: Accuracy and resolution of participant input. In *Proceedings of the Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3490149.3501309

van Halem, S., van Roekel, E., Kroencke, L., Kuper, N., & Denissen, J. (2020). Moments that matter? On the complexity of using triggers based on skin conductance to sample arousing events within an experience sampling framework. *European Journal of Personality*, *34*(5), 794–807. https://doi.org/10.1002/per.2252

Veilleux, J. C., Schreiber, R. E., Warner, E. A., & Brott, K. H. (2024). Development and validation of a brief version of the emotion reactivity scale: The B-ERS. *Current Psychology*, *43*(14), 12586–12600. https://doi.org/10.1007/s12144-023-05323-4

Verduyn, P., Delvaux, E., Coillie, H. V., Tuerlinckx, F., & Mechelen, I. V. (2009). Predicting the duration of emotional experience: Two experience sampling studies. *Emotion*, *9*(1), 83–91. https://doi.org/10.1037/a0014610

Verduyn, P., Van Mechelen, I., & Frederix, E. (2012). Determinants of the shape of emotion intensity profiles. *Cognition and Emotion*, *26*(8), 1486–1495. https://doi.org/10.1080/02699931.2012.662152

Verduyn, P., Van Mechelen, I., Tuerlinckx, F., Meers, K., & Van Coillie, H. (2009). Intensity profiles of emotional experience over time. *Cognition & Emotion*, *23*(7), 1427–1443. https://doi.org/10.1080/02699930902949031

Vogelsmeier, L. V. D. E., Cloos, L., Kuppens, P., & Ceulemans, E. (2023). Evaluating dynamics in affect structure with latent Markov factor analysis. *Emotion*, *24*, 782–794. https://doi.org/10.1037/emo0001307

Waugh, C. E., & Kuppens, P. (Eds.). (2021). *Affect dynamics*. Springer. https://doi.org/10.1007/978-3-030-82965-0

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *dplyr: A grammar of data manipulation*. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr

Wright, A. G. C., Hopwood, C. J., & Simms, L. J. (2015). Daily interpersonal and affective dynamics in personality disorder. *Journal of Personality Disorders*, *29*(4), 503–525. https://doi.org/10.1521/pedi.2015.29.4.503