Anxiety modulates event segmentation

Oded Bein[1], Nora Wolf[2], Cara Khalifeh[3], Lauren Beugelsdyk[4], Dan-Mircea Mirea[3], Gili Karni[5], and Yael Niv[3,5]*

Affiliations

[1] Weill Cornell Institute of Geriatric Psychiatry, Department of Psychiatry, Weill Cornell Medicine, New York, New York, 10065, USA.

[2] Department of Psychology, University of California, Davis, CA, 95616, USA.

[3] Department of Psychology, Princeton University, Princeton, NJ, 08540, USA.

[4] Department of Physiology and Pharmacology, University of Georgia, Athens, GA, 30605, USA.

[5] Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08540, USA.

*Correspondence: odb4002@med.cornell.edu

Keywords: anxiety, event segmentation, context, generalization.

Abstract

Anxiety is one of the most prevalent mental health concerns. Current theories suggest that anxiety may arise due to deficits in segmentation of continuous experience into discrete context representations (‘event segmentation’), which leads to overgeneralization of fear across contexts, or conversely, overly rigid segmentation that prevents safety learning. Here, in two segmentation tasks (N=1109), we found novel and direct evidence that anxiety is associated with changes in event segmentation. Individuals with higher anxiety symptoms responded more slowly to transitions between events (event boundaries). They also segmented movies into discrete events more typically and more hierarchically, two hallmarks of precise segmentation. This precision was linked to self-reporting fewer context changes in daily life, suggesting individuals with anxiety prefer stable and predictable environments. These findings challenge overgeneralization theories of anxiety, revealing instead that individuals with anxiety exhibit precise, and potentially overly rigid segmentation. Such segmentation could maintain fear by preventing generalization from safe to fearful contexts, which has important implications for interventions like exposure therapy.

Introduction

Anxiety disorders are among the most common mental health conditions worldwide, contributing significantly to global burden and disability (Alonso et al., 2018; Craske et al., 2017; Yang et al., 2021). Understanding the cognitive processes that contribute to excessive anxiety is crucial for developing more targeted interventions and improving therapeutic outcomes (Abend, 2023; Beckers et al., 2023; Hamm, 2019; LeDoux & Pine, 2016). Specifically, understanding how different contexts are distinguished and represented in one's mind might be essential. Excessive anxiety might result from difficulty in properly distinguishing between dangerous and safe contexts, leading to overgeneralizing fear from dangerous to safe situations (Cooper et al., 2022; Duits et al., 2015; Dunsmoor & Paz, 2015; Lee et al., 2024). For example, if one has a fearful experience in a specific park, and one tends to lump this experience in their mind with other parks, or even other public places, then all parks or public places might become threatening. Alternatively, distinguishing contexts too precisely might also maintain fear by preventing safety learning (Gershman et al., 2010). If, after having a fearful park experience, additional safe or positive experiences in parks are represented as part of a separate 'safe parks' context, these experiences may not update the 'dangerous parks' representation, which remains unchanged and continues to evoke fear. Thus, anxiety might be related to how we tend to segment or lump events together in our minds.

Event segmentation – the process by which we segment continuous experience into discrete context representations ('events') – is fundamental to how context is inferred and represented, and is therefore essential to adaptive learning and generalization (Botvinick, 2008; Clewett et al., 2019; Shin & DuBrow, 2021; Zacks, 2020; Zacks et al., 2007). According to event segmentation theory, our brain maintains representations of current experience within a context (e.g., a visit to a park). Changes of context, termed event boundaries, for example, when crossing the street to leave the park, signal to our brain to end one event representation and start a new one. Thus, event boundaries distinguish different contexts and create distinct representations in our minds. These context representations then become the building blocks of more abstract higher-level context representations (sometimes termed 'event schemas', Bein & Niv, 2025; Beukers et al., 2024; Franklin et al., 2020; Zacks, 2020), e.g., of 'safe parks' or 'dangerous parks'. We reason that in individuals with trait anxiety, deficits in event segmentation can lead to lumping together contexts and overgeneralization of fear. Alternatively, enhanced segmentation might lead to overly rigid context representations that prevent safety learning.

Previous research points to either insufficient or enhanced segmentation in anxiety. Some research suggests that individuals with anxiety have difficulties distinguishing stimuli that evoke fear from neutral stimuli (Duits et al., 2015; Dunsmoor & Paz, 2015; Modecki et al., 2023), which might point towards difficulty distinguishing experiences in different contexts. Further, event boundaries elicit arousal (Antony et al., 2021; Clewett et al., 2020), and the physiological response to arousing events is reduced in patients with anxiety disorders (Duits et al., 2015; Hoehn-Saric & McLeod, 2000), who also show less variation in arousal during everyday life (Hoehn-Saric & McLeod, 2000). This might lead to impaired boundary processing and segmentation. Event boundaries are also associated with uncertainty due to reduced ability to predict the coming event (Baldwin & Kosie, 2021; Eisenberg et al., 2018; Zacks et al., 2011), and individuals with anxiety show impaired uncertainty processing (Brown et al., 2023;

Browning et al., 2015; Grupe & Nitschke, 2013; Pulcu & Browning, 2019).

On the other hand, a recent study using reversal fear learning showed that individuals with high anxiety adjusted their expectations more quickly in response to changes in context (Zika et al., 2023). This behavior was captured by a computational model that accurately separated contexts when changes occurred, suggesting that anxiety symptoms might be associated with precise context segmentation. To arbitrate between these two possibilities, here we tested directly whether event segmentation is impaired or enhanced in people with symptoms of anxiety.

In the first experiment, we used a naturalistic reading task to measure response slowing due to event boundaries. We found that participants who reported higher anxiety symptoms responded more slowly to event boundaries. In a second experiment with a pre-registered analysis plan, we tested whether this slow boundary response reflects impaired boundary processing that leads to poor segmentation, or rather enhanced boundary processing that leads to precise segmentation. We found evidence for precise event segmentation in participants with symptoms of anxiety. Thus, both experiments provided converging evidence that anxiety is associated with changes in event segmentation. We also explored how segmentation might relate to participants' everyday life experiences of context changes and found that individuals with symptoms of anxiety tended to have fewer changes of context in their everyday life, a tendency that mediated more precise segmentation.

Finally, we tested how different aspects of anxiety—namely, cognitive anxiety (worrying, intrusive thoughts, and difficulty concentrating), somatic anxiety (hyperventilation, sweating, increased heart rate, and muscle tension; Grös et al., 2007; Ree et al., 2008; Styck et al., 2022), and social anxiety (excessive fear and avoidance of social situations and social evaluation; Claus et al., 2023; Hofmann, 2007; Horenstein et al., 2018; Spokas et al., 2007; Watson & Friend, 1969)—might differently impact event segmentation. We found that precise segmentation was specifically correlated with symptoms of social anxiety. These findings suggest that symptoms of anxiety, and potentially social anxiety, may result from overly rigid context representations that prevent generalization from safe contexts to fearful ones.

Results

Experiment 1: Story reading

Sample Characteristics

We recruited N=662 U.S.-based adult participants on Prolific. Demographics and breakdown according to anxiety characteristics are presented in Table 1. We measured anxiety using two subscales from the Personality Inventory of the DSM-5 (PID-5; Krueger et al., 2012): Anxiousness (4 questions, 0-3 scale) and Social withdrawal (4 questions, 0-3 scale), capturing non-social and social anxiety symptoms, respectively. Our sample was similar to previously published norms with PID-5 (Krueger et al., 2012; Miller et al., 2022), though slightly more anxious (Anxiousness: M = 5.22, SD = 4.00; Social withdrawal: M = 4.74, SD = 3.32). This is likely due to increased anxiety during the COVID-19 pandemic (Delpino et al., 2022), when the data were collected. The two subscale measures had high internal consistency in our sample (Cronbach's alpha (Cronbach, 1951): Anxiousness: .93; Social withdrawal: .87), and were correlated (Pearson's r = .54), as expected. We therefore summed both measures to create a composite anxiety score and used both the composite and the two separate scores in our analyses.

Sample Characteristics (**N = 662**)	
Gender	
Woman	372 (56.2%)
Man	271 (40.9%)
Non-binary	19 (2.9%)
Age (years)	40.34 (14.29)
Race	
African American	48 (7.3%)
American Indian	2 (0.3%)
Asian	31 (4.7%)
Multiracial	24 (3.6%)
Native Hawaiian	1 (0.2%)
Other	7 (1.1%)
White	549 (82.9%)
Ethnicity	
Hispanic	46 (6.9%)
Not Hispanic	605 (91.4%)
Other	11 (1.7%)
Anxiousness	M=5.21 (SD=4.00)
High (≥M+1 SD)	126 (19.0%)
Medium (M±1 SD)	360 (54.4%)
Low (≤M–1 SD)	176 (26.6%)
Social withdrawal	M=4.74 (SD=3.32)
High (≥M+1 SD)	97 (14.7%)
Medium (M±1 SD)	418 (63.1%)
Low (≤M–1 SD)	147 (22.2%)

Table 1. Participant characteristics for the Stories task (N = 662). Values represent counts (percentage) unless otherwise stated. Anxiety levels represent sample-based tertile splits using mean ± standard deviation cutoffs.

Replication of slower boundary processing

Participants read stories displayed sentence by sentence. Reading times were measured as participants pressed a button to see the next sentence. Some sentences included large changes in the context or situation (boundary sentences) while other sentences were identical or highly similar in wording (see Methods) to boundary sentences but did not include a context change (no-boundary sentences). Previous research established reading times as a marker of boundary processing, showing that boundary sentences take longer to process than no-boundary sentences (Heusser et al., 2018; Pettijohn & Radvansky, 2016; Zwaan, 1996; Zwaan & Radvansky, 1998).

We developed stories specifically for individual differences research, incorporating rigorous controls that allowed identical sentences to function as boundary or no-boundary sentences across different participant groups, while using structurally similar sentences within participants (see Methods, Supplementary Information). Three pilot studies validated that our boundary sentences were perceived as larger situational changes, rated as more unexpected, and produced slower reading times compared to no-boundary sentences (Supplementary Information). The stories and sentence-by-sentence ratings are available in Appendix 1.

Using these well-controlled stories, we replicated the boundary effect in the current sample (Figure 1A): sentences that included event boundaries led to slower reading times compared to sentences without boundaries (boundary: M = 197.96 ms/syllable, SD = 58.45 ms; no-boundary: M = 176.84 ms/syllable, SD = 55.47 ms, $t_{(661)} = 30.77$, $p < .00001$, Cohen's d = 1.20; all Student's t tests reported here and elsewhere in the paper were two-sided). Similar results were obtained from a mixed-level model controlling for sentence identity and order (see

Methods; β = .078, $\chi^2_{(1)}$ = 357.21, p < .00001, BIC reduction: 347; all β's reported here and elsewhere in the paper are standardized).

Slower boundary processing in anxiety

Anxiety symptoms, as measured by the composite score, correlated with a larger boundary effect (β = .092, $t_{(655)}$ = 2.39, p = .017), suggesting slower processing of boundaries in anxiety (Figure 1B). When including the two anxiety subscale scores in a single regression model, we found some evidence that the larger boundary effect was related to symptoms of social anxiety (social: β = .08, $t_{(654)}$ = 1.93, p = .055; non-social: β = .02, $t_{(654)}$ = .42, p = .68; note that this is a stringent analysis given the correlation between these subscales—when the model included only social anxiety, it was significant: β = .09, $t_{(655)}$ = 2.53, p = .012). All analyses reported here and elsewhere in the paper controlled for age and gender. All reported results remained when controlling for cognitive functions by including participants' scores on three cognitive tasks (see Methods).

We divided boundary sentences into strong versus weak boundaries based on ratings from a separate group of participants and found that social anxiety symptoms correlated with a larger boundary effect for weak boundaries, while non-social anxiety symptoms correlated with a larger boundary effect for strong boundaries (see Supplementary Information).

No correlation between boundary response and depressive symptoms

Given the known comorbidity between anxiety and depression (Kaiser et al., 2021; Kaufman & Charney, 2000), we asked whether our results were specific to anxiety. First, we found no relationship between boundary response and depressive symptoms as measured using the Depressive Attributions Questionnaire (DAQ, Kleim et al., 2011, 17 items; p = .43). When controlling for depression, anxiety remained significant (p = .01).

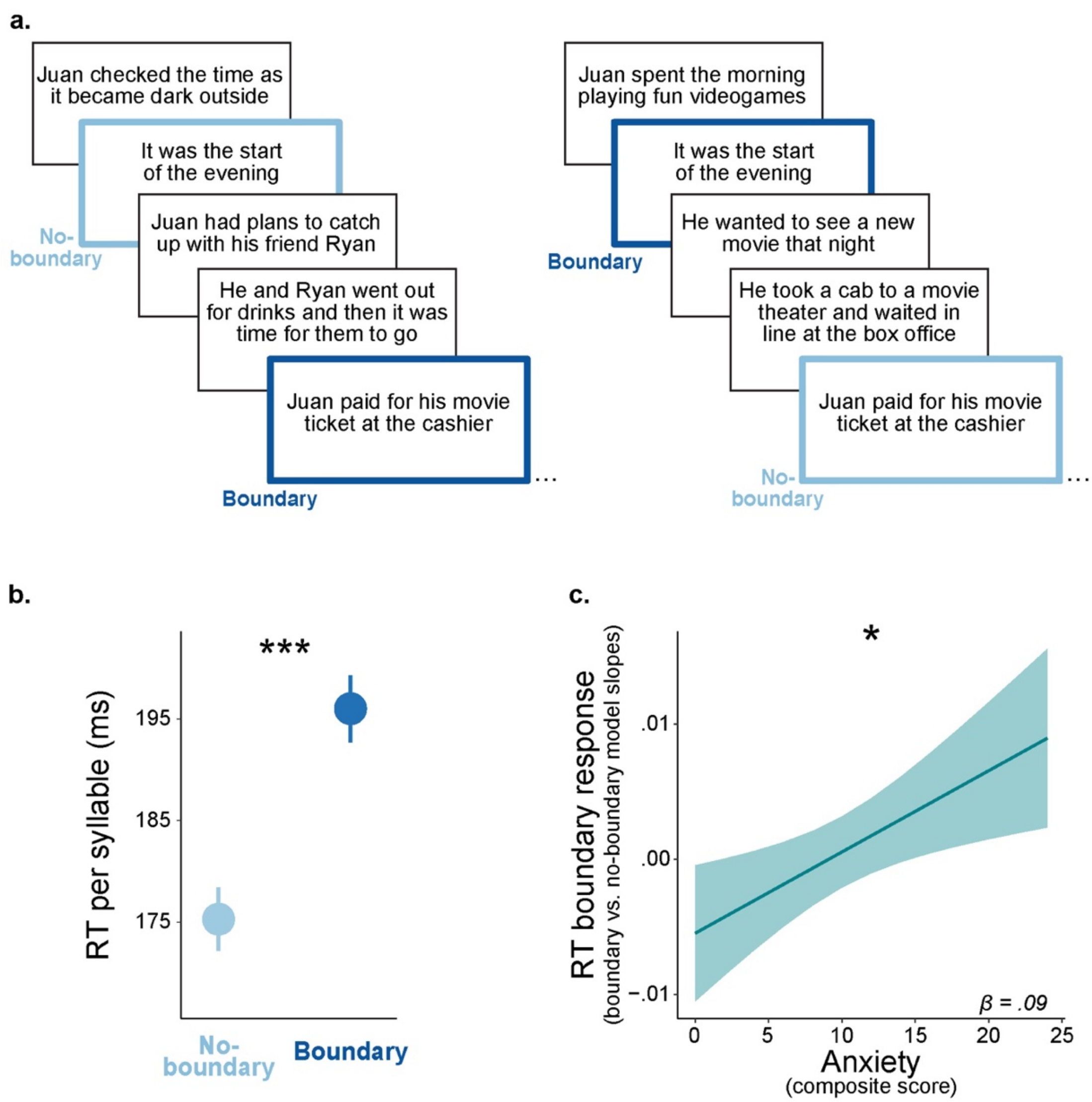

Figure 1. Anxiety correlates with slow boundary responses. **a.** Task: participants read stories sentence by sentence, pressing a button when done reading each sentence. Boundary sentences included a large change in context or situation, while matched no-boundary control sentences did not. Examples are from two versions of the same story, read by different participants, showing the same text as a boundary sentence in one version, and as a no-boundary sentence in the other version. Content was equated between versions of stories (e.g., in the version on the right side, Juan and Ryan then meet at the movie theater). See Methods for full details of how boundary and no-boundary sentences were matched both across and within participants. Colors are for emphasis and were not shown to participants. **b.** RT boundary response: longer reading times (RT per syllable) for boundary sentences compared to no-boundary sentences. **c.** We estimated the RT effect of boundaries per participant in a regression model accounting for confounds such as sentence order and sentence identity. Participants' boundary vs. no-boundary slopes from that model, reflecting the RT boundary response, were significantly correlated with anxiety symptoms. * p < .05, *** p < .001.

Experiment 2: Movie segmentation

In Experiment 1, we found that individuals with anxiety responded more slowly to event boundaries. Does the slower boundary response reflect difficulty in boundary processing or enhanced sensitivity to boundaries that results in enhanced segmentation? To answer this, in Experiment 2, N=442 participants segmented short movie clips into distinct events by pressing a button when, in their mind, changes of situations or contexts occurred (Figure 2A; Michelmann et al., 2023; Newtson, 1973; Zacks et al., 2001). Previous work established that people segment events similarly and interpreted the group-average segmentation as 'ground truth' for precise segmentation (Michelmann et al., 2023; Sargent et al., 2013; Sasmita & Swallow, 2022; Sava-Segal et al., 2023; Zacks, 2020; Zacks et al., 2001). Across participants, similarity to the group segmentation (termed segmentation typicality, also known as segmentation 'agreement'; Figure 2B) correlates with better event memory (Kurby & Zacks, 2011; Sargent et al., 2013), consistent with the view of typical segmentation as precise segmentation (Kurby & Zacks, 2011; Sargent et al., 2013).

We asked participants to segment each movie clip twice: first based on the largest meaningful units (coarse segmentation), and then, about a week later, based on the smallest meaningful units (fine segmentation). We imposed the one-week delay to minimize memory of previous segmentation decisions (Sargent et al., 2013). This procedure allowed us to capture within-participant hierarchical segmentation—a measure of the precision of event segmentation, where fine boundaries align with coarse boundaries (Figure 2C; Baldassano et al., 2017; Kurby & Zacks, 2008; Zacks et al., 2001). Analyses for this experiment were pre-registered (https://osf.io/y9c54/).

Sample Characteristics

We recruited N=442 U.S.-based adult participants on Prolific. Demographics and anxiety characteristics are presented in Table 2. Given the results of Experiment 1, here we measured anxiety more comprehensively using the trait version of the State-Trait-Inventory of Cognitive and Somatic Anxiety (STICSA; 21-items, 0-3 scale; Ree et al., 2008; Styck et al., 2022) for cognitive and somatic symptoms of anxiety and the Liebowitz Social Anxiety Scale (LSAS-SR; 48-items, 0-3 scale; Heimberg et al., 1999; Safren et al., 1999) for social anxiety symptoms. Anxiety symptoms (STICSA: M = 35.30; SD = 10.35; LSAS: M = 52.11, SD = 29.40) were slightly elevated compared to previous reports in community samples and the general population (Baxter et al., 2013; Van Dam et al., 2013), but comparable to levels reported during the COVID-19 pandemic when these data were collected (Delpino et al., 2022). STICSA and LSAS scores had excellent internal consistency (Cronbach's alpha for STICSA: .91; LSAS-SR: .96), and were correlated with each other (Pearson's r = .65), as expected. To account for the different numbers of items per scale, we computed a composite anxiety score by summing the average score in each scale.

Sample Characteristics **(N=442)**	
Gender	
Woman	177 (40.0%)
Man	257 (58.1%)
Non-binary	8 (1.8%)
Age (years)	37.93 (13.11)
Race	
African American	31 (7.0%)
American Indian	2 (0.5%)
Asian	26 (5.9%)
Multiracial	16 (3.6%)
Other	9 (2.0%)
White	358 (81.0%)
Ethnicity	
Hispanic	39 (8.8%)
Not Hispanic	395 (89.4%)
Other	8 (1.8%)
STICSA	M=35.35 (SD=10.35)
High (≥43)	111 (25.1%)
Medium (24–42)	265 (60.0%)
Low (≤23)	66 (14.9%)
LSAS-SR	M=52.51 (SD=29.40)
High (≥60)	169 (38.2%)
Medium (30–59)	161 (36.4%)
Low (<30)	112 (25.3%)

Table 2. Sample characteristics for the movie segmentation task. Values represent counts (percentage) unless otherwise stated. STICSA: State-Trait Inventory of Cognitive and Somatic Anxiety; LSAS-SR: Liebowitz Social Anxiety Scale-Self Report. We used recommended clinical thresholds for anxiety severity classifications: ≥43 for STICSA (Van Dam et al., 2013) and ≥60 for generalized social anxiety in LSAS-SR (Fresco et al., 2001; Mennin et al., 2002; Rytwinski et al., 2009).

<u>Replication of previous findings: segmentation typicality and robust hierarchical segmentation</u>

Segmentation typicality (correlations of each participant's segmentation with the group norm: <u>fine:</u> *M* = .66, *SD* = .15; <u>coarse:</u> *M* = .53, *SD* = .20) was comparable to previous studies (Kurby & Zacks, 2011; Pitts et al., 2022; Sargent et al., 2013). We also observed robust hierarchical segmentation, namely, fine and coarse boundaries were more aligned (i.e., were temporally closer to each other) than expected by chance based on each participant's boundaries (Zacks et al., 2001; chance log distance minus actual log temporal distance: M = .43, SD = .39, where larger numbers reflect more hierarchical segmentation; $\chi^2_{(1)}$ =873.47, p < .00001, BIC reduction: 865 when comparing a mixed-level model with a regressor of chance vs. actual distance to a reduced model without this regressor, with both models accounting for movie order and identity; see Supplementary Information). These results replicate prior in-lab studies (Kurby & Zacks, 2011; Pitts et al., 2022; Sargent et al., 2013), here in an online sample.

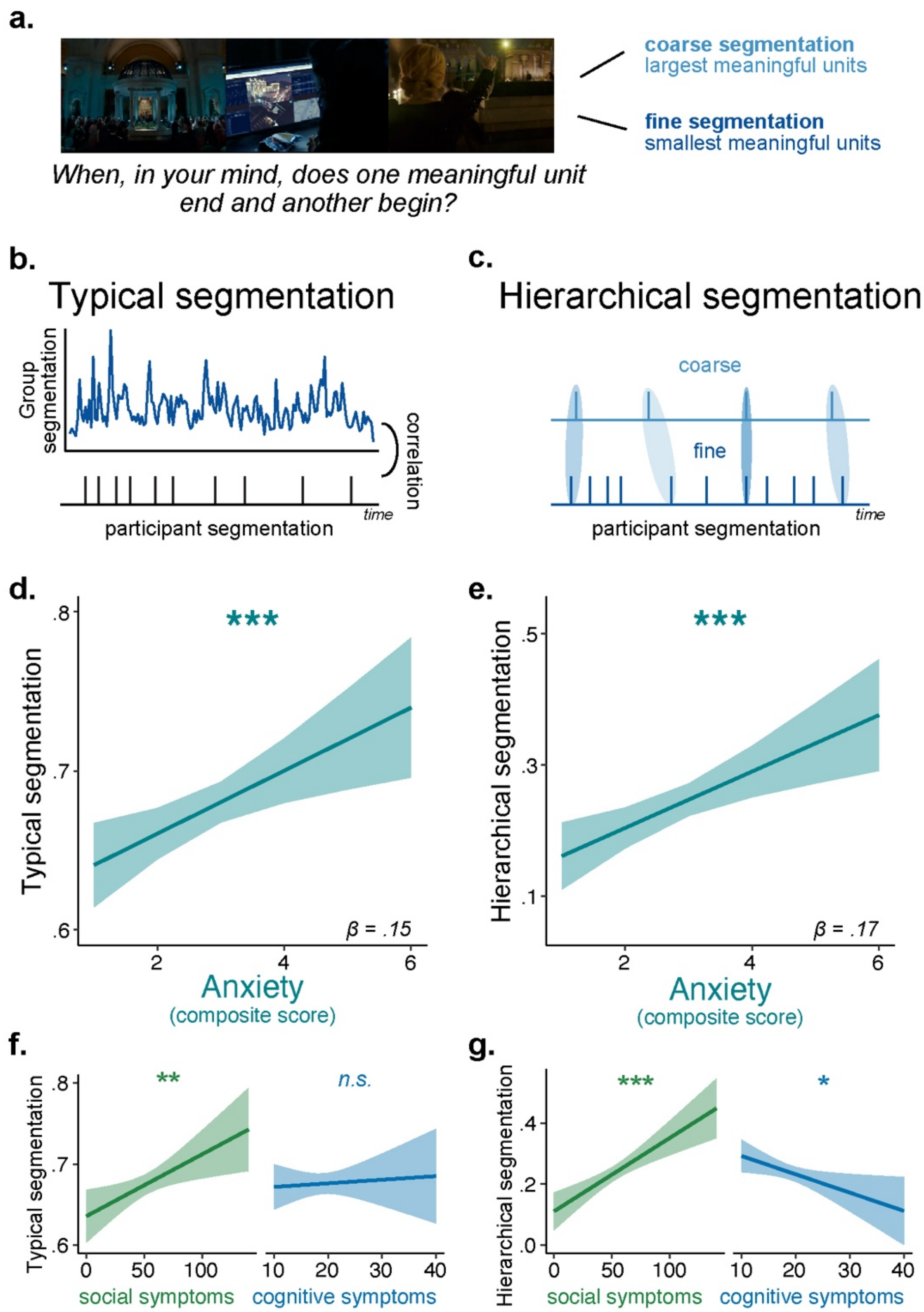

Figure 2. Anxiety correlates with typical and hierarchical segmentation. **a.** Participants segmented short movie clips based on the largest meaningful units (coarse segmentation), and the smallest meaningful units (fine segmentation). **b.** To compute typical segmentation, we first computed the group segmentation norm (i.e., the proportion of participants indicating a boundary at each time point; top) separately per fine and coarse segmentation. We then correlated each participant's segmentation time course (bottom) with the group segmentation norm. **c.** Hierarchical segmentation refers to the alignment between fine and coarse boundaries, computed as the average distance between each coarse boundary and the closest fine boundary, relative to chance. We subtracted the observed distance from the chance distance such that larger numbers reflect more hierarchical segmentation. In the illustration, darker shades reflect more aligned fine and coarse boundaries. **d.** Participants with more severe symptoms of anxiety showed higher segmentation typicality. **e.** Participants with more severe symptoms of anxiety showed a stronger segmentation hierarchy. **f.** Higher segmentation typicality correlated mostly with symptoms of social anxiety, but not cognitive anxiety. **g.** Stronger segmentation hierarchy correlated with more severe symptoms of social anxiety, but less severe symptoms of cognitive anxiety. Segmentation typicality results shown are from fine segmentation; coarse segmentation was similar. No segmentation measure correlated with somatic anxiety (not shown). * $p < .05$, ** $p < .01$, *** $p < .005$

Anxiety correlates with more typical segmentation and better hierarchical segmentation

Anxiety, as measured using the combined scores of both STICSA-trait and LSAS-SR questionnaires, correlated with more typical segmentation (fine: $\beta = .15$, $t_{(437)} = 2.97$, $p = .003$; coarse: $\beta = .12$, $t_{(437)} = 2.44$, $p = .015$) and more hierarchical segmentation ($\beta = .17$, $t_{(437)} = 3.34$, $p = .0009$). These results survived correction for multiple comparisons using the preregistered sd.minP procedure (adjusted p's $< .037$; Blakesley et al., 2009; Westfall & Young, 1993; see Methods). Anxiety symptoms did not correlate with the number of boundaries participants identified (fine: $\beta = -.04$, $t_{(437)} = .76$, $p = .44$; coarse: $\beta = -.05$, $t_{(437)} = 1.01$, $p = .31$).

Note that the typicality and hierarchical segmentation measures both account for the number of segments identified (see Methods). Nevertheless, we established that the results survive additional control by including the number of segments in the regression models (all p's $< .02$). We also repeated the analyses for segmentation typicality by computing the group segmentation profile based on all participants (rather than only non-anxious participants; see Methods) and obtained similar results. Together, these findings suggest that anxiety correlates with more structured and precise segmentation, as reflected by both more typical and more hierarchical segmentation.

When examining specific aspects of anxiety (Figure 2 F-G), symptoms of social anxiety correlated with more typical (fine: $\beta = .16$, $t_{(435)} = 2.644$, $p = .011$; coarse: $\beta = .13$, $t_{(435)} = 2.06$, $p = .04$) and more hierarchical segmentation ($\beta = .27$, $t_{(435)} = 4.39$, $p < .0001$). In contrast, cognitive symptoms of anxiety correlated with poorer hierarchical segmentation ($\beta = -.16$, $t_{(435)} = 2.28$, $p = .023$). Somatic symptoms of anxiety did not correlate with segmentation typicality or hierarchy (p's $> .19$).

The correlation between anxiety and segmentation is robust to influences of emotion and prior knowledge

Emotion, and specifically arousal, alters segmentation (Clewett et al., 2020; McClay et al., 2023; Riegel et al., 2023). Even though the movie clips we used were not meant to elicit strong emotions, it is possible that participants with anxiety had different emotional responses to the movies, and that variance in emotional response might underlie differences in segmentation. Another factor theorized to influence segmentation is prior knowledge of the situations in the movies, which is suggested to make segmentation more structured (Franklin et al., 2020; Levine et al., 2017; Newberry et al., 2021; Newberry & Bailey, 2019; Smith et al., 2020; Zacks et al., 2001). To control for potential influences of emotion and prior knowledge, we asked participants to rate valence and arousal at the end of each movie clip, as well as their familiarity with the general situation in the movie clip.

Worse anxiety symptoms correlated with lower valence ratings and higher arousal ratings across participants. Higher arousal, in turn, correlated with typicality of fine segmentation (see Supplementary Information). Importantly, all our results above held when controlling for valence, arousal, and familiarity both within and across participants: more severe symptoms of anxiety correlated with fine and coarse segmentation typicality and with hierarchical segmentation (p's $< .035$). As in the main analysis, this was driven by symptoms of social anxiety in the three segmentation measures (fine typicality: $p = .028$; coarse typicality: $p = .080$; hierarchical segmentation: $p < .001$), and cognitive symptoms of anxiety were marginally correlated with less hierarchical segmentation ($p = .059$).

Fewer context changes in everyday lives correlated with more severe anxiety and more typical and hierarchical segmentation

We were interested in whether segmentation behavior—and anxiety—relate to participants' experiences of changes in their everyday lives. For this, we developed a questionnaire asking about experience and preference for context or situation changes in everyday life (e.g., "In a typical day, I usually stay in one space most of the day"). We used bifactor analysis to extract one general factor that captures participants' experience of context changes (validation analyses in Supplementary Information, the questionnaire is available in Appendix 2), and correlated participants' scores in this factor with anxiety symptoms and with segmentation behavior in the task.

Individuals with more severe anxiety symptoms reported fewer context changes in everyday life (β = -.47, $t_{(437)}$ = 10.50, p < .0001), and this was specific to symptoms of social anxiety (β = -.52, $t_{(435)}$ = 9.40, p < .0001; cognitive symptoms of anxiety: p > .29; somatic symptoms of anxiety: β = .09, $t_{(435)}$ = 1.67, p = .09). Fewer changes in everyday life also correlated significantly with higher segmentation typicality and more hierarchical segmentation (fine segmentation typicality: β = -.14, $t_{(436)}$ = -2.61, p < .01; coarse segmentation typicality: β = -.11, $t_{(437)}$ = 1.85, p = .02; hierarchical segmentation: β = -.15, $t_{(437)}$ = 3.28, p = .001; analyses control for overall anxiety symptoms). Mindful that our online sample might include a large number of individuals with anxiety who also potentially prefer to stay home, we tested the relationship between context changes in everyday life and segmentation subsampling only participants that reported low levels of anxiety (N = 252, Methods) and found a significant relationship with fine segmentation typicality (β = -.18, $t_{(247)}$ = 2.96, p = .003; the correlations with coarse segmentation typicality and hierarchical segmentation were not significant, $p > .16$).

<u>Segmentation and depressive symptoms</u>

As in the story reading task, given the comorbidity between anxiety and depression (Kaiser et al., 2021; Kaufman & Charney, 2000), we asked whether segmentation correlates specifically with symptoms of anxiety, or also with symptoms of depression (as measured using the DAQ; Kleim et al., 2011). We found no correlation between symptoms of depression and segmentation typicality or hierarchy (p's > .20), and both relationships between segmentation and overall symptoms of anxiety and between segmentation and symptoms of social anxiety remained significant when controlling for symptoms of depression (all p's < .029). Interestingly, when examining symptoms of anxiety and depression together in one model for fine typicality, both were significant (fine: anxiety: $\beta = .28$, $t_{(436)} = 3.87$, $p = .0001$; depression: $\beta = -.012$, $t_{(436)} = 2.511$, $p = .012$; similar results were obtained with symptoms of social anxiety). While these results suggest depression correlated with less fine typical segmentation, they should be interpreted with caution due to the small effect size and recent analyses showing that spurious correlations with outcome could emerge when including two correlated measures such as symptoms of anxiety and depression in the model (Sarna et al., 2025).

Discussion

Event segmentation, namely, the parsing of continuous experience into discrete context representations in our mind, is fundamental to how we learn adaptive world models (Baldassano et al., 2018; Bein & Niv, 2025; Clewett et al., 2019; Zacks, 2020). We showed here that symptoms of anxiety are associated with differences in event segmentation, and specifically, with a slow and arguably more precise segmentation process. In a story-reading task, we found that individuals with more severe anxiety symptoms responded more slowly to transitions between contexts (event boundaries) compared to those with low levels of anxiety. We then asked whether slow boundary response reflects difficulty in event segmentation or enhanced boundary processing that leads to precise segmentation. In a movie segmentation task, we found that individuals with symptoms of anxiety showed more precise segmentation—their individual segmentation was better aligned with 'typical' group-averaged segmentation and more hierarchically consistent (i.e., their fine boundaries were better aligned with their coarse boundaries). Together, these findings suggest careful and precise event segmentation in anxious individuals.

We also explored how anxiety and segmentation relate to everyday experiences of context changes using a novel questionnaire we developed and validated. Symptoms of anxiety correlated with fewer context changes in participants' everyday lives, suggesting that individuals with anxiety may avoid context changes due to the uncertainty they engender (Antony et al., 2021; Brown et al., 2023; Browning et al., 2015; Ezzyat & Clements, 2024; Zacks et al., 2011), or, alternatively, lack of changes in daily life can give rise to anxiety symptoms. This finding was specific to symptoms of social anxiety, which could reflect that individuals with social anxiety prefer predictable environments where negative social evaluations are less likely. Fewer changes in daily life also correlated with more precise segmentation, consistent with event stability contributing to enhanced segmentation (Bein & Davachi, 2024; Clewett & Davachi, 2017; Ezzyat & Davachi, 2021). Another possibility is that precise segmentation might be cognitively taxing, which drives avoidance of changes in context in their everyday lives. Of note, participants completed the daily life changes scale in a separate session and were not informed that it was a part of the same experiment, reducing the likelihood of demand characteristics.

Specificity to social anxiety and potential mechanisms

Both the slow response to event boundaries and precise segmentation were specifically associated with symptoms of social anxiety. This could be because the stories and movies mostly described social interactions or activities done in social settings, which may have aggravated social anxiety symptoms. However, increased anxiety during the tasks was unlikely to be responsible for our results, as individuals with social anxiety are typically anxious when participating in or anticipating social evaluation, not when observing social interactions in which they will not participate (such as in a movie or a story). This was supported by our finding that social anxiety symptoms did not correlate with emotion ratings in the movie segmentation task. We also consider it unlikely that fear of negative evaluation (e.g., by the experimenter) drove more precise segmentation as participants could not have known what constituted “typical segmentation”, nor could they have anticipated that we would measure hierarchical alignment between their two segmentation sessions. Furthermore, individuals with social anxiety did not show superior performance on any of the additional cognitive tasks we administered.

Instead, several characteristics of social anxiety might have led to precise segmentation. Individuals with social anxiety exhibit heightened alertness to critical aspects of social events and engage in extensive thinking about social events (post-event processing; Brozovich & Heimberg, 2008; Spokas et al., 2007). This suggests that people with social anxiety might attend preferentially to event boundaries and process these boundaries carefully, making them both slower and more precise at event segmentation. Precise segmentation, in turn, improves memory of events (Gershman et al., 2014; Kurby & Zacks, 2011; Sargent et al., 2013), which might contribute to overthinking and maintenance of social anxiety. Such laborious and careful processing of boundaries may also make social interactions (where one cannot control the rate of change in context) more cognitively taxing and thus contribute to avoidance behavior.

Our findings align with previous research showing difficulty processing uncertainty in anxiety (Brown et al., 2023; Browning et al., 2015; Grupe & Nitschke, 2013; Hengen & Alpers, 2021; Pulcu & Browning, 2019). This difficulty could explain why individuals with anxiety respond more slowly to boundaries, which induce uncertainty about what follows (Baldwin & Kosie, 2021; Ezzyat & Clements, 2024; Zacks et al., 2011). Additionally, individuals with anxiety might be especially motivated to predict unfolding events to reduce uncertainty about the future, leading to precise segmentation.

A stable everyday life environment in individuals with anxiety may contribute to stronger event schemas (Bein & Niv, 2025; Zacks, 2020) and precise segmentation. Indeed, we found that individuals with anxiety prefer and experience fewer context changes in their everyday life. This may lead to stronger event schemas through repeated exposure to stable and predictable events (Bein & Davachi, 2024). These stronger event schemas might, in turn, contribute to precise event segmentation, since events typically follow their event schemas. Of course, it is also possible that people who have fewer context changes develop schemas that are too strong or rigid, which may lead to anxiety due to the threat of possible deviations from schemas. Interestingly, key brain regions involved in event segmentation and schema representation—the hippocampus and medial prefrontal cortex (Baldassano et al., 2018; Bein & Niv, 2025; Ben-Yakov et al., 2014; Ben-Yakov & Henson, 2018; Clewett et al., 2019; DuBrow & Davachi, 2016; Ezzyat & Davachi, 2011; Schapiro et al., 2013, 2016)—also show increased activation in individuals with social anxiety during emotional and social processing (Brühl et al., 2014; Cremers & Roelofs, 2016; Klumpp et al., 2010). This enhanced processing in schema-related and segmentation-related brain regions may contribute to more precise segmentation.

Another possibility is that the slower boundary processing we observed might reflect a surprise-based mechanism (Baldwin & Kosie, 2021; Eisenberg et al., 2018; Zacks et al., 2011). Individuals with anxiety who experience fewer boundaries in their everyday life might experience greater surprise when boundaries occur compared to individuals with lower levels of anxiety. This surprise, in turn, could recruit neural and attentional mechanisms that facilitate segmentation (Antony et al., 2021; Clewett et al., 2020, 2025; Nolden et al., 2024; Sinclair et al., 2021; Zacks et al., 2011). However, this surprise-based interpretation might predict that individuals with anxiety would be more surprised by relatively small changes and identify them as boundaries, resulting in higher number of boundaries in the movie segmentation task, which we did not observe.

These mechanistic hypotheses regarding uncertainty, surprise, and event schemas can be tested using the computational framework of latent cause inference. Latent cause inference addresses the computational process by which experiences are assigned to distinct contexts ('latent causes'; Gershman et al., 2010, 2014, 2017), and was recently suggested to underlie event segmentation (Bein & Niv, 2025; Beukers et al., 2024; Franklin et al., 2020; Shin & DuBrow, 2021). In this framework, each observation of stimuli and occurrences in the world is generated by a hypothetical latent (hidden) 'cause' that produces observed events with certain characteristics. As occurrences unfold, our brains infer which latent cause 'generated' these occurrences—one of the latent causes previously encountered, or a completely new latent cause, thereby segmenting a continuous sequence of input into distinct events (Shin & DuBrow, 2021). Future research could apply this computational modeling framework to test how specific parameters underlying segmentation change with anxiety. For example, uncertainty within events can be thought of as the variance of occurrences that an individual can tolerate within a single latent cause. Predictions of future occurrences stem from the typical characteristics of latent causes – the event schema – and deviations from that schema cause surprise. These properties could be quantified within individuals (e.g., Mirea, Shin, et al., 2024) and tested for a relationship between parameters of latent cause inference and symptoms of anxiety.

Consistent with the idea that latent cause inference is a general cognitive process that is not specific to emotional experiences, here we aimed to study segmentation as a fundamental cognitive process that might be altered in anxiety and is not specific to fearful or emotional events. Thus, our stimuli were not intended to evoke strong emotions. We found that precise segmentation was robust to participants' emotional responses to the movie clips. In an exploratory analysis (Supplementary Information), we found that arousal correlated with segmentation typicality, that is, participants segmented movie clips that they found arousing more typically than other clips. This finding merits validation as we did not directly manipulate emotion in our study but rather relied on participants' ratings. Nonetheless, our results are consistent with prior findings showing high arousal at boundaries, which influences event segmentation (Clewett et al., 2020, 2025; Clewett & McClay, 2025; McClay et al., 2025; Riegel et al., 2023),

<u>Event segmentation and latent cause inference in anxiety versus PTSD</u>

Conceptualizing event segmentation within the latent cause inference framework, our findings are consistent with a recent study that found accurate latent cause inference in trait anxiety during fear learning. Zika et al. (2023) used a fear-reversal paradigm, in which two stimuli alternated in predicting threat across blocks of the task. Participants with anxiety showed better adaptation to changes between contexts (blocks), and a latent cause inference model that captured the two different contexts explained this behavior better than a model that assumed learning within a single latent cause.

Interestingly, while both our study and Zika et al. (2023) address anxiety symptoms more generally, studies of post-traumatic stress disorder (PTSD) show impaired latent cause inference (Cisler et al., 2024; Norbury et al., 2021) and event segmentation (Eisenberg et al., 2023; Pitts et al., 2022). This discrepancy between studies of anxiety and PTSD might reflect differences in the severity of symptoms of anxiety, as PTSD is also considered an anxiety disorder. Our sample included individuals with varying levels of symptoms, but we did not test a clinical sample that meets the diagnostic criteria for an anxiety disorder. It is possible that there exists a U-shaped relationship between symptoms of anxiety and segmentation, where mild anxiety improves segmentation due to higher vigilance and attention to boundaries, but severe (clinical-level) anxiety impairs segmentation. Alternatively, the different findings might reflect different mechanisms that are specific to PTSD, for instance, the interfering effects of intrusive memories. In PTSD, impaired latent cause inference could lead to impaired context segmentation and overgeneralization of fear. In contrast, in anxiety, we have suggested that accurate latent cause inference results in precise segmentation, which may contribute to the maintenance of anxiety by preventing safe experiences from updating fearful latent causes (contexts).

Implications for fear and safety learning and exposure therapy

We showed that symptoms of anxiety are associated with precise segmentation in complex events that unfold over time. While this suggests that overgeneralization might not underlie excessive fear in anxiety, studies that used simpler stimuli did report overgeneralization of fear in anxiety. For example, individuals with anxiety disorders show reduced discrimination between stimuli associated with fearful vs. safe outcomes and generalize fear across similar stimuli more than healthy controls (Cha et al., 2014; Dunsmoor & Paz, 2015; Lissek et al., 2010, 2014). Still, in extinction paradigms, where fear-learning and extinction (safety-learning) can be thought of as two different contexts, results regarding generalization in anxiety disorders have been mixed (Beckers et al., 2023; Duits et al., 2015). One speculation is that generalizing across stimuli vs. across complex temporal contexts reflects different levels of segmentation and generalization. Overgeneralization across stimuli within a context/latent cause might maintain anxiety by expanding the application of fear responses to safe stimuli, while rigid segmentation of contexts may prevent generalization from safe experiences in one context to other contexts.

Rigid context segmentation could therefore have implications for exposure therapy. Exposure therapy involves participants confronting fearful stimuli in a safe situation with no aversive outcomes, which should gradually reduce fear through learning of the association of the stimuli with safety (Brown et al., 2023; Seuling et al., 2024; Spyridonis et al., 2024). Exposure therapy is effective, but its effectiveness is limited (Brown et al., 2023). Precise segmentation in participants with social anxiety or anxiety more broadly might be one factor limiting the effectiveness of exposure therapy due to poor generalization from the therapy context to other scenarios.

Conclusion

Event segmentation is a core cognitive process essential for learning and generalization (Botvinick, 2008; Zacks, 2020). Here, we found that individuals with symptoms of anxiety, and specifically those with symptoms of social anxiety, process event boundaries more slowly and segment continuous streams of input into separate events more precisely. This latter result was robust to emotional responses to the stimuli and was mediated by the self-reported frequency of changes in context in participants' everyday life, which was inversely correlated with anxiety symptom severity. Future research could test the relationship between segmentation and generalization of fear and safety learning (Dunsmoor et al., 2018), extend findings to clinical samples, and compare segmentation in PTSD versus other anxiety disorders. Research could also examine the mechanisms underlying segmentation changes in anxiety disorders. Such research could facilitate a better understanding of learning and generalization in anxiety and contribute to developing better interventions.

Acknowledgements

This work was supported by NIMH T32 fellowship to O.B. (T32MH065214), Princeton Neuroscience Institute Summer Internship Program (NSF # 2150171), NIMH Silvio O. Conte Center Grant (P50MH136296), and John Templeton Foundation. We thank Adele Goldberg and Isabel Berwian for insightful feedback on the design and interpretation of the findings.

Data availability

The data and appendices will be publicly available upon publication.

Code availability

The analysis code will be publicly available upon publication.

Methods

Experiment 1: Story reading

Participants. We based our sample size on a power analysis using the effect size obtained in a similar online study, Experiment 2 from Gillan et al. (2016), which tested correlations between a cognitive task and mental health symptoms, and controlling for several nuisance covariates. The effect size for the coefficient of partial determination was $f^2 = .018$ (a “small” effect; Cohen, 1988; Lakens, 2013). For $p < .05$, testing one predictor and including age, gender, and stories group (see below) as additional covariates, implies 80% power at N=438 and 90% at N=586 (G*Power a-priori power analyses for an F-test, multiple regression: Fixed model, R^2 increase).

The participants were recruited online using Prolific January-June 2022. Aiming to reach a sample size of at least 586 participants after exclusions, we invited to the study all the participants who completed and passed the quality control in a mental health symptoms questionnaire portion of another study conducted in the lab at the time (Mirea, Shin et al., 2024). The quality control of the questionnaire data included no zigzag response pattern, no straight-line responses (namely, responding the same response for all questions), and no more than 2 mistakes in 10 attention check questions (Zorowitz et al., 2023). Attention check questions were questions embedded in questionnaires that were phrased similarly to other questionnaire items but had one obvious correct answer. For example, participants should endorse “When something good happens, it makes me think about all the times I traveled to the moon” with ‘definitely disagree’.

Participants were native English speakers from the United States, age 18 and above. The study was approved by the Institutional Review Board of Princeton University, and all participants provided informed consent. The total study duration was approximately 50 minutes per participant. Participants received monetary compensation for their time ($10 for 50 minutes; rate of $12 per hour), plus $1 for completion if they completed the study and $.5 for good performance (at least 75% accuracy in the questions about the stories and reading times within the expected range, see below). In addition, our custom experiment delivery software (NivTurk, Zorowitz et al., 2023) has bot-checking functionality built into it and rejects from the start participants who are likely not human.

Seven hundred and sixty-three participants completed the story reading task. Thirty-seven participants (5%) were excluded due to showing less than 75% accuracy in the questions about the stories. Of the remaining 726, an additional 15 participants were excluded for insufficient reading time data (see below), resulting in 711 participants. While most participants completed the task within 2 weeks of completing the mental health questionnaires, some participants were invited to the study approximately 3 months after they completed the questionnaires. These participants were therefore invited to another session to complete the questionnaires again. Twenty-six participants did not come back for this additional session. An additional 18 participants did not complete the demographics survey data and were excluded from the analysis. Five more participants were excluded due to outlier reaction times (more than 3-SD larger than the mean). Thus, our final sample for testing a correlation with anxiety was N=662 participants. This gave us a power of 80% to discover an effect of $f^2 = .0119$, and 93% to discover an effect of $f^2 = .018$ as in the analysis above.

Materials. Stories: We created six stories about familiar everyday events (zoo visit, going to the movies, school day, restaurant work, amusement park, vacation) with carefully controlled boundary and no-boundary sentences. Each story had two versions (A and B; total of 12 unique stories) that used identical or structurally similar sentences, with the surrounding context manipulated to create boundary versus no-boundary sentences. The stories are provided in Appendix 1. Detailed design and story development procedures are provided in Supplementary Information, with the rationale and story characteristics described here briefly.

To control for content in boundary and no-boundary sentences across participants, identical sentences appeared as boundaries in one story version but as no-boundaries in the other version, read by different participant groups. To control for content within participants but without including identical repetition, we used structurally similar sentences (same parts-of-speech order, different specific words) that appeared as boundaries in one story and no boundaries in a paired story within the same participant's version. This design enabled comparison of identical boundary/no-boundary sentences across participants while comparing similar boundary/no-boundary sentences within participants, controlling for sentence-level factors that might influence reading times.

Boundary sentences incorporated five types of contextual changes: spatial, temporal, character, causal, and goal boundaries, following established event segmentation literature (Pettijohn & Radvansky, 2016). Each participant read 30 boundary and 30 no-boundary sentences across the six stories. Stories used everyday language, ranged from 28-34 sentences (405-457 words), and included demographically diverse characters. Three pilot studies validated that boundary sentences were perceived as larger situational changes, rated as more unexpected, and produced slower reading times compared to no-boundary sentences (Supplementary Information).

<u>Self-report instruments for assessing mental health symptoms in online participants:</u> In a separate session prior to completing the story reading task, participants completed transdiagnostic questionnaires previously used in the Niv lab in studies of latent-cause inference, including the PID-5 (Personality Inventory of the DSM-5; Krueger et al., 2012) from

which we focus here on the anxiousness and social withdrawal subscales as measures of different aspects of non-social (e.g., ‘I am a very anxious person’) and social (e.g., ‘I keep my distance from people’) anxiety, respectively. Participants also completed the DAQ (Depressive Attributions Questionnaire; Kleim et al., 2011), which we used to control for symptoms of depressive thinking style, given the known comorbidly between anxiety and depression (Kaiser et al., 2021; Kaufman & Charney, 2000).

Procedure. Story reading task: The stories were presented one sentence at a time in the center of the screen. For each sentence, participants were asked to read the sentence and press the space button immediately when they were done reading the sentence. After each story, they were asked to answer three easy questions about the story to ensure that they had read the story. N=223 participants of the N=711 sample also rated subjective valence and arousal on a 9-point scale (Bradley & Lang, 1994, 2017) to report their emotional response to the stories (we control for this in the analysis, and this addition did not influence the results, see below and Results). Participants were told in advance that they would have to answer these questions, and they practiced the task (including the questions) using two short stories before completing the main task. Optional breaks were provided between stories.

One group of participants read version A of the stories, and another group read version B. The order of the stories was determined pseudo-randomly, ensuring that structurally similar stories did not immediately follow each other to mitigate repetition effects (as unlikely as they may be, given that these are different stories). Aiming to study individual differences and to minimize task-driven differences, we constructed one order for each version. Thus, story group

A (N=353) read version A of the stories in the following order: zoo, school day, amusement park, workday, movies, vacation. To have boundary and no-boundary sentences appearing in similar positions in the task across both groups, story group B (N=358) read the stories in an order matching the similar stories: movie, workday, vacation, school day, zoo, amusement park. This meant that if in version A the zoo story had a causal boundary sentence early on, the parallel movie story also had a causal boundary early on in version B.

After completing the story reading task, participants completed three standard tasks to measure general cognitive functions: (1) Matrix Reasoning task, a version of Raven's matrices measuring reasoning and executive function (Bilker et al., 2012; Chierchia et al., 2019); (2) Verbal association task to measure verbal fluency (Hartshorne & Germine, 2015), 20 questions version; (3) Symmetry span task to measure working memory (Kane et al., 2004; Unsworth et al., 2005). Data from the 3 tasks were used to control for the respective cognitive factors in the analysis. The story reading and cognitive tasks were coded using Inquisit Lab 6.6.1 (2021) and were run on participants' desktops using Inquisit 6 (2021).

Analysis. Generally, preprocessing of the data was performed using customized code in Python, and statistical analyses were implemented in R (R core Team, 2018) using the lm (stats package), lmer and glmer functions (lme4 package; Bates et al., 2015).

Reading time per syllable was calculated for each sentence, and responses shorter than 50ms/syllabus or longer than 500ms/syllable were removed from analyses (Pettijohn & Radvansky, 2016). Participants who had too few responses remaining after exclusion (fewer than 69%, which was 3 SD below the group mean) were removed from the analysis.

To estimate the boundary effect at the group level and to extract individual participants' boundary effect while accounting for potentially confounding effects that might have remained despite our closely controlled design, we entered reading time of the target sentences (scaled without mean centering) to a general linear mixed-effects model with the inverse Gaussian link function, accounting for response time distribution (Lo & Andrews, 2015). We included as fixed effects boundary (boundary/no-boundary sentence), story order, sentence order (per story), story group (A/B), emotion questions (with/without) and their 5-way interaction, as well as sentence identity (note that we could estimate the effect of sentence identity due to our design, because sentences appeared both as boundary and no-boundaries). In addition, we included boundary, story order, and sentence order as random slopes, accounting for the possibility that the progression of the study might affect each participant differently. This analysis step was performed on all participants who had full story reading task data (N=711). The random slope of the boundary effect per participant was extracted per participant and taken for group-level analysis.

To test whether we replicated the boundary effect using our stimuli and in an online sample, the effect of Boundary was estimated using model comparison to an identical model to above, excluding the fixed effect of Boundary. We report the BIC difference between the two models and significance of the difference using a Chi-square test (as these are nested models). For simplicity, we also compared mean reading time for boundary vs. no-boundary sentences using a within-sample two-tailed t-test.

To test our main hypothesis regarding the relationship between anxiety symptoms and reading time for boundaries, the boundary effect per participant (i.e., slopes from the mixed-effect model, see above) was taken as the predicted outcome values of a multiple linear regression with the predictors of anxiety score as our main variable of interest and gender, age, story group (A/B), and emotion questions (with/without) as factors of no interest. We first analyzed the general effect of anxiety, summing the two subscales of anxiousness and social withdrawal (PID-5). Since an effect of anxiety was revealed (see Results), we conducted two follow-up analyses: (1) an analysis controlling for cognitive functions by including the participants' scores from the three cognitive tasks (reasoning, verbal fluency, working memory) as additional factors of no interest in the multiple regression model; (2) an analysis evaluating different aspects of non-social symptoms of anxiety (anxiousness subscale) and social symptoms of anxiety (social withdrawal subscale), with both subscale as two regressors in one linear regression model, with gender, age, and story group. If analysis (2) was significant, we followed up with an additional analysis (3) including both subscales and the scores from the three cognitive tasks to control for cognitive functions.

We additionally examined whether boundary strength influenced the relationship between anxiety and boundary response. Using boundary ratings from an independent sample (see Supplementary Information), we calculated boundary strength per each pair of similar Boundary/No-boundary sentences (i.e., within-participant). We then median-split the pairs into strong and weak boundaries in each story version (A/B) based on the difference in boundary strength ratings for the two sentences. Next, we entered reading time of the target sentences (scaled without mean centering) into a general linear mixed-effects model with the inverse Gaussian link function, including as fixed effects boundary (boundary/no-boundary sentence), strength (strong/weak), story number (in the experiment), sentence number (per story), story group (A/B), emotion questions (with/without), and their 6-way interaction, as well as sentence identity. As random slopes, we included boundary, strength, and their interaction, story number, and sentence number. In this model, since Boundary and Strength are categorical (0/1) factors, the weak boundary effect is the effect associated with the factor of Boundary, whereas the strong boundary effect is the effect computed from adding the effects of Boundary and the interaction of Boundary and Strength. The random slopes of the weak and strong boundary effects were computed per participant and entered separately into linear multiple regression models to test for a relationship with anxiety, as done above for the analysis including all boundaries. Results of this analysis are briefly described in the main text, and detailed in Supplemental Information.

In all multiple regression models, outlier participants (segmentation measure or anxiety score more than 3 SD from the mean) were removed from the analysis. This resulted in excluding 5 participants from the model including all boundaries, and 4 and 9 participants from the models including weak or strong boundaries, respectively (participants were removed for outlier reading time data, no participant was removed due to anxiety score). Three additional participants did not have data in the cognitive tasks and thus were removed from the analysis controlling for these tasks.

We controlled for depression by first examining whether boundary response correlated with depression symptoms (based on the DAQ; Kleim et al., 2011) using similar regression models as for anxiety. Then, we incorporated depression scores in the same models with anxiety, testing whether anxiety symptoms still correlated with boundary response when controlling for depression symptoms.

Experiment 2: movie segmentation

Participants. The sample size and power analysis were identical to Experiment 1. The recruitment (during July and August 2022), invitation, and pay rate were also identical to Experiment 1. Participants were paid $15 per each 75-minute session ($12/h) of the segmentation task (see below) and could earn another $1 for completion if they completed the segmentation and response time tasks (see below) and $.5 for good performance (at least 75% accuracy in the questions about the movie clips that served as attention checks, completed all cognitive tasks, and scored above 85% accuracy in the symmetry span task, see below; Kane et al., 2004; Unsworth et al., 2005). Participants were paid $3 for the additional questionnaire session (which took ~10 minutes to complete; see below) and were given a bonus of $.3 for completing the entire questionnaire session.

Six hundred and sixty participants started the first movie segmentation session. Of these, 10 did not complete the session and 3 more did at least a part of it more than once due to a technical error. Of the remaining 647 participants, 640 participants passed a threshold of 75% accuracy in the attention checks and were invited to the second session and to conduct additional mental health questionnaires (if they hadn't already completed these in the last couple of weeks as a part of another study in the lab). We invited participants to do the additional mental health questionnaire one day after the first session and allowed them to complete the questionnaires up to 4 days after the first session. On the 5th day after the first session, participants were invited to the second session of the movie segmentation task, which they could complete up to 8 days after the first session. Invitations to the first session and to these additional questionnaires only told participants that they were invited because they did well in a previous study in our lab and did not include a reference to this specific study (multiple studies were run in parallel at the time this study was conducted). Invitations to the second session mentioned that this was the second session of the same study. This was done because the task on both days was similar, and participants knew that, in general, they should not complete the same experiment twice.

Of the 530 participants who started the second session (82% retention), 5 did not complete it. To ensure compliance with the task, we further excluded 1 participant who had below 75% accuracy in the attention checks during this session and 22 participants (4%) for having fewer fine compared to coarse segments, indicating poor compliance with the task (see below). One additional participant did not complete the response-time task either before or after the segmentation task in the second session and was excluded from the analysis. Of the remaining 501, 14 participants (2.6%) were removed due to insufficient data for computing segmentation typicality or hierarchical segmentation (see below), and 7 participants did not complete the response time task after either the first or second session; thus, data from these participants were corrected based only on the response-time task conducted before segmentation (see below).

Our primary anxiety measures were based on the STICSA-trait and LSAS questionnaires that were conducted in a separate session, together with additional questionnaires as detailed below. Three participants were removed for not passing quality control (done as detailed above: no zigzag or straight-line responses, and no more than 2 mistakes in attention checks (Zorowitz et al., 2023). Seventeen additional participants (3%) were removed due to outlier values (more than 3 SD from the group mean; 3 for the number of coarse segments, 7 for coarse segmentation typicality, 5 for hierarchical segmentation, and 2 for total anxiety score).

Since not all participants returned for the separate session of anxiety questionnaires, in sum, our sample for the main analysis included N=442 participants (177 women, 8 non-binary, 257 men; ages 18-73, *M* = 37.93, *SD* = 13.11) who completed all tasks and questionnaires, passed all quality control thresholds, and had demographic data allowing controlling for age and gender (one participant was removed for not having demographic data due to technical error). This gives us 80% power to discover an effect of f^2 = .018.

Materials. Movie clips for movie segmentation: Eight movie clips were taken from a previous segmentation study (Baldassano et al., 2018) or from the internet, and were piloted to ensure reasonable hierarchical segmentation at the group level and a range of prior knowledge regarding the typical unfolding of events in that context. For example, we included in the experiments clips of a date in a restaurant (high knowledge) vs. astronaut training (low knowledge). The 8 movies were: a fishing scene (A River Runs Through, 90 s), a supermarket scene (Home Alone, 146 s), a heist scene (Ocean's Eight, 168 s), a scene including the decoding of encrypted messages (The Imitation Game, 121 s), a laundry scene (Friends, 179 s), a law lecture (How to Get Away with Murder, 178 s), a date in a restaurant (Derek, 157 s), and an astronaut training scene (First Man, 113 s). Two additional short clips were included as practice clips and were not analyzed. Each movie clip was preceded by a 5 s countdown (Baldassano et al., 2018).

Self-report instruments for assessing mental health symptoms in online participants: We used several previously validated instruments to achieve excellent characterization of anxiety in online participants and address additional mental health symptoms as potential covariates/alternative explanations of the data: (1) the state and trait versions of the State-Trait Inventory of Cognitive and Somatic Anxiety (STICSA; Ree et al., 2008; Styck et al., 2022); this scale can separately capture aspects of cognitive anxiety (excessive worrying) and somatic anxiety (heightened arousal and physical stress symptoms such as sweating and heart beating); (2) the self-report version of the Leibowitz Social Anxiety Scale (LSAS-SR) for both fear and avoidance (Fresco et al., 2001; Heimberg et al., 1999; Rytwinski et al., 2009) was used to capture symptoms of social anxiety; (3) the Depressive Attributions Questionnaire (DAQ; Kleim et al., 2011) was used to assess symptoms of depression due to high rates of comorbidity with anxiety (Kaiser et al., 2021; Kaufman & Charney, 2000); (4) we composed a changes-in-life questionnaire querying the extent to which participants experience changes of situations in their everyday life and their preference for such changes (e.g., we asked participants whether or not during their day of work they are typically in the same room, and in a separate question, whether or not they like staying in one room all day); we added to this questionnaire the 'attention to switch' subscale from the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001); see Supplementary Information for more details and Appendix 2 for the questionnaire.

Procedure. Movie segmentation task: For each of the 8 movie clips (90-179 s each; total: 1152 s), participants indicated when, in their mind, one meaningful unit has ended, and another began (see details below). Participants performed this task twice, 5-8 days apart, first based on the largest meaningful units (coarse segmentation) and then based on the smallest meaningful units (fine segmentation; Sargent et al., 2013). This order and time gap were determined based on a previous study examining segmentation in PTSD patients (Eisenberg et al., 2016). After each clip, participants answered 2 simple questions about the clip that we used as attention checks (see exclusion criteria above). Participants also rated their subjective valence and arousal on a 9-point scale (Bradley & Lang, 2017) to evaluate their emotional response to the movie clips and to control for these in the analysis. Since knowledge influences event segmentation (Hard et al., 2006; Zacks, 2020; Zacks et al., 2001), participants were asked how familiar they were with the situation presented (e.g., eating in a restaurant; clip topics were taken as those provided by a majority of participants in a previous pilot where participants were asked to provide a topic for each clip), how often they experience this situation or were exposed to it, and how recently this was.

Participants watched the 8 movies in the order detailed above, which spread high- and low-knowledge clips (as judged in the previous pilot) across the experiment. The main clips were preceded by instructions and practice to ensure participants understood the segmentation task. Aiming at participants' subjective segmentation, as in prior work, we instructed participants to "press the space bar every time when, in your judgment, one meaningful and natural unit ends, and another begins. There are no right or wrong answers, we want to know how you do it" (Michelmann et al., 2021; Newtson, 1973; Sargent et al., 2013). Participants were instructed that one can segment movies based on large or small events, and then practiced both coarse and fine segmentation. Then, they were informed: "Today, you will indicate the largest units that seem meaningful to you" (in the second session, the word ‘largest’ was replaced by ‘smallest’). To ensure that online participants were attentive to the instructions and understood the level of segmentation in each session, we asked them to type in whether they would segment based on the largest or smallest units in the current session, and they were not allowed to continue the task until they typed the correct answer. Then, they practiced the segmentation task once more, based on the granularity of that session. This last practice also included the knowledge and emotion questions.

In each session, before and after the segmentation task, participants performed a short reaction-time task that measured their response to a fixation cross that changed to a circle (Zalla et al., 2004). We used this to adjust each participant’s segmentation responses, correcting for baseline reaction-time differences across participants. After the post-segmentation reaction-time task, participants completed (in each of the two sessions) the STICSA questionnaire (state version), to evaluate their state anxiety. This was followed by the 3 cognitive tasks, as in Experiment 1. The segmentation and cognitive tasks were coded using Inquisit Lab 6.6.1 (2021) and were run on participants’ desktops using Inquisit 6 (2021). The self-report questionnaires were administered using proprietary NivTurk software in JsPsych (Zorowitz & Bennett, 2022). Between the two sessions of the segmentation task, participants were invited to complete additional self-report questionnaires that included STICSA (trait version), our ‘changes in everyday life’ questions and AQ attention to switches subscale, and the LSAS-SR scale, as detailed above.

Analysis.

Segmentation measures. We assessed 5 segmentation measures: the number of fine/coarse segments, fine/coarse segmentation typicality (sometimes called segmentation “agreement”, we refer to it here as typicality as this might be more intuitive to readers not versed in event segmentation literature), and hierarchical segmentation. Following the same approach as in Experiment 1, these 5 measures were calculated per participant and movie clip, entered into a mixed-level linear model to extract a subject-level estimate of the measure while controlling for confounds, and then taken to a multiple regression model to correlate segmentation with anxiety. The number of segments was the number of times participants indicated a boundary, calculated separately for fine or coarse segmentation. As this was zero or positive by definition and right skewed, to adjust for non-normality we added 1 to the numbers of segments and log-transformed them. Segmentation typicality captures the extent to which a participant's segmentation correlated with the segmentation profile of the group, computed as the proportion of participants who identified a boundary in each second of a movie clip (Eisenberg et al., 2016, 2016; Zacks et al., 2001). This measure was calculated separately for fine and coarse segmentation, and included normalizing for each participant’s potential minimal and maximal correlation given their number of segments (see Supplementary Information for details). Hierarchical segmentation refers to the notion that fine-grained events should be encapsulated within coarse-grain events. If so, coarse and nearby fine boundaries should align more than by chance. To test this, we measured the distance between coarse boundaries and their nearest fine boundary and compared it to the distance expected by chance for the number of fine and coarse boundaries a participant indicated (Hard et al., 2006; Kurby & Zacks, 2011; Zacks et al., 2001). We log-transformed the values to account for non-normality and subtracted the log(distance) from the log(chance distance) such that larger numbers reflect more hierarchical segmentation (Kurby & Zacks, 2011; see Supplementary Information for details).

Prior to the main analysis, each of the 5 segmentation measures was entered into a mixed-level model including the movies and their order as fixed effects, an intercept per participant, and a slope per participant for the order of the movies. This was done to minimize the effect of movie order as much as possible, considering also that participants might show different effects of movie order. The intercept per participant was taken to reflect each participant's segmentation score, removing the effects of movie order and of each movie. These intercepts (per each measure) were then entered into the main group-level analyses testing for a relationship with anxiety (see below).

Anxiety measures. As a primary measure of anxiety, we summed the average scores of the STICSA and LSAS. We use averages and not sum scores of each scale to account for the different number of questions per instrument. For follow-up analyses, we used the sum score of LSAS as a measure of social anxiety symptoms, and the sum scores of cognitive and somatic subscales of the STICSA questionnaire as measures of cognitive symptoms and somatic symptoms of anxiety.

Main analyses: the relationship between segmentation and anxiety. Multiple linear regressions were conducted (using "lme4" package in R; Bates et al., 2015) per each of the 5 segmentation measures. As predictors, we included the anxiety score, as well as age and gender as confounds. A two-tailed t-test on the coefficient estimate of the anxiety score was conducted to determine if anxiety significantly correlates with segmentation.[1] To correct for 5 multiple comparisons in a way that appreciates the correlation structure in the data (the range of correlation between segmentation measure was -.27 to .61), we followed the step-down minimum p (sd.minP) procedure (Westfall & Young, 1993). This process was recommended for correction of multiple dependent measures because it was shown to maintain Type 1 error (the chance of false positive) while also maintaining statistical power (Blakesley et al., 2009). Briefly, this procedure removes the effect of interest (anxiety) from the data (in each measure), uses bootstrapping to produce a null distribution of p-values, and sets the adjusted p-value as the probability of obtaining a p-value (in the null distribution) that is smaller than or equal to the actual p-value. For segmentation measures where the sum score of anxiety was significant and survived correction for multiple comparisons, we followed up by running additional regressions (one for each segmentation measure), including each of the sum scores of each of the three subscales (STICSA-cognitive, STICSA-somatic, LSAS), to examine specificity to the different aspects of anxiety.

Control for cognitive functions. The multiple linear regressions testing the relationship between segmentation and anxiety were repeated, including participants' accuracy in each of the 3

[1] The preregistration (April, 2023) included a directional hypothesis whereby segmentation agreement (fine/coarse) and hierarchical segmentation are worse in anxiety. However, soon after the preregistration was submitted, and before data analysis, a new study was published (Zika et al., July 2023), which showed that anxiety correlated with *better* response to context changes in a fear-reversal learning paradigm. Given the new findings, we modified our planned analyses and performed two-sided t-tests for these measures (and for number of segments, as in the preregistration) following recommendations for updating analysis plans in Crüwell & Evans (2021; see also Lakens, 2024).

cognitive tasks mentioned above, controlling for reasoning and executive function, verbal fluency, and working memory (17 participants who did not have data for these tasks in either session were removed from this analysis).[2] In the working memory symmetry span task, participants were required to keep their accuracy above 85% in the additional 'processing' task (namely, the indication of symmetrical pattern or not; Kane et al., 2004; Unsworth et al., 2005). However, we did not remove participants who did not reach this threshold to be conservative with exclusions and in line with research showing that such exclusions do not affect the reliability of the main working memory score (for remembering the order and location of a sequence of squares on the screen; Đokić et al., 2018). This control analysis was conducted only for the segmentation measures that survived correction for multiple comparisons in the main analysis.

Control for influence of prior knowledge and emotion on segmentation. For each movie clip, participants rated their valence and arousal and three measures of prior knowledge (familiarity, frequency of exposure, and recency, see above), which we used to control for the possibility that these factors mediated the relationship between anxiety and segmentation. The prior knowledge measures showed strong correlations across participants (Pearson's r ~.80). Thus, we used only the familiarity ratings, because these had the most variance given the scale ranges from 1-100 vs. a 4-point scale for the frequency and recency questions. We controlled for knowledge and emotion within participant by including the ratings per movie in a mixed-

---

[2] The preregistration also included education level, based on self-report. We found that this was cumbersome, since, being a categorical factor with multiple options, it introduced many additional predictors to the model. Thus, we report the results without education level. Including education level did not change the results.

level analysis (per segmentation measure) and extracting an intercept and slope per participant. The intercepts reflect segmentation while controlling for knowledge and emotion within participants. We then included these intercepts in multiple regressions as in the main analyses, expecting anxiety to correlate with segmentation. Since knowledge and emotion ratings correlated (Person's r ~.30 with valence, and less so with arousal, Pearson's r ~-.10), we included knowledge and emotion measures together as predictors in the multiple regression models, to control for each other.

Control for depression: As in Experiment 1, we controlled for depression by first examining whether segmentation correlated with depression symptoms (based on the DAQ, Kleim et al., 2011) using the same regression models as with anxiety. Then, we incorporated depression scores in the same models with anxiety, testing whether anxiety symptoms still correlated with segmentation when controlling for depression symptoms.

The relationship between anxiety, segmentation, and changes in everyday life

Participants also completed our newly developed 'changes in everyday life questionnaire', which asked participants to self-report their preference for having changes in context or situations during their everyday lives, as well as their experience with changes of situations.

The full questionnaire and its assessment are found in the Supplementary Information. Briefly, we examined the correlation between items, divergence and convergence of our scale vs. the AQ 'attention to switch' subscale (Baron-Cohen et al., 2001), and iterative removal of items to reach optimal convergence (Cronbach's alpha). This resulted in the removal of some items. Since we constructed this questionnaire aiming for one psychological construct, we ran a bifactor analysis (Cucina & Byle, 2017; Rodriguez et al., 2016b, 2016a) that assumed the existence of one general factor underlying all items in addition to more specific factors. This enabled us to extract one general factor underlying individuals' experience with changes in situations in their everyday lives. Participants' score in this general factor was then correlated with anxiety and segmentation.

First, we were interested in whether anxiety influences everyday life experiences with situational changes. Thus, we included participants' general factor score as the predicted outcome in a multiple regression model including anxiety, age, and gender as predictors. Second, we were interested in whether participants' lived experiences influenced event segmentation. Thus, we included the general factor scores in a multiple linear regression with age and gender, predicting segmentation (only for measures surviving correction for multiple comparisons in the main analysis).

References

Abend, R. (2023). Understanding anxiety symptoms as aberrant defensive responding along the threat imminence continuum. *Neuroscience & Biobehavioral Reviews*, *152*, 105305. https://doi.org/10.1016/j.neubiorev.2023.105305

Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., Abdulmalik, J., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L. H., Bruffaerts, R., Cardoso, G., Cia, A., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., de Jonge, P., … the WHO World Mental Health Survey Collaborators. (2018). Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depression and Anxiety*, *35*(3), 195–208. https://doi.org/10.1002/da.22711

Antony, J. W., Hartshorne, T. H., Pomeroy, K., Gureckis, T. M., McDougle, S. D., & Norman, K. A. (2021). Behavioral, Physiological, and Neural Signatures of Surprise during Naturalistic Sports Viewing. *Neuron*, *109*, 1–14.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of Real-World Event Schemas during Narrative Perception. *Journal of Neuroscience*, *38*(45), 9689–9699. https://doi.org/10.1523/JNEUROSCI.0251-18.2018

Baldwin, D. A., & Kosie, J. E. (2021). How Does the Mind Render Streaming Experience as Events? *Topics in Cognitive Science*, *13*(1), 79–105. https://doi.org/10.1111/tops.12502

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, Articles*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine*, *43*, 897–910.

Beckers, T., Hermans, D., Lange, I., Luyten, L., Scheveneels, S., & Vervliet, B. (2023). Understanding clinical fear and anxiety through the lens of human fear conditioning. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-023-00156-1

Bein, O., & Davachi, L. (2024). Event integration and temporal differentiation: How hierarchical knowledge emerges in hippocampal subfields through learning. *The Journal of*

Neuroscience, *44*(10), JN-RM-0627-23. https://doi.org/10.1523/JNEUROSCI.0627-23.2023

Bein, O., & Niv, Y. (2025). Schemas, reinforcement learning and the medial prefrontal cortex. *Nature Reviews Neuroscience*, *26*(3), 141–157. https://doi.org/10.1038/s41583-024-00893-z

Ben-Yakov, A., & Henson, R. N. (2018). The Hippocampal Film Editor: Sensitivity and Specificity to Event Boundaries in Continuous Experience. *J. Neurosci.*, *38*(47), 10057–10068. https://doi.org/10.1523/JNEUROSCI.0524-18.2018

Ben-Yakov, A., Rubinson, M., & Dudai, Y. (2014). Shifting gears in hippocampus: Temporal dissociation between familiarity and novelty signatures in a single event. *J. Neurosci.*, *34*(39), 12973–12981. https://doi.org/10.1523/JNEUROSCI.1892-14.2014

Beukers, A. O., Collin, S. H. P., Kempner, R. P., Franklin, N. T., Gershman, S. J., & Norman, K. A. (2024). Blocked training facilitates learning of multiple schemas. *Communications Psychology*, *2*(1), 28. https://doi.org/10.1038/s44271-024-00079-4

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of Abbreviated Nine-Item Forms of the Raven's Standard Progressive Matrices Test. *Assessment*, *19*(3), 354–369. https://doi.org/10.1177/1073191112446655

Blakesley, R. E., Mazumdar, S., Dew, M. A., Houck, P. R., Tang, G., Reynolds, C. F., & Butters, M. A. (2009). Comparisons of methods for multiple hypothesis testing in neuropsychological research. *Neuropsychology*, *23*(2), 255–264. https://doi.org/10.1037/a0012850

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.*, *12*(5), 201–208. https://doi.org/10.1016/j.tics.2008.02.009

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the Semantic Differential. *J. Behav. Ther. Exp. Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (2017). International Affective Picture System. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_42-1

Brown, V. M., Price, R., & Dombrovski, A. Y. (2023). Anxiety as a disorder of uncertainty: Implications for understanding maladaptive anxiety, anxious avoidance, and exposure therapy. *Cognitive, Affective, & Behavioral Neuroscience*, *23*(3), 844–868. https://doi.org/10.3758/s13415-023-01080-w

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.*, *18*(4), 590–596. https://doi.org/10.1038/nn.3961

Brozovich, F., & Heimberg, R. G. (2008). An analysis of post-event processing in social anxiety disorder. *Clinical Psychology Review*, *28*(6), 891–903. https://doi.org/10.1016/j.cpr.2008.01.002

Brühl, A. B., Delsignore, A., Komossa, K., & Weidt, S. (2014). Neuroimaging in social anxiety disorder—A meta-analytic review resulting in a new neurofunctional model. *Neuroscience & Biobehavioral Reviews*, *47*, 260–280. https://doi.org/10.1016/j.neubiorev.2014.08.003

Cha, J., Greenberg, T., Carlson, J. M., DeDora, D. J., Hajcak, G., & Mujica-Parodi, L. R. (2014). Circuit-Wide Structural and Functional Measures Predict Ventromedial Prefrontal Cortex Fear Generalization: Implications for Generalized Anxiety Disorder. *Journal of Neuroscience*, *34*(11), 4043–4053. https://doi.org/10.1523/JNEUROSCI.3372-13.2014

Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019). The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, *6*, 190232. http://dx.doi.org/10.1098/rsos.190232

Cisler, J. M., Dunsmoor, J. E., Fonzo, G. A., & Nemeroff, C. B. (2024). Latent-state and model-based learning in PTSD. *Trends in Neurosciences*, *47*(2), 150–162. https://doi.org/10.1016/j.tins.2023.12.002

Claus, N., Takano, K., & Wittekind, C. E. (2023). The interplay between cognitive biases, attention control, and social anxiety symptoms: A network and cluster approach. *PLOS ONE*, *18*(4), e0282259. https://doi.org/10.1371/journal.pone.0282259

Clewett, D., & Davachi, L. (2017). The Ebb and Flow of Experience Determines the Temporal Structure of Memory. *Curr Opin Behav Sci*, *17*, 186–193. https://doi.org/10.1016/j.cobeha.2017.08.013

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*(3), 162–183. https://doi.org/10.1002/hipo.23074

Clewett, D., Gasser, C., & Davachi, L. (2020). Pupil-linked arousal signals track the temporal organization of events in memory. *Nat. Commun.*, *11*(1), 4007. https://doi.org/10.1038/s41467-020-17851-9

Clewett, D., Huang, R., & Davachi, L. (2025). Locus coeruleus activation 'resets' hippocampal event representations and separates adjacent memories. *Neuron*, *113*. https://doi.org/10.1101/2024.08.15.608148

Clewett, D., & McClay, M. (2025). Emotional arousal lingers in time to bind discrete episodes in memory. *Cognition and Emotion*, *39*(1), 97–116. https://doi.org/10.1080/02699931.2023.2295853

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Cooper, S. E., Van Dis, E. A. M., Hagenaars, M. A., Krypotos, A.-M., Nemeroff, C. B., Lissek, S., Engelhard, I. M., & Dunsmoor, J. E. (2022). A meta-analysis of conditioned fear generalization in anxiety-related disorders. *Neuropsychopharmacology*, *47*(9), 1652–1661. https://doi.org/10.1038/s41386-022-01332-2

Craske, M. G., Stein, M. B., Eley, T. C., Milad, M. R., Holmes, A., Rapee, R. M., & Wittchen, H.-U. (2017). Anxiety disorders. *Nature Reviews Disease Primers*, *3*(1), 17024. https://doi.org/10.1038/nrdp.2017.24

Cremers, H. R., & Roelofs, K. (2016). Social anxiety disorder: A critical overview of neurocognitive research: Social anxiety disorder. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(4), 218–232. https://doi.org/10.1002/wcs.1390

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3).

Crüwell, S., & Evans, N. J. (2021). Preregistration in diverse contexts: A preregistration template for the application of cognitive models. *Royal Society Open Science*, *8*, 210155.

Cucina, J., & Byle, K. (2017). The Bifactor Model Fits Better Than the Higher-Order Model in More Than 90% of Comparisons for Mental Abilities Test Batteries. *Journal of Intelligence*, *5*(3), 27. https://doi.org/10.3390/jintelligence5030027

Delpino, F. M., Da Silva, C. N., Jerônimo, J. S., Mulling, E. S., Da Cunha, L. L., Weymar, M. K., Alt, R., Caputo, E. L., & Feter, N. (2022). Prevalence of anxiety during the COVID-19 pandemic: A systematic review and meta-analysis of over 2 million people. *Journal of Affective Disorders*, *318*, 272–282. https://doi.org/10.1016/j.jad.2022.09.003

Đokić, R., Koso-Drljević, M., & Đapo, N. (2018). Working memory span tasks: Group administration and omitting accuracy criterion do not change metric characteristics. *PLOS ONE*, *13*(10), e0205169. https://doi.org/10.1371/journal.pone.0205169

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory*, *134*, 107–114. https://doi.org/10.1016/j.nlm.2016.07.011

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., & Engelhard, I. M. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, *32*, 239–253.

Dunsmoor, J. E., Kroes, M. C. W., Moscatelli, C. M., Evans, M. D., Davachi, L., & Phelps, E. A. (2018). Event segmentation protects emotional memories from competing experiences encoded close in time. *Nature Human Behaviour*, *2*(4), 291–299. https://doi.org/10.1038/s41562-018-0317-4

Dunsmoor, J. E., & Paz, R. (2015). Fear Generalization and Anxiety: Behavioral and Neural Mechanisms. *Biological Psychiatry*, *78*(5), 336–343. https://doi.org/10.1016/j.biopsych.2015.04.010

Eisenberg, M. L., Rodebaugh, T. L., Flores, S., & Zacks, J. M. (2023). Impaired prediction of ongoing events in posttraumatic stress disorder. *Neuropsychologia*, *188*, 108636. https://doi.org/10.1016/j.neuropsychologia.2023.108636

Eisenberg, M. L., Sargent, J. Q., & Zacks, J. M. (2016). Posttraumatic stress and the comprehension of everyday activity. *Collabra*, *2*(1), 11. https://doi.org/10.1525/collabra.43

Eisenberg, M. L., Zacks, J. M., & Flores, S. (2018). Dynamic prediction during perception of everyday events. *Cogn Res Princ Implic*, *3*(1), 53. https://doi.org/10.1186/s41235-018-0146-z

Ezzyat, Y., & Clements, A. (2024). Neural activity differentiates novel and learned event boundaries. *The Journal of Neuroscience*, e2246232024. https://doi.org/10.1523/JNEUROSCI.2246-23.2024

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychol. Sci.*, *22*(2), 243–252. https://doi.org/10.1177/0956797610393742

Ezzyat, Y., & Davachi, L. (2021). Neural Evidence for Representational Persistence Within Events. *The Journal of Neuroscience*, *41*(37), 7909–7920. https://doi.org/10.1523/JNEUROSCI.0073-21.2021

Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M., & Gershman, S. J. (2020). Structured Event Memory: A neuro-symbolic model of event cognition. *Psychol. Rev.*, *127*(3), 327–361. https://doi.org/10.1037/rev0000177

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., & Goetz, D. (2001). The Liebowitz Social Anxiety Scale: A comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine*, *31*(6), 1025–1035. https://doi.org/10.1017/S0033291701004056

Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychol. Rev.*, *117*(1), 197–209. https://doi.org/10.1037/a0017808

Gershman, S. J., Monfils, M.-H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *Elife*, *6*. https://doi.org/10.7554/eLife.23763

Gershman, S. J., Radulescu, A., Norman, K. A., & Niv, Y. (2014). Statistical computations underlying the dynamics of memory updating. *PLoS Comput. Biol.*, *10*(11), e1003939. https://doi.org/10.1371/journal.pcbi.1003939

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife*, *5*. https://doi.org/10.7554/eLife.11305

Grös, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment*, *19*(4), 369–381. https://doi.org/10.1037/1040-3590.19.4.369

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, *14*(7), 488–501. https://doi.org/10.1038/nrn3524

Hamm, A. O. (2019). Fear, anxiety, and their disorders from the perspective of psychophysiology. *Psychophysiology*, *57*(2), e13474. https://doi.org/10.1111/psyp.13474

Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Mem. Cognit.*, *34*(6), 1221–1235. https://doi.org/10.3758/bf03193267

Hartshorne, J. K., & Germine, L. T. (2015). When Does Cognitive Functioning Peak? The Asynchronous Rise and Fall of Different Cognitive Abilities Across the Life Span. *Psychological Science*, *26*(4), 433–443. https://doi.org/10.1177/0956797614567339

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R., & Liebowitz, M. R. (1999). Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine*, *29*(1), 199–212. https://doi.org/10.1017/S0033291798007879

Hengen, K. M., & Alpers, G. W. (2021). Stress Makes the Difference: Social Stress and Social Anxiety in Decision-Making Under Uncertainty. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.578293

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *J. Exp. Psychol. Learn. Mem. Cogn.*, *44*(7), 1075–1090. https://doi.org/10.1037/xlm0000503

Hoehn-Saric, R., & McLeod, D. R. (2000). Anxiety and arousal: Physiological changes and their perception. *Journal of Affective Disorders*, *61*(3), 217–224. https://doi.org/10.1016/S0165-0327(00)00339-6

Hofmann, S. G. (2007). Cognitive Factors that Maintain Social Anxiety Disorder: A Comprehensive Model and its Treatment Implications. *Cognitive Behaviour Therapy*, *36*(4), 193–209. https://doi.org/10.1080/16506070701421313

Horenstein, A., Butler, R. M., & Heimberg, R. G. (2018). Social Anxiety Disorder. In B. Olatunji (Ed.), *The Cambridge Handbook of Anxiety and Related Disorders* (1st ed., pp. 480–516). Cambridge University Press. https://doi.org/10.1017/9781108140416.018

Inquisit (Version 6). (2021). [Computer software]. https://www.millisecond.com

Kaiser, T., Herzog, P., Voderholzer, U., & Brakemeier, E. (2021). Unraveling the comorbidity of depression and anxiety in a large inpatient sample: Network analysis to examine bridge symptoms. *Depression and Anxiety*, *38*(3), 307–317. https://doi.org/10.1002/da.23136

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Kaufman, J., & Charney, D. (2000). Comorbidity of mood and anxiety disorders. *Depression and Anxiety*, *12*(S1), 69–76. https://doi.org/10.1002/1520-6394(2000)12:1+%253C69::AID-DA9%253E3.0.CO;2-K

Kleim, B., Gonzalo, D., & Ehlers, A. (2011). The Depressive Attributions Questionnaire (DAQ): Development of a Short Self-Report Measure of Depressogenic Attributions. *Journal of Psychopathology and Behavioral Assessment*, *33*(3), 375–385. https://doi.org/10.1007/s10862-011-9234-9

Klumpp, H., Angstadt, M., Nathan, P. J., & Phan, K. L. (2010). Amygdala reactivity to faces at varying intensities of threat in generalized social phobia: An event-related functional MRI study. *Psychiatry Research: Neuroimaging*, *183*(2), 167–169. https://doi.org/10.1016/j.pscychresns.2010.05.001

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine*, *42*(9), 1879–1890. https://doi.org/10.1017/S0033291711002674

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends Cogn. Sci.*, *12*(2), 72–79. https://doi.org/10.1016/j.tics.2007.11.004

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Mem. Cognit.*, *39*(1), 75–91. https://doi.org/10.3758/s13421-010-0027-2

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Lakens, D. (2024). When and How to Deviate From a Preregistration. *Collabra: Psychology*, *10*(1), 117094. https://doi.org/10.1525/collabra.117094

LeDoux, J. E., & Pine, D. S. (2016). Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *American Journal of Psychiatry*, *173*(11), 1083–1093. https://doi.org/10.1176/appi.ajp.2016.16030353

Lee, Y. I., Lee, D., Kim, H., Kim, M. J., Jeong, H., Kim, D., Glotzbach-Schoon, E., & Choi, S.-H. (2024). Overgeneralization of conditioned fear in patients with social anxiety disorder. *Frontiers in Psychiatry*, *15*, 1415135. https://doi.org/10.3389/fpsyt.2024.1415135

Levine, D., Hirsh-Pasek, K., Pace, A., & Michnick Golinkoff, R. (2017). A goal bias in action: The boundaries adults perceive in events align with sites of actor intent. *J. Exp. Psychol. Learn. Mem. Cogn.*, *43*(6), 916–927. https://doi.org/10.1037/xlm0000364

Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized Anxiety Disorder Is Associated With Overgeneralization of Classically Conditioned Fear. *Biological Psychiatry*, *75*(11), 909–915. https://doi.org/10.1016/j.biopsych.2013.07.025

Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., & Grillon, C. (2010). Overgeneralization of Conditioned Fear as a Pathogenic Marker of Panic Disorder.

American Journal of Psychiatry, *167*(1), 47–55. https://doi.org/10.1176/appi.ajp.2009.09030410

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Front. Psychol.*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

McClay, M., Rouhani, N., & Clewett, D. (2025). Negative emotional events retroactively disrupt semantic scaffolding of temporal-order memory. *Emotion*.

McClay, M., Sachs, M. E., & Clewett, D. (2023). Dynamic emotional states shape the episodic structure of memory. *Nature Communications*, *14*(1), 6533. https://doi.org/10.1038/s41467-023-42241-2

Mennin, D. S., Fresco, D. M., Heimberg, R. G., Schneier, F. R., Davies, S. O., & Liebowitz, M. R. (2002). Screening for social anxiety disorder in the clinical setting: Using the Liebowitz Social Anxiety Scale. *Journal of Anxiety Disorders*, *16*(6), 661–673. https://doi.org/10.1016/S0887-6185(02)00134-2

Michelmann, S., Hasson, U., & Norman, K. A. (2023). Evidence That Event Boundaries Are Access Points for Memory Retrieval. *Psychological Science*.

Michelmann, S., Price, A. R., Aubrey, B., Strauss, C. K., Doyle, W. K., Friedman, D., Dugan, P. C., Devinsky, O., Devore, S., Flinker, A., Hasson, U., & Norman, K. A. (2021). Moment-by-moment tracking of naturalistic learning and its underlying hippocampo-cortical interactions. *Nature Communications*, *12*(1), 5394. https://doi.org/10.1038/s41467-021-25376-y

Miller, J. D., Bagby, R. M., Hopwood, C. J., Simms, L. J., & Lynam, D. R. (2022). Normative data for PID-5 domains, facets, and personality disorder composites from a representative sample and comparison to community and clinical samples. *Personality Disorders: Theory, Research, and Treatment*, *13*(5), 536–541. https://doi.org/10.1037/per0000548

Mirea, D.-M., Shin, Y. S., DuBrow, S., & Niv, Y. (2024). The Ubiquity of Time in Latent-cause Inference. *Journal of Cognitive Neuroscience*, 1–13. https://doi.org/10.1162/jocn_a_02231

Modecki, K. L., Ryan, K. M., & Waters, A. M. (2023). Fear learning and extinction predicts anxiety in daily life: A study of Pavlovian conditioning and ecological momentary assessment. *Psychological Medicine*, *53*(11), 5301–5311. https://doi.org/10.1017/S0033291722002379

Newberry, K. M., & Bailey, H. R. (2019). Does semantic knowledge influence event segmentation and recall of text? *Mem. Cognit.*, *47*(6), 1173–1187. https://doi.org/10.3758/s13421-019-00926-4

Newberry, K. M., Feller, D. P., & Bailey, H. R. (2021). Influences of domain knowledge on segmentation and memory. *Mem. Cognit.*, *49*(4), 660–674. https://doi.org/10.3758/s13421-020-01118-1

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*(1), 28–38. https://doi.org/10.1037/h0035584

Nolden, S., Turan, G., Güler, B., & Günseli, E. (2024). Prediction Error and Event Segmentation in Episodic Memory. *Neuroscience & Biobehavioral Reviews*, 105533. https://doi.org/10.1016/j.neubiorev.2024.105533

Norbury, A., Brinkman, H., Kowalchyk, M., Monti, E., Pietrzak, R. H., Schiller, D., & Feder, A. (2021). Latent cause inference during extinction learning in trauma-exposed individuals with and without PTSD. *Psychological Medicine*, 1–12. https://doi.org/10.1017/S0033291721000647

Pettijohn, K. A., & Radvansky, G. A. (2016). Narrative event boundaries, reading times, and expectation. *Mem. Cognit.*, *44*(7), 1064–1075. https://doi.org/10.3758/s13421-016-0619-6

Pitts, B. L., Eisenberg, M. L., Bailey, H. R., & Zacks, J. M. (2022). PTSD is associated with impaired event processing and memory for everyday events. *Cognitive Research: Principles and Implications*, *7*(1), 35. https://doi.org/10.1186/s41235-022-00386-6

Pulcu, E., & Browning, M. (2019). The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences*, *23*(10), 865–875. https://doi.org/10.1016/j.tics.2019.07.007

R core Team. (2018). *R: A language and environment for statistical computing* [R]. R Foundation for Statistical Computing. https://www.r-project.org/

Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing Cognitive and Somatic Dimensions of State and Trait Anxiety: Development and Validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy*, *36*(03). https://doi.org/10.1017/S1352465808004232

Riegel, M., Granja, D., Amer, T., Vuilleumier, P., & Rimmele, U. (2023). Opposite effects of emotion and event segmentation on temporal order memory and object-context binding. *Cognition and Emotion*, 1–19. https://doi.org/10.1080/02699931.2023.2270195

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying Bifactor Statistical Indices in the Evaluation of Psychological Measures. *Journal of Personality Assessment*, *98*(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rytwinski, N. K., Fresco, D. M., Heimberg, R. G., Coles, M. E., Liebowitz, M. R., Cissell, S., Stein, M. B., & Hofmann, S. G. (2009). Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. *Depression and Anxiety*, *26*(1), 34–38. https://doi.org/10.1002/da.20503

Safren, S. A., Heimberg, R. G., Horner, K. J., & Schneier, F. R. (1999). *Factor Structure of Social Fears: The Liebowitz Social Anxiety Scale*.

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., Eisenberg, M. L., & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*(2), 241–255. https://doi.org/10.1016/j.cognition.2013.07.002

Sarna, N., Dar, R., & Mazor, M. (2025). *Biased and Inattentive Responding Drive Apparent Metacognitive Biases in Mental Health*. Computational Psychiatry Conference, Tübingen, Germany.

Sasmita, K., & Swallow, K. M. (2022). Measuring event segmentation: An investigation into the stability of event boundary agreement across groups. *Behavior Research Methods*, *55*(1), 428–447. https://doi.org/10.3758/s13428-022-01832-5

Sava-Segal, C., Richards, C., Leung, M., & Finn, E. S. (2023). Individual differences in neural event segmentation of continuous experiences. *Cerebral Cortex*. https://doi.org/10.1101/2022.09.09.507003

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nat. Neurosci.*, *16*(4), 486–492. https://doi.org/10.1038/nn.3331

Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, *26*(1), 3–8. https://doi.org/10.1002/hipo.22523

Seuling, P. D., Czernin, N. S., & Schiele, M. A. (2024). Virtual Reality exposure therapy in the treatment of public speaking anxiety and social anxiety disorder. *Neuroscience Applied*, *3*, 104074. https://doi.org/10.1016/j.nsa.2024.104074

Shin, Y. S., & DuBrow, S. (2021). Structuring Memory Through Inference-Based Event Segmentation. *Top. Cogn. Sci.*, *13*(1), 106–127. https://doi.org/10.1111/tops.12505

Sinclair, A. H., Manalili, G. M., Brunec, I. K., Adcock, R. A., & Barense, M. D. (2021). Prediction errors disrupt hippocampal representations and update episodic memories. *Proceedings of the National Academy of Sciences*, *118*(51), e2117625118. https://doi.org/10.1073/pnas.2117625118

Smith, M. E., Newberry, K. M., & Bailey, H. R. (2020). Differential effects of knowledge and aging on the encoding and retrieval of everyday activities. *Cognition*, *196*, 104159. https://doi.org/10.1016/j.cognition.2019.104159

Spokas, M. E., Rodebaugh, T. L., & Heimberg, R. G. (2007). Cognitive biases in social phobia. *Social Phobia*.

Spyridonis, F., Daylamani-Zad, D., & Nightingale, J. (2024). PublicVR: A virtual reality exposure therapy intervention for adults with speech anxiety. *Virtual Reality*, *28*(2), 105. https://doi.org/10.1007/s10055-024-00998-x

Styck, K. M., Rodriguez, M. C., & Yi, E. H. (2022). Dimensionality of the State–Trait Inventory of Cognitive and Somatic Anxiety. *Assessment*, *29*(2), 103–127. https://doi.org/10.1177/1073191120953628

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/BF03192720

Van Dam, N. T., Gros, D. F., Earleywine, M., & Antony, M. M. (2013). Establishing a trait anxiety threshold that signals likelihood of anxiety disorders. *Anxiety, Stress & Coping*, *26*(1), 70–86. https://doi.org/10.1080/10615806.2011.631525

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, *33*(4), 448–457. https://doi.org/10.1037/h0027806

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Vol. 279). John Wiley & Sons.

Yang, X., Fang, Y., Chen, H., Zhang, T., Yin, X., Man, J., Yang, L., & Lu, M. (2021). Global, regional and national burden of anxiety disorders from 1990 to 2019: Results from the Global Burden of Disease Study 2019. *Epidemiology and Psychiatric Sciences*, *30*, e36. https://doi.org/10.1017/S2045796021000275

Zacks, J. M. (2020). Event Perception and Memory. *Annu. Rev. Psychol.*, *71*, 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *J. Cogn. Neurosci.*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychol. Bull.*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29–58. https://doi.org/10.1037/0096-3445.130.1.29

Zalla, T., Verlut, I., Franck, N., Puzenat, D., & Sirigu, A. (2004). Perception of dynamic action in patients with schizophrenia. *Psychiatry Res.*, *128*(1), 39–51. https://doi.org/10.1016/j.psychres.2003.12.026

Zika, O., Wiech, K., Reinecke, A., Browning, M., & Schuck, N. W. (2023). Trait anxiety is associated with hidden state inference during aversive reversal learning. *Nature Communications*, *14*(1), 4203. https://doi.org/10.1038/s41467-023-39825-3

Zorowitz, S., & Bennett, D. (2022). *NivTurk* (Version v1.2-Prolific) [Zenodo]. https://doi.org/10.5281/zenodo.6609218

Zorowitz, S., Solis, J., Niv, Y., & Bennett, D. (2023). Inattentive responding can induce spurious associations between task behaviour and symptom measures. *Nature Human Behaviour*, *7*(10), 1667–1681. https://doi.org/10.1038/s41562-023-01640-7

Zwaan, R. A. (1996). Processing narrative time shifts. *J. Exp. Psychol. Learn. Mem. Cogn.*, *22*(5), 1196–1207. https://doi.org/10.1037//0278-7393.22.5.1196

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol. Bull.*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162