Supplemental Information: Norming Rhythm Boundaries
[De Freitas et al. (2013), *JEP:General*]

JULIAN DE FREITAS, BRANDON M. LIVERENCE, & BRIAN J. SCHOLL
Yale University

In order to confirm that naïve observers would perceive boundaries in our rhythmic stimuli in the same places that we did when designing them, we normed these stimuli in an independent group of subjects. While listening to each repeating phrase, subjects simply pressed a key whenever they perceived a boundary between repetitions of the phrase.

METHOD

Subjects

144 subjects (students and other members of the Yale and New Haven community) participated in exchange for $1. We chose this relatively large number of subjects since (a) each subject made only a few responses in a between-subjects design; (b) we were using a type of measure and stimuli here that to our knowledge had never before been used (so that we had no a priori guide as to what sample size would be required, unlike the experiments reported in the main text); and (c) this experiment was only meant for norming stimuli for the primary experiments, rather than as a test for our hypothesis about same-'object' advantages in time rather than space.

Stimuli

Auditory stimuli were created in MATLAB using the PsychToolbox libraries (Brainard, 1997; Pelli, 1997), and played on a Macintosh computer through headphones. Subjects listened to repeating sequences of 279 Hz tones (at moderate volume), arranged into four rhythmic phrases (approximately 2.78, 3.24, 2.91, and 3.56 s, respectively), with each subject hearing only one of the four phrases. The four phrases are each depicted in musical notation in Figure 2 of the main text. To ensure that responses were not influenced by which portion of the phrase subjects heard first, the phrase gradually faded in from one of four potential starting points (counterbalanced across subjects). On average, the fade-in period lasted 7.4 s (approximately 2.4 phrase repetitions), during which the volume linearly increased from silence to full volume. The fade-in began with the same notes for each pair of rhythmic phrases as was noted in Footnote 2 of the main text (see also Figure 2).

Procedure and Design

Subjects were asked to press a key at the end of each repetition of the rhythmic phrase. (Because it was also possible to hear 'sub-phrases' — just as visual objects can have parts — the instructions clarified that keypresses were only to be made at the end of each "largest pattern that you hear repeating".) Subjects first listened to eight practice repetitions of the rhythmic phrase, after which a brief high-pitched (450 Hz) tone signaled when they should begin responding. So that the tone did not by itself yield a segmentation cue, it occurred at one of three possible points (counterbalanced across subjects) during the last two practice repetitions. The phrase then repeated six more times, while keypresses were recorded.

RESULTS AND DISCUSSION

36 subjects participated for each of the 4 rhythmic phrases. Since the task featured only six repetitions of a single phrase per subject, we excluded 16 additional subjects for making either fewer than four or greater than eight keypresses total. The average moment of the keypresses is depicted in Figure 2 of the main text via blue shading (horizontal extent indicates 95% confidence intervals). As can be appreciated from the figure, the points where subjects indicated phrase boundaries were impressively consistent, and mirrored our initial intuitions. Statistically, the average offset between keypress and Between-Phrase rest was only 512ms, whereas the average offset between keypress and Within-Phrase rest was almost twice as long (959ms) — a highly significant difference ($t(143)=5.68$, $p<.001$, $d=.47$).