

Event Segmentation Structures Temporal Experience: Simultaneous Dilation and Contraction in Rhythmic Reproductions

Joan Danielle K. Ongchoco, Tristan S. Yates, and Brian J. Scholl
Department of Psychology, Yale University

We experience the world in terms of both (continuous) time and (discrete) events, but time seems especially primitive—since we cannot perceive events without an underlying temporal medium. It is all the more intriguing, then, to discover that event segmentation can itself influence how we perceive the passage of time. We demonstrated this using a novel “rhythmic reproduction” task, in which people listened to irregular sequences of musical tones, and then immediately reproduced those rhythmic patterns from memory. Each sequence contained a single salient (and entirely task-irrelevant) perceptual event boundary, but the temporal placement of that boundary varied across multiple trials in which people reproduced the same underlying rhythmic pattern. Reproductions were systematically influenced by event boundaries in two complementary ways: tones immediately following event boundaries were delayed (being effectively played “too late” in the reproductions), while tones immediately preceding event boundaries were sped up (being effectively played “too early”). This demonstrates how event segmentation influences time perception in subtle and nonuniform ways that go beyond global temporal distortions—with dilation across events, but contraction within events. Events *structure* temporal experience, facilitating a give-and-take between the subjective expansion and contraction of time.

Public Significance Statement

Our experience of the world unfolds in time, but not always at the same subjective rate: time can seem to pass more quickly or slowly at different moments. In this study, we show that one of the drivers of such effects is the degree to which we experience the world in terms of discrete bounded *events*: durations containing a perceived boundary between two events can seem longer, while durations within two boundaries can seem shorter. This may help to explain common experiences of the plasticity of time—as when an hour sitting in a waiting room seems to last forever, while the same hour at an amusement park seems to fly by.

Keywords: event segmentation, time perception, rhythmic reproduction, auditory perception

Supplemental materials: <https://doi.org/10.1037/xge0001447.supp>

Perhaps the most salient fact about time from the perspective of physics is how immutable it is: the rate at which time passes seems like an inviolable constant, and nothing we do can speed it up or slow it down. In light of this, it is striking that perhaps the most salient fact about time from the perspective of psychology is how plastic it is: while the rate at which time passes may be constant, the rate at which we *perceive* it to pass can vary dramatically. We are all intuitively familiar with this: an hour spent in a waiting room seems to drag on, while an hour spent at an amusement park seems to flash by before you know it. (Indeed, such

experiences are so familiar that they have become embodied in common idioms: “Time flies when you’re having fun.”) Accordingly, psychologists have identified a host of factors that systematically influence the subjective experience of time (for reviews, see [Eagleman, 2008](#); [Grondin, 2010](#); [Matthews & Meck, 2014](#)). For example, time seems to dilate (with durations seeming longer) when stimuli are more complex (e.g., [Block, 1978](#)), more attention-grabbing (e.g., [Tse et al., 2004](#)), or less predictable (e.g., [Pariyadath & Eagleman, 2007](#)); when problems are less difficult (e.g., [Wöllner & Hammerschmidt, 2021](#));

This article was published Online First September 28, 2023.

Joan Danielle K. Ongchoco  <https://orcid.org/0000-0001-5839-651X>

Brian J. Scholl  <https://orcid.org/0000-0003-3610-0890>

This work was presented at the 2020 Object Perception, Attention, and Memory (OPAM) meeting.

For helpful comments, we thank Michael Bronstein, Ty Cannon, William Koller, Laurie Paul, and members of the Yale Perception and Cognition Laboratory. This project was funded by ONR MURI Grant N00014-16-1-2007 awarded to Brian J. Scholl.

Joan Danielle K. Ongchoco served as lead for data curation, formal analysis, methodology, and visualization. Tristan S. Yates served in a

supporting role for formal analysis, methodology, and visualization. Brian J. Scholl served as lead for funding acquisition, project administration, supervision, and validation and served in a supporting role for formal analysis, methodology, and visualization. Joan Danielle K. Ongchoco, Tristan S. Yates, and Brian J. Scholl contributed equally to conceptualization, writing—original draft, writing—review and editing, and investigation.

Correspondence concerning this article should be addressed to Brian J. Scholl or Joan Danielle K. Ongchoco, Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, United States. Email: brian.scholl@yale.edu or joan.ongchoco@psych.ubc.ca

or even just when people are paying more attention to the passage of time in the first place (e.g., Macar et al., 1994).

From Time to Events (and Back?)

Many of these factors (e.g., complexity or familiarity) seem relatively unrelated or extrinsic to temporal perception. Here, in contrast, we explore a factor that seems especially foundational to temporal experience: the perception of discrete *events*. What we perceive is so often not raw temporal durations themselves, but rather the individual events that occupy those durations: *this* happens, then *that* happens, and so forth. And these events structure our memory as well: you might recall this morning in terms of discrete temporally extended events such as eating breakfast, taking a shower, or brushing your teeth—rather than as a single continuously unfolding stream of input and activity. Indeed, discrete events seem so powerful and prevalent in our experience that we may sometimes perceive them as “everyday hallucinations,” even when there are no sensory cues to event boundaries present in the stimuli themselves (as in the phenomenon of temporal “scaffolded attention”; Ongchoco & Scholl, 2022).

How does event segmentation influence the perception of time? This might initially seem like an odd question: it is natural to think of time as the underlying medium which must exist prior to the construction of event representations. (How could you perceive a temporally extended event without perceiving it as somehow arrayed in time in the first place?) It has thus been especially provocative to discover that event segmentation can influence the subjective passage of time. Indeed, a central reason why the hour in the waiting room feels longer than the hour at the amusement park might be precisely because the latter contains so many clearer event boundaries. (At the amusement park, you ride the Ferris wheel, then the roller coaster, then take a break for cotton candy, etc. But in the waiting room, waiting is just followed by ... more waiting ... with nothing but a ticking clock to break up the monotony.)

Empirical research has also explored the impact of event segmentation on the subjective passage of time. In general, when recalling an experienced duration, event boundaries serve as helpful temporal markers for when and how much *change* has occurred (e.g., Poynter, 1989). Some studies suggest that discrete event boundaries dilate perceived time. For example, durations are reproduced as longer when an interval contains different kinds of tasks, compared to when the same interval contains only one kind of task (e.g., Block & Reed, 1978; Poynter & Homa, 1983), or when subevents are dissimilar compared to when they are repeated (e.g., Faber & Gennari, 2015). Other studies, however, suggest that event boundaries can also contract perceived time. For example, the remembered duration of an object moving on a complex trajectory through a display will be shorter when that motion is perceived to have multiple independent trajectory segments (e.g., Liverence & Scholl, 2012). These various studies on events and time are not necessarily in conflict with each other, since they used very different types of stimuli and subtly different response measures—but they nevertheless collectively paint a confusing picture of how event segmentation may influence the experience of time.

The Current Study: Using Rhythmic Reproductions to Reveal How Events Structure Time

To explore how events might structure our experience of time, we employed a novel reproduction task using sequences of simple

musical tones. When listening to music, we never hear a continuous, unbroken stream of notes; rather, we hear those notes as packaged into discrete temporal phrases, segmented by various properties such as pitch. Imagine, for example, hearing an irregular rhythm of many tones, with a sudden pitch shift—such that the first several tones play at a higher pitch, and the next several tones play at a lower pitch (see Figure 1). You can still hear the individual tones, but you will also *experience* them as two discrete musical phrases—that is, two events (e.g., Ongchoco & Scholl, 2020). (Of course, many other segmentation cues are also at play in actual music—including shifts in tempo, volume, lyrics, et al. Here our goal was to effectively titrate a form of event segmentation in simple temporal sequences down to its essence, in order to more carefully manipulate the nature and timing of the event boundaries.) In our experiments, subjects heard sequences of this type, with the timing of the tones always being irregular. Critically, subjects heard the same sequences multiple times throughout each experiment, but with the particular temporal placement of the event boundary varied systematically (as depicted in Figure 1).

Past studies of event segmentation and perceived duration have typically had people simply reproduce a single duration, when it either did or did not contain one or many event boundaries (as in Liverence & Scholl, 2012). But this “all-or-nothing” approach may obscure a far richer story in terms of the local temporal dynamics that may occur just before an event boundary, at an event boundary, and just after an event boundary—all of which may be mixed together when considering reproductions only of long monolithic durations (with some local effects being additive, and others potentially canceling each other out). Accordingly, in the current study, we used a task that was both richer and more natural—and which was also much less explicitly *about* the passage of time, per se: after hearing a sequence of tones, subjects simply reproduced its rhythm, by immediately tapping it out with a key on their keyboard. This can in principle reveal both local dilation *and* local contraction of perceived time, perhaps at different moments relative to the event boundary.

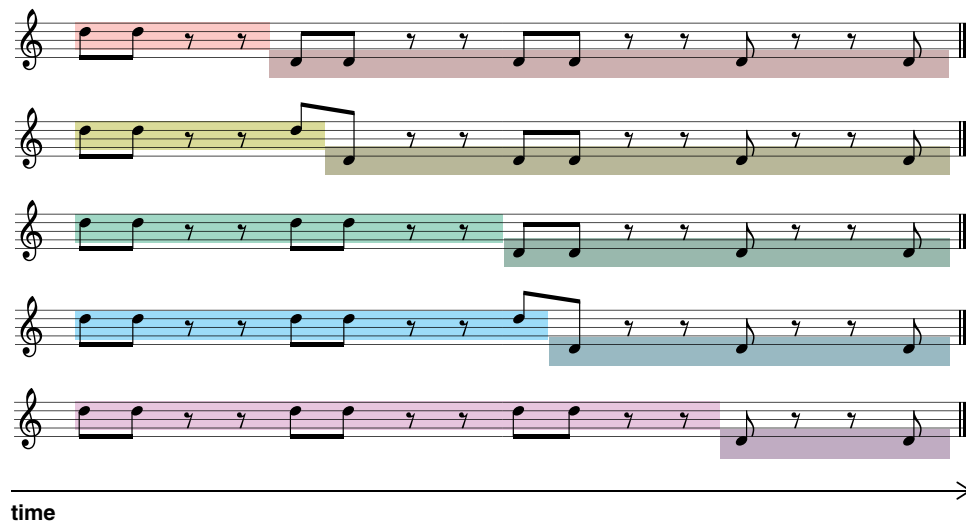
Such temporal distortions were measured by comparing the timing of each initial tone with the timing of each subsequent keypress—always normalizing to the start times of their reproductions. Positive errors thus reflected tones that were reproduced later than in the original sequence, while negative errors reflected tones that were reproduced earlier than in the original sequence. To visualize and quantify these effects, we subtracted out the previous errors from each subsequent reproduction, so that the errors did not accumulate. If reproductions are reasonably accurate and are unbiased by the event boundary, then we might see errors randomly occurring as positive or negative, but always close to zero (as in the first graph of Figure 2a). If reproductions are biased to be either uniformly dilated or contracted (as in previous work), then we would see either negative or positive errors across all tones (as in the second and third graphs of Figure 2a). However, if reproductions reflect a more subtle ebb and flow in perceived time, as a function of the specific temporal placement of the event boundary, then we might instead expect to see some parts of the sequences contracted and other parts dilated (as in Figure 2b).

We explored such reproductions in the context of event boundaries that were implemented via either a sudden shift in pitch octave (Experiment 1), a sudden shift in timbre (Experiment 2), or a change in the rhythmic structure itself (Experiment 3).

Figure 1

An Example of Event Segmentation in a Single Rhythmic Sequence, Depicted in Standard Musical Notation, With Time Proceeding to the Right

Varying the Boundary Position in a Sample Rhythmic Sequence



Note. Small elongated circles with stems depict musical notes (tones), with higher circles representing higher pitches. Small inverted “checks” between the notes represent silent periods (or musical “rests”). After a certain number of tones, an event boundary is introduced in the form of a sudden change in pitch, as in Experiment 1. Each row depicts the same sequence, with the event boundary placed at one of five different moments (corresponding to five different trials), and with the two resulting rhythmic phrases (“events”) depicted by different colors of shading. See the online article for the color version of this figure.

Experiment 1: Event Boundaries via Pitch Changes

On each trial, subjects listened to a sequence of eight musical tones, each of which had a single event boundary, implemented via a pitch change: as depicted in Figure 1, several notes were played at one pitch, followed by several notes at a different pitch. This sort of sudden change causes subjects to hear the tones as packaged into two discrete phrases (e.g., Ongchoco & Scholl, 2020). Each rhythmic sequence was repeated multiple times across trials, varying only where the event boundary was introduced. After listening to each rhythmic sequence, subjects simply reproduced the timing of each subsequent tone by repeatedly pressing a key.

Method

Subjects

Thirty subjects from the Yale and New Haven community participated in exchange for monetary payment or class credit. This sample size (determined from pilot testing) was preregistered prior to data collection and was fixed to be identical across all three experiments reported here. Subjects were asked their age and their sex in a free-response postexperiment debriefing phase ($M_{\text{age}} = 23.9$ years; 57% female, 43% male).

Apparatus

Stimuli were presented using custom software written in Python with the PsychoPy libraries (Peirce et al., 2019) and were displayed

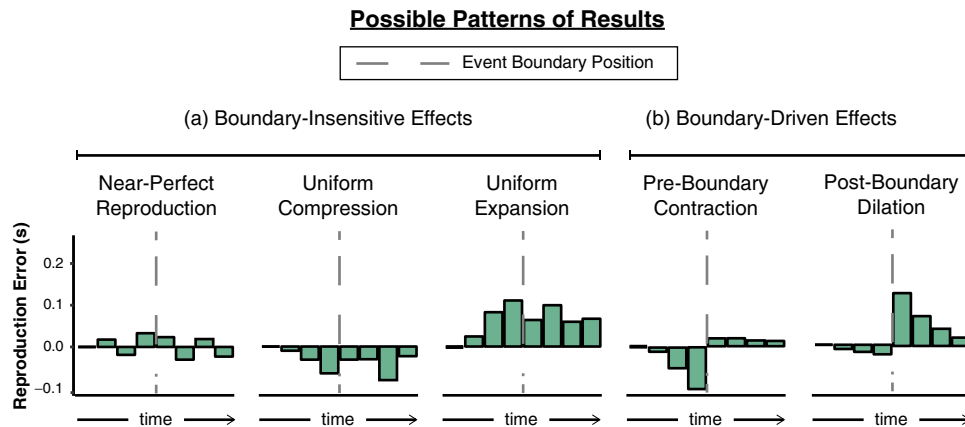
on a monitor with a 60 Hz refresh rate. Subjects sat in a dimly lit room without restraint approximately 60 cm from the display (with all visual extents reported below based on this approximate viewing distance). The functional part of the display subtended $34.87^\circ \times 28.21^\circ$. Subjects listened to the rhythmic sequences at a fixed volume (50% of the maximum possible volume) through headphones.

Stimuli and Conditions

For each subject, a unique set of 16 “base rhythms” was generated, where each base rhythm consisted of sixteen 250-ms periods (for a total of 4 s), into which eight tones and eight silent periods (“rests”) were randomly distributed (with the single constraint that the first period had to include a tone). (In musical terms, each base rhythm thus consisted of eight eighth notes and eight rests, randomly distributed across two 4/4 time-signature measures, where the first eighth note always began on the first downbeat of the sequence.) Tones were rendered using the “Marimba” instrument setting in MuseScore software, played (with an initial “attack” followed by a slow decay) at one of two pitches—either a high (587.33 Hz or D5) or a low (293.66 Hz or D4) pitch.

Each rhythmic sequence began with one of the two pitches (randomly determined on each trial) and then switched at some point to the other pitch (i.e., switching from a high to a low pitch, or vice versa). This sudden switch occurred after two, three, four, five, or six tones. These five conditions (varying only in the temporal placement of the event boundary) were implemented for each of the 16 base rhythms, such that

Figure 2
Possible Patterns of Results for Subjects' Reproduction Errors



Note. Positive values indicate temporal dilation (with a latency between reproduced tones being longer than the true latency) whereas negative values indicate temporal contraction (with a latency between reproduced tones being shorter than the true latency). (a) Results may be boundary-insensitive, as in the case of a near-perfect reproduction, or uniform compression or dilation that does not vary depending on the position of the event boundary. (b) Alternatively, subjects may show systematic temporal distortions that are not uniform across the trial—for example with preboundary contraction or postboundary dilation. See the online article for the color version of this figure.

subjects effectively reproduced each of the 16 rhythms five times, but with a different event boundary placement each time.

Procedure and Design

Each trial began with a single-tone sequence (along with a static black visual speaker/sound icon on a white background). Immediately after the sequence finished, the icon was replaced by the response prompt (centered, black, 0.5° height, written in Monaco font): “Use the spacebar to reproduce the rhythm.” On each subsequent keypress, a blue feedback bar (0.5° in height, with its center placed 2° below the response prompt) increased its width by 0.5° (growing from the center). After eight keypresses, subjects were asked (with a visual prompt) to rate their performance (“How close was your reproduction to the original?”) by typing a number from 1 (*not at all*) to 5 (*near-perfect reproduction*). After a blank 2-s interval, the next trial began.

Subjects completed 10 practice trials (each with a different randomly generated rhythm, and with the pitch change occurring at a random time), the results from which were not recorded. They then completed 80 experimental trials (16 base rhythms \times 5 event boundary conditions), presented in a different random order for each subject, with a short self-timed break after every 20 trials. Finally, subjects completed a debriefing procedure during which they were asked about their basic demographic information, such as their age and gender. They were also asked about their musical and dance experience (in terms of years of training), what they thought the experiment was testing, whether they noticed anything about the experiment (e.g., repeating rhythms), and any other strategies they might have used.

Transparency and Openness

All raw data (for this experiment, and all other experiments reported here) are available in the [online supplemental materials](#). The

preregistered methods and analyses for each experiment can be viewed at: <https://aspredicted.org/blind.php?x=xh5us2> (for Experiment 1), <https://aspredicted.org/blind.php?x=9562dc> (for Experiment 2), and <https://aspredicted.org/blind.php?x=y9n9a6> (for Experiment 3).

Results

Per the preregistration, reproductions were excluded if (a) subjects reported a reproduction rating below 3 for that particular trial ($M = 20.17$, $SD = 12.54$), and (b) the reproduction duration was 2 SD s greater than the subject's mean reproduction duration, after the initial below 3 rating exclusions ($M = 1.93$, $SD = 1.04$). The start time of each reproduction was first standardized to 0 (i.e., by subtracting any initial delay before the first keypress, such that the first keypress had no error by definition).

Dilation Across Events

The reproduction error for each tone after the initial keypress was computed by subtracting its true latency (i.e., the elapsed time between the current tone's onset and the previous tone's onset) from its reproduced latency (i.e., the elapsed time between the subject's current keypress and the previous keypress). (This procedure implemented a kind of normalization, such that errors did not accumulate over time, and that each tone's computed error reflected only the *incremental* contraction or dilation introduced by that tone's reproduction alone. If we had simply compared the raw timestamps for each actual vs. reproduced tone, in contrast, then a long or anomalous error for an early tone would be effectively “inherited” by each subsequent keypress, which would in turn mask the effects of interest.) As such, a negative error for a given tone indicates that the subject effectively pressed the key “too early” (relative to the previous keypress, indicating temporal contraction), whereas a positive error indicates that they pressed the key “too late” (indicating temporal dilation).

The average reproduction errors for each of the eight tones across the five event boundary conditions are depicted in Figure 3a, with the temporal placement of the event boundary drawn as a dashed vertical line. Inspection of this figure reveals a single salient pattern: the greatest change in the reproduction error (an increase in temporal dilation) systematically occurred for the tone immediately after the event boundary. This can be appreciated by the fact that the bar to the right of the dashed line in the graph is always far taller than the bar immediately to the left of the dashed line—with that dramatic difference effectively “moving” later as the event boundary also moves later. And it can also be appreciated by inspecting Figure 4a, in which the pairwise temporal placement of the greatest change in reproduction error is graphed as a function of the temporal placement of the event boundary, in the form of a heatmap (simply summing the number of trials in each cell): in this graph, the salient diagonal line indicates that the greatest difference reliably spanned the event boundary, for five out of the five event boundary conditions (e.g., so that when the pitch change occurred at the fourth tone in the sequence, the greatest difference in reproduction error spanned tones 3 and 4). In other words, event boundaries introduced anomalous reproduction delays, specifically for the next tone.

These impressions were confirmed with the following analyses. The temporal placement of the event boundary reliably predicted the adjacent pair of reproduction errors that had the greatest difference, computed simply as the reproduction error of the latter item in the pair, subtracting the reproduction error of the former item in the pair ($\beta = .12$, $p < .001$). Additionally, when we shuffled the order of each subject’s keypresses for each trial 1,000 times, the probability that the majority of tone pairs with the greatest difference spanned the boundary for all five event boundary conditions was 0/1,000 ($p < .001$). In fact, on average, only 0.85 out of the five event boundary conditions showed the same pattern as did the actual data (i.e., where the majority of tone pairs with the greatest difference spanned the event boundary). And perhaps most important, our primary pre-registered analysis revealed that reproduction errors were reliably greater for tones occurring immediately after the event boundary than for tones occurring immediately prior to the event boundary, $t(29) = 4.73$, $p < .001$, $d = 0.86$. (This pattern was true for 26/30 subjects [two-tailed binomial test, $p < .001$], and did not significantly correlate with musical or dance experience [musical training: $r = -.25$, $p = .181$; dance training: $r = -.01$, $p = .940$].)

Contraction Within Events

Given these results, we can further ask whether this systematic time dilation at event boundaries was accompanied elsewhere in the sequence by any corresponding temporal contraction, as a form of effective compensation. In other words, given that there is extra (dilated) time being added to the sequence, is that “extra” time being systematically deducted from other intervals in the reproduced sequence? (For example, the tone immediately prior to the event boundary could be played too early in the reproductions, or all of the reproductions could be effectively “repelled” from the event boundary—perhaps as a function of the distance to the boundary.) A quick inspection of Figure 3 suggests some temporal contraction within the events, especially for the tones *before* the event boundary. In particular, the two reproduction errors just before the event boundary are on average *below* 0—demonstrating contraction on average for five out of the five event boundary conditions.

To confirm these initial impressions, though, we must first somehow equate the total duration of each reproduced sequence. Thus, we standardized subjects’ reproductions to have the same overall durations as the actual sequences—effectively stretching or squeezing the reproductions as wholes, while retaining all of the relative tone-by-tone differences. To distinguish this from the previous analyses, we refer to these as the *duration-normalized* reproduction errors. These duration-normalized reproduction errors essentially ensure that the “compensatory” logic mentioned above must apply: any temporal dilation *must* be “paid for” with some corresponding temporal contraction elsewhere.

Given the systematic dilation effect that we observed, the key question then becomes just *where* the compensatory contraction may be systematically located. Contraction with respect to the event boundary in the first event (i.e., before the pitch change) would manifest as *negative* duration-normalized reproduction errors, as reproduced tones would be played earlier in time, being effectively repelled by the later event boundary. In contrast, contraction with respect to the event boundary in the second event (i.e., after the pitch change) would manifest as *positive* duration-normalized reproduction errors, as reproduced tones would be played later in time, being effectively repelled by the past event boundary. Note that neither of these systematic patterns is foreordained by our duration-normalization process: the compensatory contraction in this analysis must exist, but it need not be *systematically* localized to any particular set of tones—and so for example, it could be concentrated to only the intertone intervals immediately before, or immediately after the postboundary interval—or it could be more evenly distributed across the intertone intervals in both the first and second events, and so forth.

Without yet taking the distance to the boundary into account, duration-normalized reproduction errors in the first event were reliably negative, $t(29) = 2.31$, $p = .028$, $d = 0.42$, but duration-normalized reproduction errors in the second event were only marginally positive at best, $t(29) = 1.75$, $p = .090$, $d = 0.32$. However, these patterns became clearer once we took the distance of each tone to the boundary into account: the magnitude of the duration-normalized contraction was a function of the distance to the event boundary for both the first (i.e., preboundary) event (with reproduction errors becoming more negative when they approached the event boundary: Pearson’s $r = -.18$, $p = .025$; see the left side of Figure 5a) and the second (i.e., postboundary) event (with reproduction errors becoming more negative as they receded from the event boundary: Pearson’s $r = -.19$, $p = .023$; see the right side of Figure 5a).¹

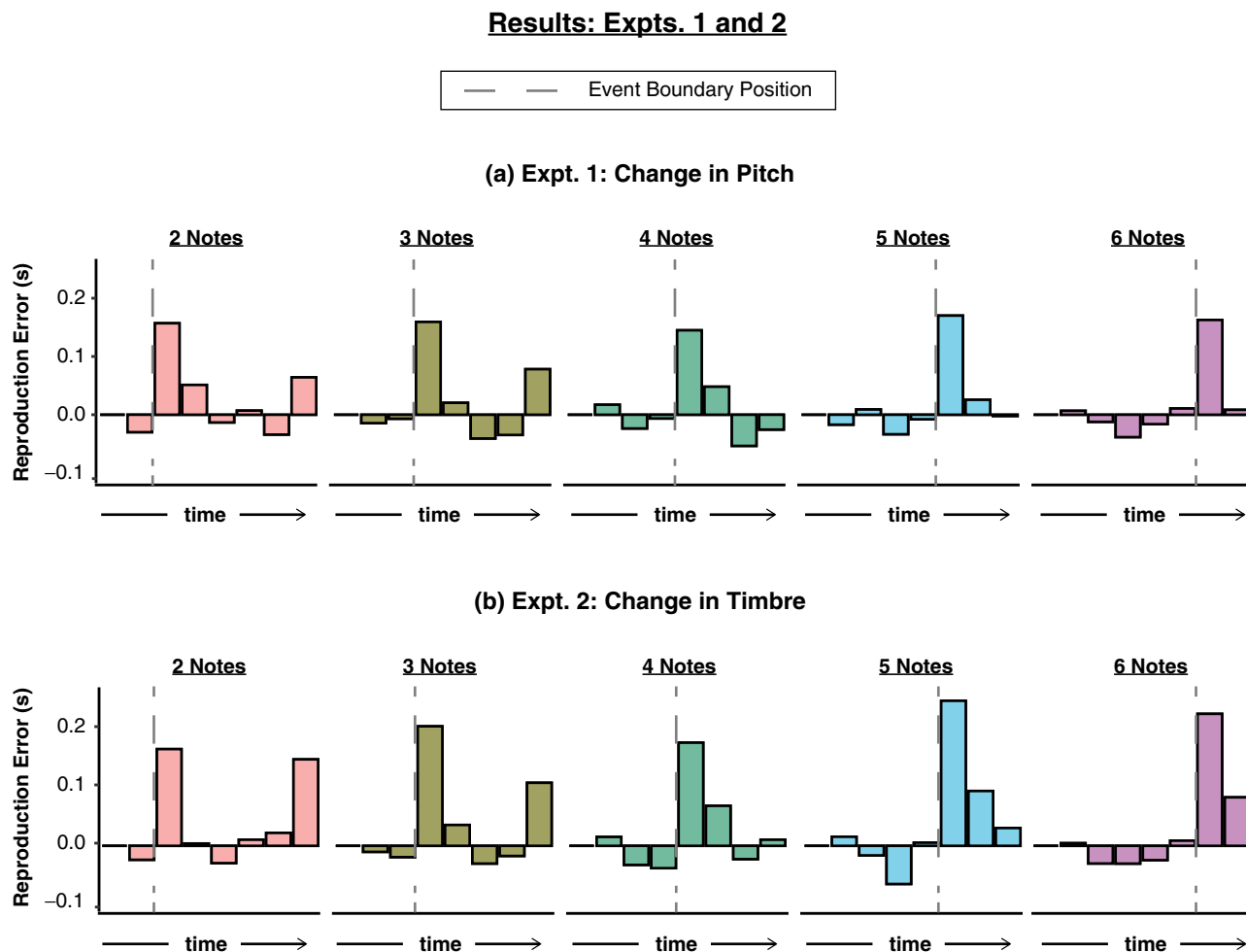
Discussion

In this initial experiment, event segmentation influenced time perception in subtle and nonuniform ways that went beyond global temporal distortions: there was temporal dilation across events (with

¹ When we ran the same contraction analyses on the raw reproduction errors—that is when they were not yet normalized for overall duration—we observed the same general pattern. The magnitude of contraction in the raw reproduction errors was a function of the distance to the event boundary for the second (i.e., postboundary) event (with reproductions errors becoming more negative as they receded from the event boundary: Pearson’s $r = -.36$, $p < .001$), but not for the first (i.e., preboundary) event (Pearson’s $r = .05$, $p = .536$).

Figure 3

Average Reproduction Errors for (a) Experiment 1 (With Event Boundaries Induced by Sudden Pitch Changes) and (b) Experiment 2 (With Event Boundaries Induced by Sudden Timbre Changes)



Note. Positive values indicate that reproduced tones were played earlier in time (relative to their actual timestamps), whereas negative values indicate that reproduced tones were played later in time. The colors/shadings of different event boundary conditions (as indicated by the position of the dashed line) match their counterparts in Figure 1. Expt. = experiment. See the online article for the color version of this figure.

delayed keypresses reliably occurring immediately after the event boundaries), but also reliable contraction within events (on both sides of the boundary, as a function of the distance between each tone and the boundary itself).

Experiment 2: Event Boundaries via Timbre Changes

To ensure that the systematic patterns of temporal dilation and contraction observed in Experiment 1 reflect a more general effect of event boundaries (and are not somehow specific to pitch changes, *per se*), we conducted a conceptual replication in which event boundaries were introduced via sudden changes in timbre.

Method

This experiment was identical to Experiment 1 except where noted. Thirty new subjects participated (with this preregistered

sample size chosen to exactly match that of Experiment 1). Subjects were asked their age and their sex in a free-response post-experiment debriefing phase ($M_{\text{age}} = 21.1$ years; 50% female, 50% male). Tones (now all played at the same frequency of 293.66 Hz, or D4) were rendered via the “Hand Bells” and “Acoustic Guitar” instrument settings in MuseScore software: each rhythmic sequence began with one of the two instruments (randomly determined on each trial), and then switched at some point to the other instrument (i.e., switching from hand bells to acoustic guitar, or vice versa).

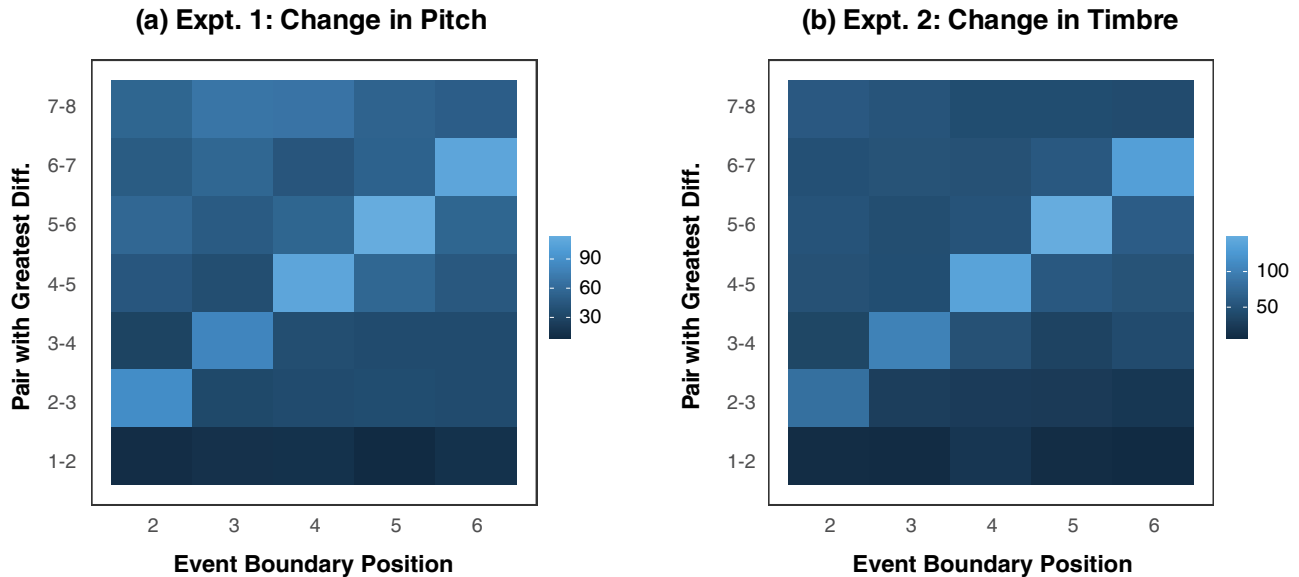
Results and Discussion

Per the preregistration, reproductions were excluded if (a) subjects reported a reproduction rating below 3 for that particular trial ($M = 19.46$, $SD = 12.82$), and (b) the reproduction duration was 2 *SDs* greater than the subject’s mean reproduction duration, after the initial

Figure 4

Trial Counts Depicting the Temporal Placement of the Largest Difference Between Adjacent Reproduction Errors (With Brighter Colors Corresponding to More Trials) Across Different Event Boundary Placements for (a) Experiment 1 (With a Pitch Change) and (b) Experiment 2 (With a Timbre Change)

Greatest Change in Reproduction Error as a Function of Boundary Placement



Note. The salient diagonal line of bright boxes indicates that the largest difference in reproduction errors typically occurred for pairs spanning the event boundary. Expt. = experiment; Diff. = difference. See the online article for the color version of this figure.

below 3 rating exclusions ($M = 2.19$, $SD = 1.08$). The data were analyzed using the same method as in Experiment 1.

Dilation Across Events

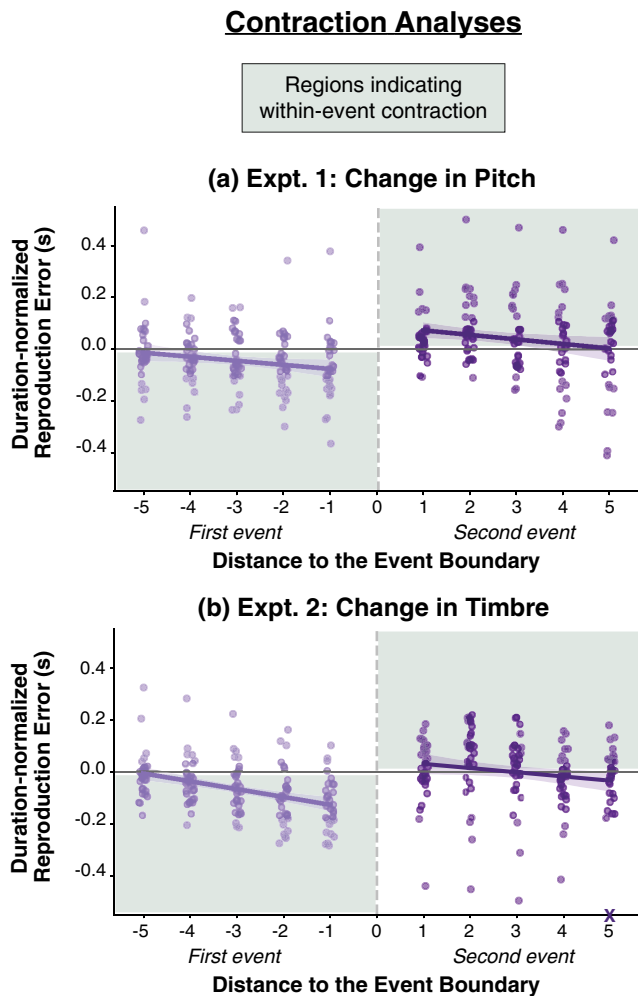
The average reproduction errors for each of the eight tones across the five event boundary conditions are depicted in Figures 3b and 4b. These figures reveal the same patterns as in Experiment 1, with systematic temporal dilation for the tone immediately after the event boundary. The temporal placement of the event boundary again reliably predicted the adjacent pair of reproduction errors that had the greatest difference ($\beta = .18$, $p < .001$). Additionally, when we shuffled the order of each subject's keypresses for each trial 1,000 times, the probability that the majority of tone pairs with the greatest difference spanned the boundary for all five event boundary conditions was again 0/1,000 ($p < .001$). On average, only 0.99 out of the five event boundary conditions showed the same effect as did the actual data (where the majority of tone pairs with the greatest difference spanned the event boundary). And the primary preregistered analysis revealed that reproduction errors were again reliably greater for tones occurring immediately after the event boundary than for tones occurring immediately prior to the event boundary, $t(29) = 6.26$, $p < .001$, $d = 1.14$. (This pattern was true for 28/30 subjects [two-tailed binomial test, $p < .001$] and did not significantly correlate with musical or dance experience [musical training: $r = -.23$, $p = .226$; dance training: $r = -.07$, $p = .731$].)

Contraction Within Events

Without taking the distance to the boundary into account, duration-normalized reproduction errors in the first event were reliably negative, $t(29) = 3.87$, $p = .001$, $d = 0.71$, but duration-normalized reproduction errors in the second event did not differ reliably from zero, $t(29) = .020$, $p = .983$, $d < .001$. Again, these patterns became clearer once we took the distance of each tone to the boundary into account: the magnitude of the duration-normalized contraction was a function of the distance to the event boundary for both the first (i.e., preboundary) event (with reproduction errors becoming more negative as they approached the event boundary: Pearson's $r = -.39$, $p < .001$; see the left side of Figure 5b) and the second (i.e., postboundary) event (with reproduction errors becoming more negative as they receded from the event boundary: Pearson's $r = -.16$, $p = .044$; see the right side of Figure 5b).²

² When we ran the same contraction analyses on the raw reproduction errors—that is when they were not yet normalized for overall duration—we observed the same general pattern: the magnitude of contraction in the raw reproduction errors was a function of the distance to the event boundary for both the first (i.e., preboundary) event (with reproductions errors becoming more negative as they approached the event boundary: Pearson's $r = -.22$, $p < .001$) and the second (i.e., postboundary) event (with reproductions errors becoming more negative as they receded from the event boundary: Pearson's $r = -.32$, $p < .001$).

Figure 5
Magnitude of Temporal Contraction and Temporal Distance to the Boundary in (a) Experiment 1 (With a Pitch Change) and (b) Experiment 2 (With a Timbre Change)



Note. The direction corresponding to within-event temporal contraction in each case is shaded gray (negative for the first event, positive for the second event). The results indicate greater contraction with closer proximity to the boundary. One subject in the timbre plot was included in the analyses themselves, but was outside the plotted range and so is indicated with an X on the negative edge. Expt. = experiment. See the online article for the color version of this figure.

This experiment thus replicated the key patterns of results from Experiment 1, demonstrating that event boundaries structure temporal experience, and that this is not dependent on a particular type of event segmentation cue.

Experiment 3: Event Boundaries via Rhythmic Structure

Event segmentation in the previous experiments was induced by *extrinsic* cues (either pitch or timbre)—in the sense that the cue itself was not part of what was being reproduced. But of course in real-world experience—for example when listening to music—we may frequently hear tones segmented into different discrete phrases simply

because of the underlying rhythmic structure itself. Perhaps the most direct example of this is brute repetition: suppose that you are listening to a sequence of tones that never changes in pitch or timbre, but repeats (i.e., with same intertone timings) multiple times. This repetition will cause you to readily hear the tones as arriving in higher order discrete rhythmic phrases. Thus, to determine whether temporal dilation at event boundaries can also occur through such *intrinsic* cues, we constructed new tone sequences from arrangements of three-tone base rhythms, where pairs of these base rhythms were repeated twice in immediate succession. In practice this caused subjects to hear the tones as coming in discrete and repeating musical phrases (i.e., temporal groups of multiple tones) despite the individual tones all being identical to each other. Critically, these underlying sequences were constructed so that there was always a within-phrase silence that was equal in duration to the key across-phrase silence—but we nevertheless predicted that the latter would be temporally dilated relative to the former in subjects' reproductions.

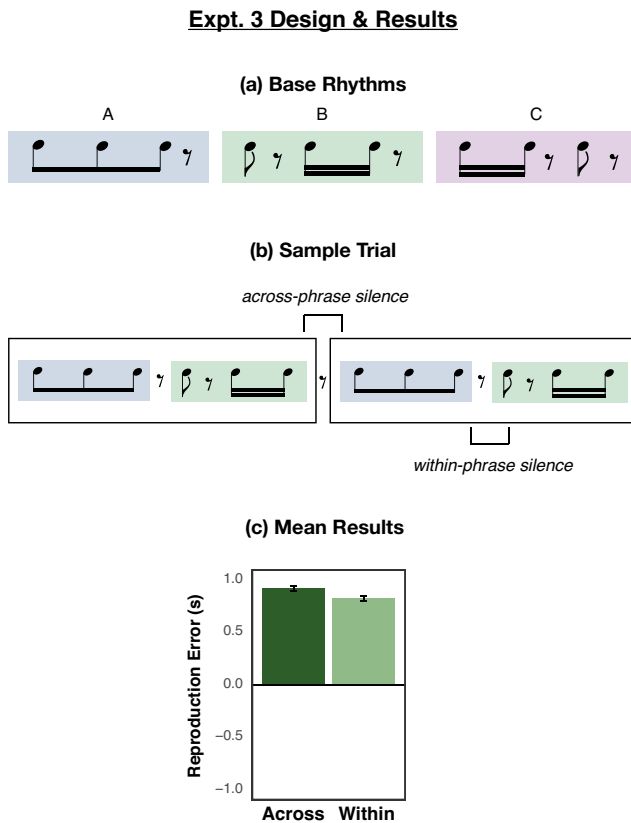
Method

This experiment was identical to Experiment 1 except where noted. Thirty new subjects participated (with this preregistered sample size chosen to exactly match that of Experiments 1 and 2). Subjects were asked their age and their sex in a free-response post-experiment debriefing phase ($M_{\text{age}} = 19.8$ years; 53% female, 47% male). Across subjects, three base rhythms were used (A, B, and C; each specified in Figure 6a), with three tones (293.66 Hz or D4) in each rhythm. Base rhythms were paired into six different rhythmic phrases (AB, AC, BA, BC, CA, and CB), with each sequence as a whole then consisting of the repetition of a single phrase (e.g., ABAB), as in the example depicted in Figure 6b. The duration of the within-phrase silence (i.e., between either A and its corresponding B) was then always identical to the duration of the repetition-induced across-phrase silence (i.e., between the two instances of AB) on that same trial—and this period of silence was randomly selected on each trial from the following possibilities: 125 ms, 250 ms, 500 ms, or 750 ms. Subjects completed two blocks of 24 experimental trials (6 rhythmic phrases \times 4 durations), presented in a different random order for each subject, with a short self-timed break in between the blocks.

Results and Discussion

As before, reproductions were excluded if (a) subjects reported a reproduction rating below 3 for that particular trial ($M = 8.40$, $SD = 9.59$), and (b) the reproduction duration was 2 SD s greater than the subject's mean reproduction duration, after the initial below 3 rating exclusions ($M = 1.63$, $SD = 0.74$). Per the preregistered analysis plan, and as depicted in Figure 6c (where a positive value again reflects subjective temporal dilation, and a negative value reflects subjective temporal contraction), the average reproduction error for the single tone (per trial) immediately following the across-phrase silence was reliably greater than the average reproduction error for the two tones (per trial) immediately following the two within-phrase silences, 923 ms versus 828 ms, $t(29) = 4.36$, $p < .001$, $d = 0.80$. (These reproduction errors are consistent with the wider range of intertone silences used in this experiment, with the longest duration being 750 ms.) This pattern was true for 23/30 subjects (two-tailed binomial test, $p = .005$), and did not

Figure 6
Experiment 3 Design and Results



Note. (a) Three base rhythms consisting of three tones each. Rhythmic phrases were composed of two base rhythms (e.g., AB or BC). (b) Rhythmic phrases were repeated twice such that there was both a silence between the phrases (across-phrase silence) and a silence within the phrases (within-phrase silence). (c) Positive values reflect temporal dilation, whereas negative values reflect temporal contraction. The average reproduction error for the across-phrase silence was significantly greater than for the within-phrase silence. Expt. = experiment. See the online article for the color version of this figure.

significantly correlate with musical or dance experience (musical training: $r = .12$, $p = .515$; dance training: $r = -.01$, $p = .960$). Thus, temporal dilation by event boundaries can also be driven by intrinsic event segmentation cues based on brute repetition.

General Discussion

Throughout any given day, we constantly keep track of time. We do this in obvious ways that are both external (e.g., when we check our watches) and internal (e.g., via circadian rhythms), but we may also do so in subtler ways—as when we judge the passage of time as a function of the discrete events that we experience *happening*. Some previous work has supported this possibility, suggesting that event segmentation leads to overall temporal dilation (e.g., Faber & Gennari, 2015) or contraction (e.g., Liverence & Scholl, 2012). The current work, in contrast, suggests that both of these patterns are partly true, but that neither is the full story: rather than having an all-or-nothing impact on our experience of time, the current

results reveal that event segmentation *structures* subjects' rhythmic reproductions in systematic, nonuniform ways. In particular, we observed that event segmentation had different effects on different moments of the underlying rhythms, in two ways that were highly dependent on the placement of the event boundary on any given trial.

Boundary-Driven Dilation

First, temporal dilation in subjects' reproductions was systematically localized to the moments just after the event boundary—with an extra delay between the two notes spanning that boundary. And this effect generalized in at least three ways: (a) it occurred regardless of whether the event boundary was implemented with an extrinsic cue (such as a sudden change in pitch or timbre, as in Experiments 1 and 2) or an intrinsic cue (such as repeated rhythmic structure, as in Experiment 3); (b) it occurred independently of just where the event boundary occurred in the sequence, with robust effects for event boundaries that occurred after the second note or the penultimate note in the sequence (as in the left vs. right sides of Figure 3); and (c) it occurred regardless of the extent of subjects' experience with rhythms in the context of training in music or dance.

This boundary-driven dilation may reflect the very purpose of event segmentation in the first place. One reason for representing the world in terms of discrete events (rather than just a continuous stream of stimulation) is that not every moment is equally predictive of what is about to happen next, and prediction error may radically increase precisely when one event transitions to another (for a review, see Kurby & Zacks, 2008). Consider again a day spent at an amusement park: what you are experiencing near the beginning of one ride (say, the roller coaster) may be highly predictive of what you will be experiencing later on that same ride, but what you are experiencing near the end of that ride may not be very predictive at all of what you will be experiencing initially on the next ride (say, the Ferris wheel)—even when those two latencies are identical. As such, it is helpful to represent the two events as *distinct*, in part to help avoid having now-obsolete statistical information transfer from one event to another. And perhaps the most direct way to make two events *distinct* is to literally separate them from each other in time—via the introduction of an extra delay in the underlying representation.

This logic also seems consistent with prior work showing that people are less likely to detect anomalous pauses inserted into music when they are placed at important positions in musical score (Repp, 1992), since those positions may correspond directly to the transition from one phrase (or verse) of a song to another. This introduction of extra perceptual delays that are not present in the underlying signal is also reminiscent of our perception of native speech: the acoustic signal itself doesn't contain reliable pauses between words, but we nevertheless hear such illusory pauses precisely at the transition from one word to another. And the usefulness of such pauses has also been directly demonstrated when they are not illusory. For example, (a) inserting actual delays between words, phrases, or story fragments improves memory (e.g., Gold et al., 2017; van de Ven et al., 2022), and (b) hearing pauses between words can help children learn object names (Estes & Hurley, 2013).

Boundary-Driven Contraction

The second way in which reproduction errors in the current study were dependent on the particular placement of the event boundary

was that temporal dilation was not the whole story. In addition, we observed temporal *contraction* for tones *within* a given event, with tones in the first event (before the boundary) reproduced “too early,” and tones in the second event (after the boundary) reproduced “too late”—with this contraction scaling as a function of the temporal proximity of a given tone to the boundary (in both Experiments 1 and 2).³

This temporal contraction may reflect the same underlying logic of making different events more distinct from each other in terms of their underlying representations. After all, when there are many different items, there are always two complementary ways to enhance distinctiveness: (a) *segmentation*: the items across categories can be moved further from each other, or (b) *grouping*: the items within a category can be moved closer to each other. This is exactly what we observed in the present study, where the “categories” are the discrete rhythmic phrases, and the metric is the latency from one tone to the next. And these two complementary phenomena may also have independently helpful effects for cognition (see also Chunharas et al., 2022). For example, time dilation can improve processing of information within a duration (e.g., Wutz et al., 2015), while time contraction can aid in binding across different stimulus modalities (e.g., Zimmermann et al., 2016).

Constraints on Generality

The effects reported here generalized across multiple types of auditory boundary cues, across several different boundary timings, and the effects seemed unrelated to music and dance experience. However, there remain several open questions about the degree to which these results might generalize in other ways. First, it remains unclear whether these results are modality-specific. Some past work has shown comparable event segmentation effects across vision and audition (e.g., De Freitas et al., 2014), but the present results tested only auditory sequences. Future work could test these same effects with “visual rhythms,” action sequences, or even meaningful stimuli such as linguistic passages. Second, the current study tested only sequences of eight tones, and so it remains unknown whether or how these effects would generalize to shorter or longer sequences. Finally, although we did not observe reliable correlations in a general population with music or dance expertise (see also Repp, 2008), it may be interesting to explore the existence and magnitude of such effects in true experts (e.g., professional musicians or dancers).

Conclusions

This project began by asking about how event segmentation interacts with time perception: does event segmentation lead to subjective time dilation, or contraction? The clear answer from the current study is that this may not be the right question, because it introduces a false dichotomy. In fact, event segmentation *structured* perceived time in a nuanced way, with *both* dilation and contraction occurring at different moments of the reproductions, relative to the underlying event boundary. Such effects would not have been possible to observe in a task which simply measured a single monolithic duration (e.g., the overall duration of the entire sequence), but it fell out directly from the simple rhythmic reproduction task introduced here.

Accordingly, these results may help to reconcile the seemingly inconsistent results from the very few previous studies of temporal distortions due to event segmentation. In particular, we speculate that when multiple short events occur in quick succession, the

internal event structure may be more salient than the transitions across boundaries, and so contraction dominates during duration reproduction (e.g., Liverence & Scholl, 2012). And conversely, when there are only two events and a single event boundary in a sequence, or when the events spanning a boundary are perceived as highly dissimilar or novel, the boundaries themselves may be more salient, and so dilation dominates during reproduction (e.g., Faber & Gennari, 2015). These nuanced ways in which event segmentation structures time perception may help to explain how our perception of time can be so subjectively malleable while also allowing us to stay in sync with what is happening in the external world.

³We did not explore within-event temporal contraction in Experiment 3 simply because any such effects would necessarily be confounded or corrupted by other factors. In particular, whereas the intertone latencies were always independent of each other in Experiments 1 and 2, Experiment 3 required those latencies to repeat, which can in turn fuel a task demand to reproduce them in the same way. Critically, this did not apply to the intertone latency including the event boundary, since that was the single latency from the sequence that did not repeat.

References

- Block, R. A. (1978). Remembered duration: Effects of event and sequence complexity. *Memory & Cognition*, 6(3), 320–326. <https://doi.org/10.3758/BF03197462>
- Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning & Memory*, 4(6), 656–665. <https://doi.org/10.1037/0278-7393.4.6.656>
- Chunharas, C., Rademaker, R., Brady, T., & Serences, J. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology: General*, 151(10), 2300–2323. <https://doi.org/10.1037/xge0001191>
- De Freitas, J., Liverence, B. M., & Scholl, B. J. (2014). Attentional rhythm: A temporal analogue of object-based attention. *Journal of Experimental Psychology: General*, 143(1), 71–76. <https://doi.org/10.1037/a0032296>
- Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, 18(2), 131–136. <https://doi.org/10.1016/j.conb.2008.06.002>
- Estes, K. G., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, 18(5), 797–824. <https://doi.org/10.1111/inf.12006>
- Faber, M., & Gennari, S. P. (2015). In search of lost time: Reconstructing the unfolding of events from memory. *Cognition*, 143, 193–202. <https://doi.org/10.1016/j.cognition.2015.06.014>
- Gold, D. A., Zacks, J. M., & Flores, S. (2017). Effects of cues to event segmentation on subsequent memory. *Cognitive Research: Principles and Implications*, 2(1), 1–15. <https://doi.org/10.1186/s41235-016-0043-2>
- Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception & Psychophysics*, 72(3), 561–582. <https://doi.org/10.3758/APP.72.3.561>
- Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 12(2), 72–79. <https://doi.org/10.1016/j.tics.2007.11.004>
- Liverence, B. M., & Scholl, B. J. (2012). Discrete events as units of perceived time. *Journal of Experimental Psychology: Human Perception and Performance*, 38(3), 549–554. <https://doi.org/10.1037/a0027228>
- Macar, F., Grondin, S., & Casini, L. (1994). Controlled attention sharing influences time estimation. *Memory & Cognition*, 22(6), 673–686. <https://doi.org/10.3758/bf03209252>

- Matthews, W. J., & Meck, W. H. (2014). Time perception: The bad news and the good. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(4), 429–446. <https://doi.org/10.1002/wcs.1298>
- Ongchoco, J. D. K., & Scholl, B. J. (2020). Enumeration in time is irresistibly event-based. *Psychonomic Bulletin & Review*, 27(2), 307–314. <https://doi.org/10.3758/s13423-019-01680-z>
- Ongchoco, J. D. K., & Scholl, B. J. (2022). Scaffolded attention in time: ‘everyday hallucinations’ of rhythmic patterns from regular auditory beats. *Attention, Perception, & Psychophysics*, 84(2), 332–340. <https://doi.org/10.3758/s13414-021-02409-8>
- Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS One*, 2(11), Article e1264. <https://doi.org/10.1371/journal.pone.0001264>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Poynter, D. (1989). Judging the duration of time intervals: A process of remembering segments of experience. In L. Levin & D. Zakay (Eds.), *Time & human cognition* (pp. 305–331). Elsevier Science.
- Poynter, W. D., & Homa, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics*, 33(6), 548–560. <https://doi.org/10.3758/bf03202936>
- Repp, B. H. (1992). Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Cognition*, 44(3), 241–281. [https://doi.org/10.1016/0010-0277\(92\)90003-Z](https://doi.org/10.1016/0010-0277(92)90003-Z)
- Repp, B. H. (2008). Metrical subdivision results in subjective slowing of the beat. *Music Perception*, 26(1), 19–39. <https://doi.org/10.1525/mp.2008.26.1.19>
- Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics*, 66(7), 1171–1189. <https://doi.org/10.3758/bf03196844>
- van de Ven, V., Jäckels, M., & De Weerd, P. (2022). Time changes: Timing contexts support event segmentation in associative memory. *Psychonomic Bulletin & Review*, 29(2), 568–580. <https://doi.org/10.3758/s13423-021-02000-0>
- Wöllner, C., & Hammerschmidt, D. (2021). Tapping to hip-hop: Effects of cognitive load, arousal, and musical meter on time experiences. *Attention, Perception & Psychophysics*, 83(4), 1552–1561. <https://doi.org/10.3758/s13414-020-02227-4>
- Wutz, A., Shukla, A., Bapi, R. S., & Melcher, D. (2015). Expansion and compression of time correlate with information processing in an enumeration task. *PLoS One*, 10(8), Article e0135794. <https://doi.org/10.1371/journal.pone.0135794>
- Zimmermann, E., Derichs, C., & Fink, G. R. (2016). The functional role of time compression. *Scientific Reports*, 6(1), Article 25843. <https://doi.org/10.1038/srep25843>

Received April 2, 2022

Revision received April 26, 2023

Accepted May 9, 2023 ■