

# Implicit learning of between-group intervals in auditory temporal structures

J. Terry<sup>1,2</sup> · C. J. Stevens<sup>1,3</sup> · G. Weidemann<sup>3</sup> · B. Tillmann<sup>1,4</sup>

Published online: 14 June 2016  
© The Psychonomic Society, Inc. 2016

**Abstract** Implicit learning of temporal structure has primarily been reported when events within a sequence (e.g., visual-spatial locations, tones) are systematically ordered and correlated with the temporal structure. An auditory serial reaction time task was used to investigate implicit learning of temporal intervals between pseudorandomly ordered syllables. Over exposure, participants identified syllables presented in sequences with weakly metrical temporal structures. In a test block, the temporal structure differed from exposure only in the duration of the interonset intervals (IOIs) *between groups*. It was hypothesized that reaction time (RT) to syllables following between-group IOIs would decrease with exposure and increase at test. In Experiments 1 and 2, the sequences presented over exposure and test were counterbalanced across participants (Pattern 1 and Pattern 2 conditions). An RT increase at test to syllables following between-group IOIs was only evident in the condition that presented an exposure structure with a slightly stronger meter (Pattern 1 condition). The Pattern 1 condition also elicited a *global* expectancy effect: Test block RT slowed to earlier-than-expected syllables (i.e., syllables shifted to an earlier beat) but not to later-than-expected

syllables. Learning of between-group IOIs and the global expectancy effect extended to the Pattern 2 condition when meter was strengthened with an external pulse (Experiment 2). Experiment 3 further demonstrated implicit learning of a new weakly metrical structure with only earlier-than-expected violations at test. Overall findings demonstrate learning of weakly metrical rhythms without correlated event structures (i.e., sequential syllable orders). They further suggest the presence of a global expectancy effect mediated by metrical strength.

**Keywords** Temporal structure · Rhythm · Meter · Auditory modality · Serial reaction time task · Music

## Introduction

Sequences of auditory events, such as speech and music, unfold in time and, via mere exposure, humans are able to perceive and learn the temporal structure of these sequences (Bigand & Poulin-Charronnat, 2006; Jones & Boltz, 1989). Learning leads to the development of expectations about *when* events are going to occur, and temporal expectations influence perception of a complex array. Expectations allow listeners to direct attention to the points in time when an event is expected to occur, facilitating event processing (Boltz, 1993; Jones & Boltz, 1989; Large & Jones, 1999; Penel & Jones, 2005).

Using an auditory serial reaction time task (SRTT), we investigated the learning of temporal structures when the order of events (i.e., event structure or *what* dimension) was pseudorandom. More specifically, we investigated learning of the temporal intervals between groups of auditory events, defined here as *between-group IOIs* (interonset intervals). The rationale is that even if the *identity* of an upcoming auditory event is unpredictable, having an expectation of *when* it will occur will lead to faster responses, compared to if that event occurs

✉ J. Terry  
j.terry@westernsydney.edu.au

<sup>1</sup> MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Richmond, Australia

<sup>2</sup> Western Sydney University - Bankstown, Locked Bag 1797, Penrith, NSW 2751, Australia

<sup>3</sup> School of Social Sciences and Psychology, Western Sydney University, Richmond, Australia

<sup>4</sup> Lyon Neuroscience Research Center, CNRS-UMR 5292, INSERM U1028, Université de Lyon, Auditory Cognition and Psychoacoustics Team, Lyon, France

at an unexpected point in time. We tested the hypothesis that listeners acquire a perceptual benefit from exposure to a repeating temporal structure. In other words, we investigated whether reaction time (RT) to identify syllables (presented in a pseudorandom order) would become faster over exposure blocks, as the temporal structure of the sequence was learned. As this decrease in RT can also be driven by task learning, we examined whether violating the temporal structure by shortening or lengthening the between-group IOIs would lead to slower syllable identification (test block). An additional aim was to determine whether learning of temporal structure can occur implicitly. That is, we examined whether learning could occur without awareness being drawn to the temporal structure and without requiring participants to have an intention to learn (Cleeremans, Destrebecqz, & Boyer, 1998).

### Perceptual advantages of temporal expectations

A temporal context (e.g., a short isochronous sequence of tones) can establish expectations of when a subsequent event might occur. Judgements of an auditory event (e.g., pitch) occurring at an expected compared to unexpected time point are faster and more accurate (Jones, Moynihan, MacKenzie, & Puente, 2002; Jones & Yee, 1997; Penel & Jones, 2005; Tillmann & Lebrun-Guillaud, 2006). According to the dynamic attending theory, attentional energy is directed to expected points in time so that the perception of events is optimized (Bausenhart, Rolke, & Ulrich, 2007; Jones & Boltz, 1989; Large & Jones, 1999). These effects appear to be driven by facilitation at the perceptual stage rather than at a motor response stage (e.g., Bausenhart et al., 2007; Correa, Lupiáñez, Madrid, & Tudela, 2006; Correa, Lupiáñez, & Tudela, 2005; Rolke & Hofmann, 2007; Sanabria, Capizzi, & Correa, 2011). Our reported experiments investigate whether complex temporal structure can be learned, giving rise to temporal expectations that elicit perceptual benefits (i.e., faster RT to identify events).

### Grouping structure and between-group intervals

One way in which listeners perceptually organize complex temporal structures is to group events together that have similar time spans between them, with longer time spans or between-group intervals marking the boundaries between these groups (Deutsch, 1999; Lerdahl & Jackendoff, 1983). An isolated event bounded by an equivalent time span is also considered a group (Handel, 1998; Jackendoff & Lerdahl, 2006). Hence, a sequence of groups and isolated events gives a temporal pattern its *grouping structure* (see Fig. 1). It has been demonstrated that listeners are significantly better than chance at explicitly discriminating between two sequences of tones with temporal structures that differ according to their grouping structure, but are less able to do so when the grouping structure is maintained while the between-group intervals are violated (Handel, 1998).

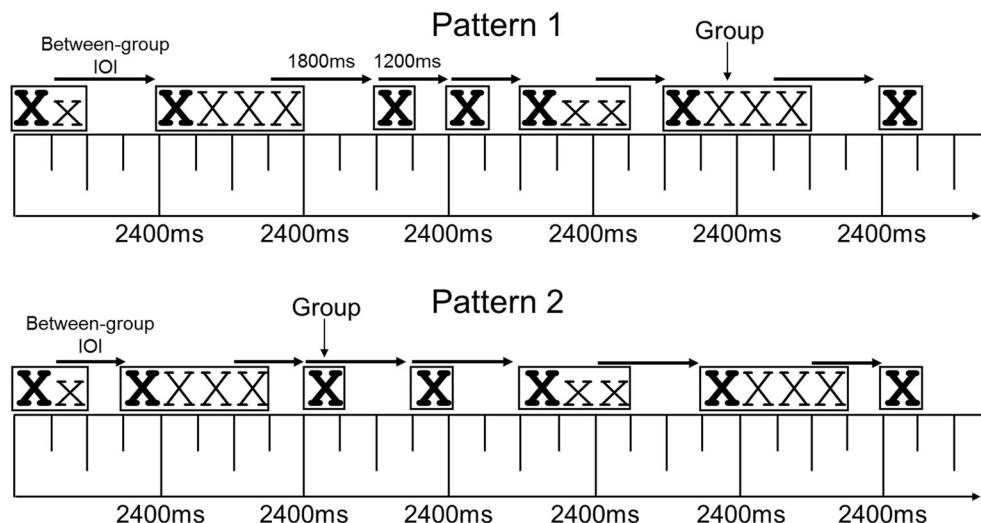
However, there is some evidence to suggest that strong meters facilitate the encoding of between-group intervals and improve discrimination (Handel, 1998; Hébert & Cuddy, 2002; Ross & Houtsma, 1994). Meter is a hierarchical structure that gives rise to the perception of regular patterns of strong and weak beats embedded within larger scale patterns of strong and weak beats (Lerdahl & Jackendoff, 1983; see Fig. 1). When events align with beats at multiple levels in the hierarchy, the more readily the meter is abstracted and perceived as strong (Jones, 1987; Lerdahl & Jackendoff, 1983; Palmer & Krumhansl, 1987). A strong meter facilitates production and perception, and it has been proposed that a strong meter may also facilitate learning (Jones, 2010).

### Implicit learning of temporal structures

Previous findings have led to the hypothesis that implicit learning may be a powerful means through which listeners learn temporal structure and develop temporal expectations (Salidis, 2001; Tillmann, Stevens, & Keller, 2011). To test the strength of implicit learning, we investigated the abstraction and learning of between-group IOIs within complex temporal structures that are *weakly metrical*. This was done using an implicit learning task (i.e., SRTT) rather than explicit discrimination. Implicit learning occurs outside of awareness and without an intention to learn (Boyer, Destrebecqz, & Cleeremans, 2005; Perruchet, 2008; Reber, 1993). Learners are often unable to describe the regularities in the materials to which they have been exposed despite showing evidence of a benefit from these regularities when performing a behavioral task.

Implicit learning of the order of events in a sequence (i.e., ordinal or event structure) in auditory and visual sequences has been demonstrated using SRTTs (e.g., A. Cohen, Ivry, & Keele, 1990; Destrebecqz & Cleeremans, 2001; Mayr, 1996; Nissen & Bullemer, 1987; Perruchet, 2008; Reed & Johnson, 1994). However, less empirical research has investigated implicit learning of temporal structure, and some studies even suggest that temporal structure can only be learned when correlated with a systematic event structure (e.g., auditory pitch sequence, visuospatial locations; Buchner & Steffens, 2001; Shin & Ivry, 2002, respectively).

However, a small number of studies using an auditory adaptation of the SRTT have demonstrated learning of temporal structure without a correlated event structure (e.g., Brandon, Terry, Stevens, & Tillmann, 2012; Schultz, Stevens, Keller, & Tillmann, 2013; Salidis, 2001; Tillmann et al., 2011). For instance, Salidis (2001) required participants to press a key every time they heard a tone. Over exposure, the tones were presented according to a nonmetrical but fixed sequence of response-stimulus intervals. RT decreased with exposure but increased at a test block that presented tones according to a *random* sequence of response-stimulus intervals, indicating that



**Fig. 1** Temporal structures used in the current experiment (Pattern 1 and Pattern 2): “X” represents a syllable, bold font represents positions following between-group IOIs, lowercase font represents controlled within-group positions. Syllable groups are enclosed in boxes.

Horizontal arrows represent short (→) and long (→) between-group IOIs (onset to onset), and vertical lines represent the hypothesized metrical structure (the longer the line, the stronger the perceived beat)

acquired temporal expectations had been violated. This effect was most notably for the shortest intervals (180 ms). The inherent temporal variability arising from the response-stimulus interval sequence, and nonmetricality of the temporal structures, likely prohibited the potential benefit of a perceived temporal grid, or pulse. This may have impeded learning of the longer and more difficult-to-predict intervals (i.e., medium: 450 ms, long: 1,125 ms). Indeed, these longer intervals may have acted as between-group intervals and, as a consequence, were difficult to process in the absence of a beat grid (Handel, 1998; Hébert & Cuddy, 2002; Ross & Houtsma, 1994).

It is important to note that Salidis (2001) required participants to *detect* the onset of each event. Earlier research, failing to find independent learning of temporal structure, instead required *identification* of the events presented according to the temporal structure (Buchner & Steffens, 2001; Shin & Ivry, 2002). Two studies have shown that implicit learning can occur in an identification task, but, in both instances, the temporal structures were based on a sequence of IOIs and were strongly metrical (Brandon et al., 2012; Tillmann et al., 2011; see also Schultz et al., 2013, for a discussion on identification vs. detection in auditory SRTTs). The events were syllables (e.g., “Pa”, “Ta”) presented according to either a simple (Tillmann et al., 2011) or complex (Brandon et al., 2012) repeating sequence of IOIs. Importantly, the syllable orders were pseudorandom and uncorrelated with the temporal structures. In other words, the syllable order was unpredictable/unlearnable, and consequently, a perceptual advantage could only be achieved by learning the temporal structure. Brandon et al. (2012) employed an exposure/test block design and found that RT to identify syllables decreased with exposure to the temporal sequences but increased at the test block when the order of the IOIs and the grouping

structure was changed. Independent temporal structure learning most likely emerged because the use of IOIs eliminated the response-based variability inherent in response-stimulus interval structures (as in Salidis, 2001; Shin & Ivry, 2002), and the strong meter allowed for the exploitation of a beat grid to facilitate temporal prediction.

### Aims of present research

The present study further investigated implicit learning of temporal structures presented in the context of a pseudorandom sequence of events (i.e., uncorrelated temporal and event structures). The aim was to examine implicit learning of complex (i.e., nonisochronous) weakly metrical structures based on repeating sequences of multiple IOIs with simple integer ratios (rather than response-stimulus intervals). Specifically, we used weakly metrical sequences to examine whether the intervals between groups of auditory events can be learned.

### Design

Using an auditory SRTT, syllables were presented in exposure sequences with temporal structures defined by a sequential ordering of IOIs (see Fig. 1). A test block sequence was presented that had an identical grouping structure to the exposure sequence (i.e., the same number of groups, number of events within groups, and the order of groups) but differed in terms of the duration of the between-group IOIs. More specifically, short (1,200 ms) and long (1,800 ms) between-group IOIs within an exposure sequence were interchanged in the test sequence, so that a short IOI became a long IOI and a long IOI became a short IOI.

In all experiments, the dependent variable was reaction time to correctly identify syllables (see each experiment for more detail). The experimental design consisted of seven blocks: Blocks 1 to 5 presented an exposure sequence (Pattern 1 or Pattern 2), and Block 6—the test block—presented the alternate sequence. Block 7 presented again the exposure sequence. As is classically done in serial reaction time tasks, the disturbing effect of the test block was measured by comparing RT at Block 6 to the mean of exposure Blocks 5 and 7 (e.g., Buchner & Steffens, 2001; Curran & Keele, 1993; Destrebecqz & Cleeremans, 2001; Salidis, 2001; Shin & Ivry, 2002). Importantly, we counterbalanced the patterns across exposure and test so that one group of participants was presented with Pattern 1 during exposure and Pattern 2 during test, and another group was presented with Pattern 2 during exposure and Pattern 1 during test. This meant that we could (a) assess the effects of violating temporal expectation at the test block and (b) ensure that the patterns were equivalent in terms of overall difficulty by comparing performance across the two exposure patterns (see also Destrebecqz & Cleeremans, 2001; Saffran, Johnson, Aslin, & Newport, 1999).

In a posttest phase, a questionnaire was administered to determine the presence of explicit knowledge of the temporal structure. Finally, in Experiments 1 and 2, participants rated their familiarity with the exposure, test, and novel sequences, and indicated the certainty with which they made the rating of familiarity.

## Hypotheses

Across three experiments, it was hypothesized that (a) correct RT would decrease over exposure blocks as participants learned the temporal structure, (b) RT would be slower at the test block compared to the adjacent exposure blocks, and (c) if learning remained implicit, participants would be unable to describe the temporal structure in an open-ended questionnaire and would report equivalent levels of familiarity across exposure, test, and novel sequences.

## Experiment 1

### Method

#### Participants

Forty-three participants from Western Sydney University took part in the experiment in exchange for course credit. The 38 female and five male participants had a mean age of 21.8 years ( $SD = 6.6$  years, range = 18–46 years), and all had self-reported normal hearing. Participants were randomly assigned to either the Pattern 1 ( $N = 21$ ) or Pattern 2 ( $N = 22$ ) exposure condition. The two participant groups were equivalent in terms of years of

musical training (Pattern 1 condition:  $M = 1.78$  years,  $SD = 2.63$ , median = .50; Pattern 2 condition:  $M = 1.32$  years,  $SD = 2.15$  years, median = 0). Levels of musical training were not statistically different across the two pattern conditions ( $p = .53$ ).

### Materials

**SRTT** Two weakly metrical patterns were used in a counterbalanced design. One group of participants (Pattern 1 condition) was exposed to Pattern 1 and tested on Pattern 2, while this was reversed for the other group of participants (Pattern 2 condition). The temporal structure of each sequence was a chaining of 15 IOIs, which was repeated nine times in a block. The shortest IOI was 600 ms and the longer IOIs were multiples of this base temporal unit (i.e., 1,200 ms and 1,800 ms). The base temporal unit (600 ms) has been shown to be a “preferred tempo,” with spontaneous finger tapping and clapping often being performed around this rate (Fraisne, 1982; van Noorden & Moelants, 1999). The IOI chaining of Pattern 1 was 600-1,800-600-600-600-1,800-1,200-1,200-600-600-1,200-600-600-1,800, and the IOI chaining of Pattern 2 was 600-1,200-600-600-600-1,200-1,800-1,800-600-600-1,800-600-600-1,200 (see Fig. 1). The longer IOIs (1,200 ms and 1,800 ms), formed boundaries between groups of syllables and isolated syllables (between-group IOIs). The two patterns were identical in terms of grouping structure. That is, the number of groups, number of events within groups, and the order of groups were identical in the two patterns. This is illustrated in Fig. 1. The difference between Pattern 1 and Pattern 2 was the duration of the IOIs between the groups and the isolated events: IOIs of 1,200 ms in Pattern 1 were substituted for IOIs of 1,800 ms in Pattern 2. Likewise, IOIs of 1,800 ms in Pattern 1 were substituted for IOIs of 1,200 ms in Pattern 2.

Events in the sequences were three syllables—“Pa,” “Ta,” and “Ka”—spoken with a male voice generated using a text-to-speech synthesizer, Mbrola. The syllables had a fundamental frequency of 120 Hz, a duration of 218 ms, and were normalized for intensity.

Each block consisted of a sequence of 136 syllables presented in a pseudorandom order. Because the first syllable in each block did not follow a between-group IOI, it was not included in the analyses. Instead an additional syllable was added at the end of each block to create the final between-group IOI. The following constraints were applied from the second to the final syllable in each sequence. There were 45 presentations of each syllable, with no adjacent repetitions and with second- and third-order repetitions roughly equated across blocks. Second-order repetitions are instances when, for example, two “Pas” are interposed with one of the other syllables (“Pa Ta Pa”), and third-order repetitions are instances when, for example, two “Pas” are interposed with two of the other syllables (“Pa Ta Ka Pa”). Finally, as the temporal structure of the presentation of syllables was defined by a repeating basic sequence of 15 IOIs, syllables were

distributed equally across all serial positions in each sequence. To ensure that any effects could not be attributed to syllable order, half the participants were presented with a second syllable order that was a reversal of the order presented to the other half of participants.

For each block, Audio Interchange File Format (AIFF) files of each syllable sequence, with the defined repeating IOI chaining were created in MATLAB. Each sequence had a duration of 2 mins and 10 s. Two practice sequences with two repetitions of the temporal structure of Pattern 1 and Pattern 2 were also created. The experiment was run in Psyscope (J. Cohen, MacWhinney, Flatt, & Provost, 1993), and the sequences were presented over Sennheiser HD 25 closed headphones.

Although Experiment 2 investigated learning of between-group IOIs, and was not aimed at manipulating meter, a measure of metrical strength (C score) was applied (Povel & Essens, 1985)<sup>1</sup> to ensure that both patterns were weakly metrical. The C scores for Pattern 1 and Pattern 2 were 15 and 19, respectively, indicating that although both patterns were weakly metrical, Pattern 2 had a slightly weaker meter (a score of zero indicates maximal metrical strength). This difference was due to Pattern 2 having one less event on a beat, compared to Pattern 1. While it would have been optimal to have identical C scores for both patterns, the choice of patterns was limited by the following constraints: (a) Both patterns needed to have identical grouping structures, and (b) there needed to be an even number of between-group IOIs (three long and three short) directly transposed across the two patterns (i.e., a long between-group IOI in Pattern 1 became a short between-group IOI in Pattern 2, and vice versa). Nonetheless, both patterns were weakly metrical and as closely matched in their degree of metrical strength as possible.

<sup>1</sup> The measure of metrical strength, or C score (Povel & Essens, 1985), is based on an internal clock model of subjective accent induction and has two key principles/assumptions: (1) Rhythms have an underlying beat or pulse grid, and (2) the grouping structure of a rhythm gives rise to the perception of subjectively accented tones (i.e., isolated tones, the second in a group of two tones, and the first and last of a group of three or more tones are perceived as accented). The C score is generated by considering the relationship between the hypothesized beat grid (or internal clock pulse) and the subjective accent structure. Specifically, the counterevidence-based measure considers the number of internal clock pulses that are not accompanied by an event or are accompanied by a subjectively unaccented event using the formula:  $C = (W * -ev) + (1 * 0ev)$ , where -ev is the number of beats that coincide with a silence and 0ev is the number of beats that coincide with an unaccented auditory event. W is a weight generally set at four. Given the relatively slow tempo of the patterns used in the current experiments (base IOI of 600 ms), the clock pulse was considered to occur at 1,200-ms intervals, with 600-ms intervals being subdivisions of the beat. The lower the C score, the stronger the meter of the sequence, with zero indicating maximal metrical strength. C scores are also influenced by the length of a rhythm and typically have been calculated on rhythms with 16 beats (e.g., Patel, Iversen, Chen, & Repp, 2005; Povel & Essens, 1985). In the current study, our rhythms extend over 20 beats and, therefore, the C scores are subsequently inflated. Thus, they act as a measure of the relative strength of the two rhythms.

**Posttests** The purpose of the posttest phase was to measure participants' explicit knowledge of the temporal structures to which they were exposed in the SRTT. Posttest stimuli were short syllable sequences ("Pa," "Ta," "Ka") presented with the temporal structure of Pattern 1 and Pattern 2. Each posttest sequence was one cycle of the chaining of 15 IOIs. Because the temporal structure was repeated in the SRTT, it was possible that participants may have segmented the sequence using any of the groups or isolated syllables as a starting point. Hence, six versions of each pattern were created, each starting at a different group or isolated syllable.

In addition to the posttest sequences based on the temporal structure of Pattern 1 and Pattern 2, two sets of six sequences were created that had novel grouping structures while still using the same number of groups and group sizes. Novel 1 and Novel 2 had grouping structures of 1-2-1-4-4-3 and 1-4-4-3-1-2, respectively. The first novel set of sequences (Novel 1) differed from Pattern 1, and the second novel set (Novel 2) differed from Pattern 2 in terms of the group-to-between-group IOI chaining. The six versions of the two novel posttest patterns began on different groups or isolated syllables. Both novel patterns were weakly metrical. The Novel 1 and Novel 2 patterns respectively were "X..XX.X..XXXX.XXXX.XXX.." and "X.XXXX..XXXX..XXX.X.XX.." (X" represents a syllable, and a full stop(s) represents a between-group interval).

### Procedure

**SRTT** As a cover story, participants were advised that the experiment was investigating the speed and accuracy of syllable identification. They were not informed of the temporal structure of the sequences. Participants listened to the sequences of syllables and identified each syllable as quickly and accurately as possible. As soon as they knew their response after the onset of each syllable, they pressed the 1, 2, or 3 key (corresponding to the identified syllable) on the numeric keypad of the computer keyboard. Labels identifying the appropriate response key were placed on the keys above. Using the right hand, participants kept their index, middle, and ring fingers on the keys at all times. Participants were instructed to not correct themselves if they made a mistake or if they missed any syllables. To minimize motor-based effects, three key-to-syllable mappings were counterbalanced across participants.

A practice block with the temporal structure of the exposure sequence was completed, and participants were reminded of the instructions before beginning the experimental blocks. The participants in the *Pattern 1* condition were presented with five exposure blocks of Pattern 1. This was followed by a test block presentation of Pattern 2 (Block 6) and a final block of Pattern 1 (Block 7). Conversely, the participants in the *Pattern 2* condition were presented five exposure blocks of Pattern 2, followed by a test block presentation of Pattern 1

and a final block of Pattern 2. A short break with a minimum duration of 30 s was provided between each block.

**Posttests** After completing the SRTT, participants reported any temporal regularity they noticed in the sequences. They were then informed that the sequences followed a repeating timing or rhythmic pattern. In the posttest task, participants identified syllables as per the SRTT. This ensured that the testing conditions were as equivalent as possible across the SRTT and posttest task, thus giving participants the best opportunity to elicit any explicit knowledge. Participants exposed to Pattern 1 in the SRTT responded to the six posttest sequences from the Pattern 1, Pattern 2, and Novel 1 sets. Participants exposed to Pattern 2 in the SRTT responded to the six posttest sequences from the Pattern 1, Pattern 2, and Novel 2 sets. The sequences were presented in random order.

After responding to each sequence, participants were asked to think back to the blocks in the SRTT of the exposure phase of the experiment and rate their familiarity with the timing pattern of the posttest sequence just heard using a scale: 1 = *very unfamiliar*, 2 = *unfamiliar*, 3 = *somewhat unfamiliar*, 4 = *somewhat familiar*, 5 = *familiar*, 6 = *very familiar*. They were then asked to indicate the certainty with which they made their familiarity rating: 1 = *complete guess*, 2 = *very uncertain*, 3 = *somewhat uncertain*, 4 = *somewhat certain*, 5 = *very certain*, 6 = *completely certain* (adapted from Destrebecqz & Cleeremans, 2001). The experiment took 50 minutes.

## Results

### SRTT

To measure learning across exposure blocks, we examined correct RT to syllables in serial positions that directly followed the between-group IOIs (see Fig. 1). To determine that learning extended beyond task learning and was specifically of between-group IOIs, RT to syllables following IOI changes at the test block (Block 6) was also examined. This was done by comparing RT to positions following between-group intervals at Block 6 to these same positions in the two adjacent exposure blocks (mean of Blocks 5 and 7).

To measure learning beyond the between-group IOIs, we also examined RT to particular serial positions within groups of syllables (within-group positions), that is, positions 2, 10, and 11 (see Fig. 1). These positions were selected because they maintained their global temporal position and consequently provided a measure of global temporal structure learning rather than a measure of local temporal violation (i.e., because of a change in the preceding IOI). RT to positions 2, 10, and 11 were collapsed to give a mean for each block. These controlled within-group positions were compared in the same way as the positions following between-group IOIs.

Correct responses were included in the analyses if they occurred 250 to 900 ms after the onset of the syllable. RTs less than 250 ms were assigned to the previous syllable if it was missed and, if multiple responses were made within the 250 to 900 ms window, only the first response was kept. The upper limit (900 ms) allowed us to retain any longer RTs made to the syllables immediately preceding a between-group interval. Mean accuracy to syllables across all serial positions was significantly above chance (33.33 %) for the Pattern 1 condition, 58 % ( $SD = 14\%$ ),  $t(20) = 7.75$ ,  $p = .00$ , and Pattern 2 condition, 60 % ( $SD = 14\%$ ),  $t(21) = 9.28$ ,  $p = .00$ . For syllables following between-group IOIs, accuracy was 69 % and 70 % in the Pattern 1 and Pattern 2 conditions, respectively.<sup>2</sup>

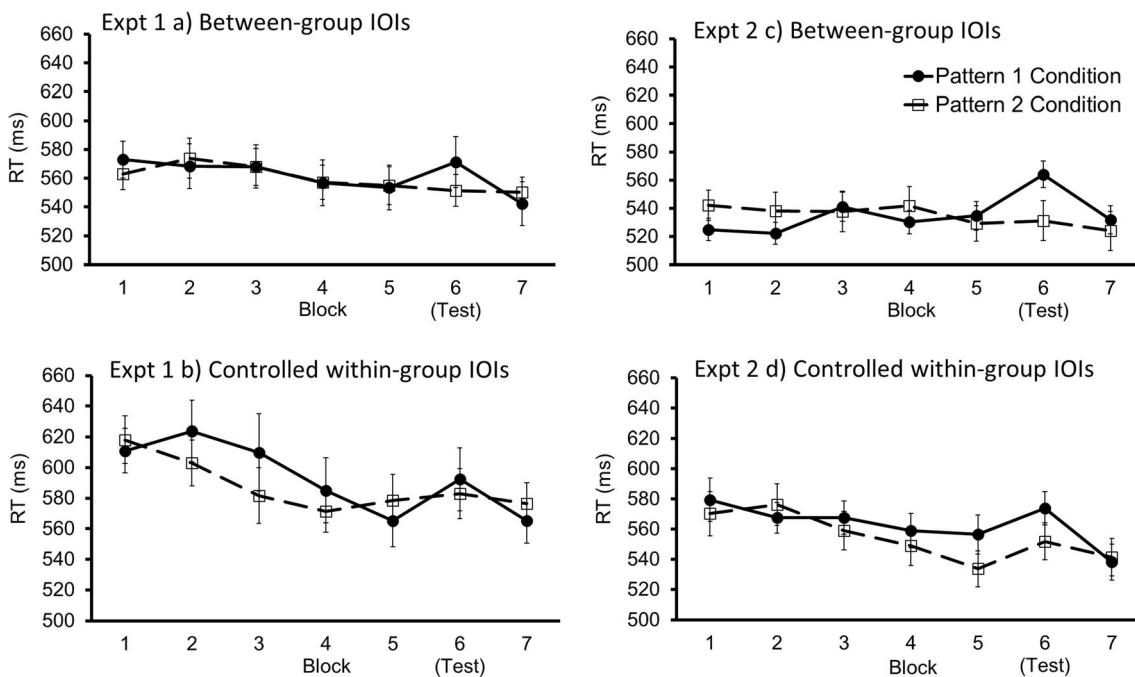
### Learning of between-group IOIs

**Exposure blocks** RT to syllables immediately following between-group IOIs was averaged for each block. Testing the hypothesis that RT would decrease over the first five exposure blocks, a  $5 \times 2$  ANOVA with block (Blocks 1–5) as a within-subjects factor and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor was conducted (see Fig. 2a)<sup>3</sup>. The main effect of block was significant,  $F(4, 164) = 2.95$ ,  $p = .02$ , partial  $\eta^2 = .07$ . There was no main effect of pattern condition,  $F(1, 41) = 0.00$ ,  $p = .97$ , partial  $\eta^2 = .00$ , and no interaction,  $F(4, 164) = .41$ ,  $p = .80$ , partial  $\eta^2 = .01$ . The difference between Block 1 and Block 5 fell just short of significance,  $F(1, 41) = 3.54$ ,  $p = .06$ , partial  $\eta^2 = .08$ . However, as can be seen in Fig. 2a, RT increased at Block 2 and then decreased over the remaining blocks.

**Test block** To investigate the effect of introducing changes to the between-group IOIs at the test block, a  $2 \times 2$  ANOVA with block as a within-subjects factor (mean of Blocks 5 and 7, Block 6) and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor was conducted. This revealed a significant effect of block,  $F(1, 41) = 3.98$ ,  $p = .05$ , partial  $\eta^2 = .08$ , and a significant Block  $\times$  Pattern condition interaction,  $F(1, 41) = 4.74$ ,  $p = .04$ , partial  $\eta^2 = .10$ . A contrast analysis revealed an effect of block in the Pattern 1 condition, with RT

<sup>2</sup> For all experiments, correlations between accuracy and correct RT were conducted for each block. All significant correlations were negative ( $p < .05$ ) revealing that the faster the RT, the higher the accuracy. Hence, no speed/accuracy trade-off was evident.

<sup>3</sup> RT was slower to controlled within-group positions, compared to positions following between-group intervals. This is likely because two within-group positions (positions 2 and 11) were the last in a group, and participants may have taken advantage of the immediately following long between-group IOI to make their response (slowing down to do so). In addition, syllables following between-group intervals indicate the start of a group and participants have additional preparation time to make their response. This particularly may be the case once between-group IOIs are learned. The longer preparation time (due to the longer between-group intervals) may also explain the higher accuracy to these syllables compared to the controlled within-group syllables.



**Fig. 2** Experiment 1 (left panel) and 2 (right panel): Correct RT to syllables following between-group IOIs (top panel) and controlled within-group syllable positions 2, 10, and 11 (bottom panel), presented

as a function of block and pattern condition. Error bars show standard error of the mean

slowing significantly at the test block compared to the mean of the two adjacent exposure blocks,  $F(1, 20) = 8.34, p = .01$ , partial  $\eta^2 = .29$ .<sup>4</sup> There was no significant increase in RT at the test block in the Pattern 2 condition,  $F(1, 21) = .02, p = .90$ , partial  $\eta^2 = .00$  (see Fig. 2a).

#### Learning of controlled within-group IOIs

**Exposure blocks** A 5 (block)  $\times$  2 (pattern) ANOVA revealed a main effect of block,  $F(4, 164) = 6.75, p = .00$ , partial  $\eta^2 = .14$ . RT at Block 5 was significantly faster than RT at Block 1,  $F(1, 41) = 18.22, p = .00$ , partial  $\eta^2 = .31$ . The main effect of pattern condition and its interaction with block were not significant,  $ps > .23$  (see Fig. 2b).

**Test block** An ANOVA revealed that RT at the test block (Block 6) was slower than at the adjacent exposure blocks (mean Blocks 5 and 7),  $F(1, 41) = 4.10, p = .05$ , partial  $\eta^2 = .09$ . The main effect of pattern condition and its interaction with block was not significant,  $ps > .19$ . Although, as can be seen in Fig. 2, the Pattern 1 condition showed a larger increase at the test block compared to Pattern 2 (see Fig. 2b).

<sup>4</sup> Additional analyses were conducted to ensure that the RT increase at the test block in the Pattern 1 condition was due to temporal violation and not simply driven only by between-group IOIs that were lengthened (1,200 ms to 1,800 ms). These analyses revealed that the RT increase at the test block occurred for IOIs that became shorter (changed from 1,800 ms to 1,200 ms) and also for IOIs that become longer (changed from 1,200 ms to 1,800 ms) at the test block, relative to exposure blocks.

#### Posttests

Although 25 % of participants reported that there were pauses between some syllables, none were able to report the temporal structure of the exposure or test sequences. To measure participants' familiarity with the exposure sequence, test sequence, and novel sequence (see Table 1), a mean rating of familiarity for each sequence type was calculated by collapsing across the six presentations of each sequence. Means were compared with 3  $\times$  2 ANOVA, with sequence as a within-subjects factor (exposure, test, novel) and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor. There was no effect of sequence or pattern condition and the interaction was not significant,  $ps > .10$ . This indicates that participants had not acquired explicit knowledge of the temporal structures.

#### Discussion

The aim of Experiment 1 was to investigate implicit learning of between-group IOIs in weakly metrical temporal structures. As hypothesized, RT to syllables presented according to a repeating temporal structure decreased over exposure blocks as participants became familiar with the task, learned the temporal structure, and, consequently, developed expectations of when the syllables would occur.

In the Pattern 1 condition, the violation of the between-group IOIs at the test block resulted in RT to the immediately following syllables becoming significantly slower compared to the mean of the adjacent exposure blocks. The effect of

**Table 1** Experiments 1 and 2: Posttest familiarity (1 = *very unfamiliar* to 6 = *very familiar*) and certainty (1 = *complete guess* to 6 = *completely certain*) mean ratings and standard deviations

Experiment	Pattern	Familiarity						Certainty					
		Exposure		Test		Novel		Exposure		Test		Novel	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	1	3.73	1.00	3.33	1.02	3.63	0.92	4.01	0.56	3.89	0.58	3.94	0.74
	2	3.68	0.86	3.70	0.88	3.52	0.95	4.24	0.64	4.14	0.67	4.14	0.68
2	1	4.19	0.79	3.58	0.65	3.54	0.83	4.29	0.63	3.84	0.63	3.95	0.66
	2	4.04	0.65	3.85	0.69	3.83	0.55	4.23	0.57	4.13	0.68	4.05	0.60

altering the between-group IOIs at the test block also carried over to the controlled within-group syllables: syllables that maintained their location within the global temporal structure and the metric structure. These findings indicate that the development of temporal expectation was not only of the between-group IOIs but also of the global temporal structure. Furthermore, explicit awareness was not a necessary requirement for learning the temporal structure of Pattern 1 because participants were no more familiar with the exposure sequence than with the test or novel sequence.

Contrary to our hypothesis, there was no RT increase at the test block in the Pattern 2 condition. This was despite an RT decrease over exposure blocks to both positions following between-group IOIs and controlled within-group positions. Although metrical strength was not experimentally manipulated across the two pattern conditions, Pattern 2 did have a slightly weaker metrical structure than Pattern 1 ( $C = 19$  and  $C = 15$ , respectively),<sup>5</sup> and this may have made learning the between-group IOIs in Pattern 2 more difficult. However, the relatively weaker meter did not result in overall slower syllable identification, because RT over exposure blocks was equivalent for both patterns. In other words, the weaker meter of Pattern 2 did not necessarily make the structure more difficult to respond to. The important implication for the Pattern 1 condition is that the RT increase at the test block cannot be attributed to Pattern 2 being more difficult to respond to (i.e., slower syllable identification). Rather, there may have been differences in the degree to which the temporal structures could be learned.

For instance, participants may have learned the grouping structure *and* between-group IOIs in the Pattern 1 condition but only the grouping structure in the Pattern 2 condition. This interpretation is supported by previous work suggesting that *grouping* is the dominant strategy used in perceptually organizing temporal structure in the absence of a perceived metrical grid (Hébert & Cuddy, 2002; Ross & Houtsma, 1994). Given this, it

is possible that strengthening the meter of Pattern 2 may induce a perceived metrical grid that allows listeners to integrate between-group intervals into the learned structure.

Pulses have been used previously to support metrical structure (Handel, 1998; Hannon & Trehub, 2005). Indeed, Handel found in an explicit discrimination task that the presence of an additional exogenous/external pulse facilitated the abstraction of between-group intervals in tone sequences. Hence, Experiment 2 presented an isochronous and amplitude-accented woodblock pulse concurrently with the to-be-learned temporal structures (Pattern 1 and 2). It was hypothesized that the presence of the pulse, facilitating the abstraction of a metrical structure, would enable participants to learn the between-group and controlled within-group IOIs also in Pattern 2 (i.e., RT decreases over exposure and increases at the test block).

## Experiment 2

### Method

#### Participants

Forty-five participants from the Western Sydney University took part in the experiment in exchange for course credit. The 35 female and 10 male participants had a mean age of 21.57 years ( $SD = 6.55$  years range = 18–52 years). All participants had self-reported normal hearing. Participants were randomly assigned to either the Pattern 1 ( $N = 22$ ) or Pattern 2 ( $N = 23$ ) condition. The two participant groups were equivalent in terms of years of musical training (Pattern 1 condition:  $M = .93$  years,  $SD = 2.3$ , median = 0; Pattern 2 condition:  $M = 1.82$  years,  $SD = 3.63$ , median = 0). These levels of musical training were not statistically different across the two conditions ( $p = .33$ ).

### Materials

**SRTT** Experiment 2 was identical to Experiment 1 in all respects except that an isochronous woodblock pulse with IOIs of 600 ms (i.e., a woodblock sound presented every 600 ms)

<sup>5</sup> It was not possible to precisely match metrical strength across the two patterns as the priority was to generate patterns that had identical grouping structures and that allowed us to directly transpose the short and long between-group intervals whilst maintaining weak metricality.

was played concurrently with the syllable sequences. This pulse was added to induce a quadruple meter and to reduce the metrical ambiguity of Pattern 2. The 218-ms woodblock sound was sourced from Logic Pro 8. An amplitude-accent structure was added to the pulse so that woodblock sounds aligning with strong beats were 3 dB louder than sounds aligning with moderate beats, and 6 dB louder than sounds aligning with weak beats. This resulted in an alternation of a strong accent, no accent, moderate accent, no accent, followed again by a strong accent, and so on. Hence, a woodblock sound could be heard with each syllable and during each between-group interval.

The syllable sequences were presented binaurally, and the woodblock pulse was presented to the left channel only. Measured at the headphones using a sound pressure level meter, the syllables were approximately 5 dBA louder than the strong accents of the woodblock pulse.

**Posttests** The posttest phase in Experiment 2 was identical to Experiment 1 except for the addition of the accented isochronous woodblock pulse to each of the posttest sequences.

#### Procedure

**SRTT and posttests** The SRTT and posttest procedures were identical to Experiment 1 except that participants were informed that this experiment was investigating the speed and accuracy of syllable identification in the presence of a stream of nonspeech sounds. Participants were instructed to focus on the syllables and to respond as quickly and accurately as possible.

#### Results

##### SRTT

The criteria for identifying correct responses and for including RT data were as in Experiment 1. Mean accuracy was significantly better than chance (33.3 %) in the Pattern 1 condition, 67 % ( $SD = 12\%$ ),  $t(21) = 12.56$ ,  $p = .00$ , and in the Pattern 2 condition, 66 % ( $SD = 15\%$ ),  $t(22) = 10.57$ ,  $p = .00$ . For syllables immediately following between-group IOIs, accuracy was 75 % and 74 % in the Pattern 1 and Pattern 2 conditions, respectively<sup>2</sup>.

##### Learning of between-group IOIs

**Exposure blocks** A  $5 \times 2$  ANOVA with block (Blocks 1–5) as a within-subjects factor and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor was conducted (see Fig. 2c). There was no main effect of block or pattern condition, and the interaction was not significant,  $ps > .11$ . Mean RTs over blocks are displayed in Fig. 2c and suggest that there

was a trend toward an RT decrease from Block 1 to Block 5 in the Pattern 2 condition.

**Test block** To investigate the effect of introducing changes to the between-group IOIs at the test block, RT to syllables following between-group IOIs at Block 6 was compared to these same syllable positions in the two adjacent exposure blocks (mean of Blocks 5 and 7). A  $2 \times 2$  ANOVA with block as a within-subjects factor (mean of Blocks 5 and 7, Block 6) and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor was conducted and revealed a significant effect of block,  $F(1, 43) = 11.21$ ,  $p = .00$ , partial  $\eta^2 = .21$ , and a significant Block  $\times$  Pattern condition interaction,  $F(1, 43) = 6.11$ ,  $p = .01$ , partial  $\eta^2 = .12$ . A contrast analysis revealed a significant effect of block for Pattern 1, with RT becoming significantly slower at the test block compared to the mean of the two adjacent exposure blocks,  $F(1, 21) = 16.17$ ,  $p = .00$ , partial  $\eta^2 = .44$ . As in Experiment 1, there was no significant increase at the test block in the Pattern 2 condition ( $p = .53$ ).

##### Learning of controlled within-group IOIs

As for Experiment 1, RT to within-group syllables (positions 2, 10, and 11) was examined for evidence of global structure learning. These syllables maintained their location within the global temporal and metric structure and are therefore a sensitive measure of global structure learning.

**Exposure blocks** A  $5$  (block)  $\times$   $2$  (pattern) ANOVA revealed a main effect of block,  $F(4, 172) = 3.83$ ,  $p = .01$ , partial  $\eta^2 = .08$ , with Block 5 significantly faster than Block 1,  $F(1, 43) = 8.05$ ,  $p = .01$ , partial  $\eta^2 = .16$ . The main effect of pattern condition and its interaction with block were not significant,  $ps > .32$  (see Fig. 2d).

**Test block** A  $2 \times 2$  ANOVA with block as a within-subjects factor and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor was conducted. There was a main effect of block with a significant RT increase at the test block, compared to the mean of the adjacent exposure blocks,  $F(1, 43) = 10.64$ ,  $p = .00$ , partial  $\eta^2 = .20$ . The main effect of pattern condition and its interaction with block were not significant,  $ps > .28$  (see Fig. 2d).

##### Posttests

While 42 % of participants reported that there were pauses between some syllables, none were able to describe the temporal structures of the exposure or test sequences. A  $3 \times 2$  ANOVA was conducted on mean familiarity ratings with sequence as a within-subjects factor (exposure, test, novel) and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor (see Table 1). The results show a main effect of

sequence,  $F(2, 86) = 6.38, p = .00$ , partial  $\eta^2 = .13$ , but no effect of pattern condition and no interaction ( $p > .19$ ). Additional comparisons revealed that participants in both conditions rated the exposure sequences as significantly more familiar than the test,  $F(1, 43) = 10.49, p = .00$ , partial  $\eta^2 = .20$ , and novel sequences,  $F(1, 43) = 8.20, p = .01$ , partial  $\eta^2 = .16$ . To examine the effect of awareness on learning in the Pattern 1 condition, the data was reanalyzed with awareness as a between-subjects factor with two levels (aware, unaware; Scott & Dienes, 2008). Participants in the aware condition met two criteria. First, their mean rating to the exposure sequences was greater than their overall mean familiarity rating. Second, their certainty ratings to the exposure sequences were equal to, or greater than four (4 = *somewhat certain*, 5 = *very certain*, 6 = *completely certain*). Thirteen of the 22 participants in the Pattern 1 condition met these criteria. For following between-group and controlled within-group positions, there was no main effect of group (aware, unaware),  $p > .29$ , and no Group  $\times$  Block interactions,  $p > .58$ . In summary, these results reveal that explicit knowledge of the temporal structure was not necessary for learning to occur.

## Discussion

Experiment 2 investigated whether the presence of an accented isochronous woodblock pulse facilitated implicit learning of weakly metrical structures. This aim was drawn from previous research showing that listeners are better able to perceive the between-group intervals in relatively strong compared to weak metrical structures (Hébert & Cuddy, 2002; Ross & Houtsma, 1994) and that a concurrently presented pulse aids this perception (Handel, 1998).

A decrease in RT over exposure blocks was evident only in the Pattern 2 condition, whereas RT in the Pattern 1 condition was fast starting right at Block 1. The addition of the pulse seemingly allowed participants in the Pattern 1 condition to develop temporal expectations early in the exposure phase. Given also that Pattern 1 had a slightly stronger meter than Pattern 2, this finding provides support for the proposition that the elicitation of a perceived metrical grid facilitated the development of temporal expectation (Jones, 2009).

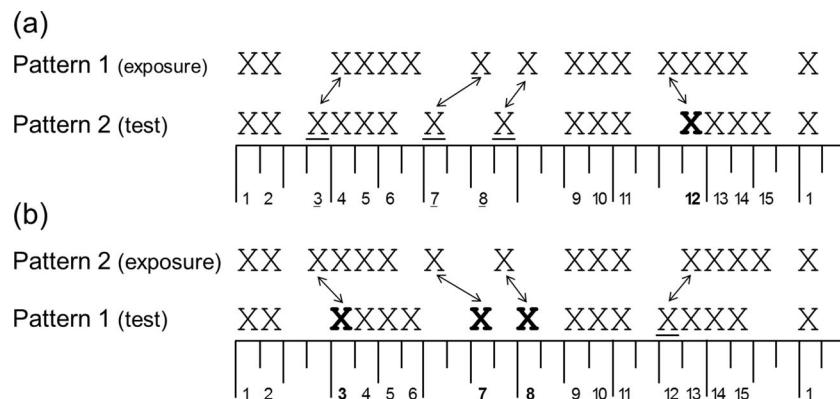
For syllables following between-group IOIs, there was the expected RT increase at the test block in the Pattern 1 condition but not for Pattern 2 condition. However, for the controlled within-group positions, there was the expected RT increase at the test block for both Pattern 1 and Pattern 2 conditions, indicating that learning of the global structure had occurred in both pattern conditions.

Consistent with Experiment 1, explicit awareness of the temporal structure was not required to elicit learning. Posttests revealed that participants in both the Pattern 1 and Pattern 2 conditions rated the exposure sequence as more familiar than the test sequence. However, this acquired

awareness did not induce the expected RT increase to syllables following between-group IOIs at the test block in the Pattern 2 condition. Thus, there may be an additional factor that underpins the lack of evidence of learning of between-group IOIs in this condition.

A possible candidate to explain the lack of apparent learning of the between-group IOIs in Pattern 2 is the nature of the violation, that is, whether syllables were presented earlier- or later-than-expected across the entire pattern. Previous research investigating temporal orienting of attention and the development of temporal expectations has demonstrated that that in the context of an isochronous event presentation, RT is slower to events (e.g., visual targets and chords) presented earlier-than-expected compared to events presented on time or later than expected, despite the later-than-expected event also being a temporal violation (e.g., Capizzi, Sanabria, & Correa, 2012; Correa, Lupiáñez, Milliken, & Tudela, 2004; Correa, Lupiáñez, & Tudela, 2006; Tillmann & Lebrun-Guillaud, 2006). In these studies, the violations were at a *local* level: The isochronous sequence elicited an expectation of regularity such that a final event occurred earlier or later than expected relative to the previous event. Conversely, in the present experiment, the violations occurred within a complex temporal structure (i.e., not isochronous) and can be considered as occurring at a *global* level. That is, the violating events occur *earlier* or *later* than expected in reference to the extracted underlying beat grid. In other words, the violations are to the relationship between an event in the temporal pattern and its corresponding position along the beat grid. The effect of a global violation is irrespective of whether an event occurs earlier or later in reference to the preceding event. For instance, in Pattern 1, the third event occurred on the fifth beat but occurred on the fourth beat in Pattern 2 (see Fig. 3). Thus, at the test block in the Pattern 1 condition, the third event occurred earlier than expected (see Fig. 3, Pattern 2). Indeed, at the test block in the Pattern 1 condition, Positions 3, 7, and 8 were presented *earlier* than expected, and Position 12 was presented *later* than expected. Conversely, for Pattern 2 condition (see Fig. 3, Pattern 1), at the test block, Position 12 was presented *earlier* than expected, and Positions 3, 7, and 8 were presented *later* than expected.

It is possible that the specific effects of early and late violations may occur not only at a local level (as shown by Capizzi et al., 2012; Correa et al., 2004; Correa, Lupiáñez, & Tudela, 2006; Tillmann & Lebrun-Guillaud, 2006) but can also occur at a global, and beat-based level. In musical contexts, it is known that judgments of melodic completion are lower when the duration of some musical notes are altered so that the final note occurs earlier compared to later than expected along the beat grid. In other words, the violation of timing is to the relationship between the temporal location of the note and the metric structure, not simply to the temporal relationship between the final note and the preceding note (Boltz, 1989). Similarly, chords are



**Fig. 3** (a) Pattern 1 condition and (b) Pattern 2 condition: X represents a syllable. Earlier-than-expected positions at the test block indicated in underlined font, and later than expected positions indicated with bold

font. Arrows indicate the direction of the violation at the test block (early, late), and digits represent the serial position of each syllable within the basic temporal structure

judged as less “appropriate” presented earlier compared to later than expected within the global metric structure, and, furthermore, judgement RTs are slower to early versus on time and late chords (Schmuckler & Boltz, 1994). In our first two experiments, the difference in the number of syllables presented earlier than expected and later than expected (i.e., those after between-group IOIs) across the Pattern 1 and Pattern 2 conditions might have contributed to our results. As more syllables following between-group intervals were presented earlier, compared to later than expected at the test block in the Pattern 1 condition, slower RT would be expected. Conversely, as more syllables following between-group intervals were presented *later* than expected at the test block in the Pattern 2 condition, a weaker (or no) increase would be expected—essentially any learning effect would be obscured.

To investigate this possibility, we conducted additional analyses of RT to syllables following between-group IOIs presented earlier and later than expected. For each pattern condition, the mean of Positions 3, 7, and 8 (earlier than expected at the test block in the Pattern 1 condition, and later than expected at the test block in the Pattern 2 condition) was calculated at the test block (Block 6) and at the two adjacent exposure blocks (Blocks 5 and 7). These means were compared to RT at Position 12 (later than expected at the test block in the Pattern 1 condition, and earlier than expected at the test block in the Pattern 2 condition). ANOVAs were conducted with block (Block 6 vs. mean of Blocks 5 and 7) and violation (early, late) as within-subjects factors, and pattern condition (Pattern 1, Pattern 2) as a between-subjects factor (for Experiment 1 and Experiment 2, respectively).

For Experiment 1, there was a significant three-way interaction between pattern condition, violation, and Block,  $F(1, 41) = 6.32, p = .02$ , partial  $\eta^2 = .13$ . In the Pattern 1 condition, a significant increase in RT at the test block was evident for the earlier than expected positions,  $F(1, 20) = 8.05, p = .01$ , partial  $\eta^2 = .29$  (mean of Blocks 5 and 7 = 566.38 ms,  $SD = 79.64$ ; Block 6 = 598.50 ms,  $SD = 97.89$ ), but not for the later-than-

expected position ( $p = .40$ ; mean of Blocks 5 and 7 = 528.27 ms,  $SD = 71.28$ ; Block 6 = 514.25 ms,  $SD = 58.84$ ). In the Pattern 2 condition, there was no RT increase at the test block for either the earlier-than-expected position, or the later-than-expected positions ( $p > .60$ ).

For Experiment 2 (with the addition of an accented pulse), there was a significant two-way interaction between violation (early, late), and block (mean of Blocks 5 and 7, Block 6),  $F(1, 43) = 5.73, p = .02$ , partial  $\eta^2 = .12$ , but no interactions with pattern condition. Follow-up contrasts revealed that, for earlier-than-expected positions, there was a significant increase in RT at the test block,  $F(1, 44) = 7.34, p = .01$ , partial  $\eta^2 = .14$  (mean of Blocks 5 and 7 = 526.85 ms,  $SD = 62.20$ ; Block 6 = 554.08 ms,  $SD = 75.20$ ), but for later-than-expected positions, there was no significant effect ( $p = .68$ ).

Overall, these findings support the proposition that the nature of the temporal violation (early, late) can drive RT at a global level, not just at a local level as previous research has demonstrated (e.g. Capizzi et al., 2012; Correa et al., 2004; Correa, Lupiáñez, & Tudela 2006; Tillmann & Lebrun-Guillaud, 2006). In other words, RT increases when there is an early violation in reference to a beat grid (i.e., at a global level) but less so when there is a late violation. This finding is consistent with research showing in musical contexts that events (i.e., notes, chords) are judged as less appropriate when they occur earlier, compared to later than expected (Boltz, 1989; Schmuckler & Boltz, 1994). Importantly, these effects can only occur when temporal expectations have been acquired via exposure to a complex temporal structure and when the underlying metrical grid has been extracted. Indeed, the specific effects of global violation may also be moderated by metrical strength, as the condition with the exposure pattern eliciting the weakest meter (Pattern 2) did not show an RT increase to earlier than expected positions at the test block in Experiment 1. However, there was an RT increase to these positions in Experiment 2 where the beat grid was reinforced with the addition of a pulse.

Based on this observation, Experiment 3 aimed to replicate the learning of a weakly metrical structure with a new set of temporal sequences designed to not elicit late global violations at the test block. During the exposure blocks, listeners were presented a temporal structure with a comparable metrical strength to Pattern 2 (with the additional accented pulse, as in Experiment 2). The test block sequence violated temporal expectations by presenting all but one syllable *earlier* than expected<sup>6</sup> (i.e., syllables following between-group intervals). We hypothesized that RT to these earlier-than-expected syllables would decrease over exposure blocks but increase (compared to the adjacent exposure blocks) at the test block (Block 6).

## Experiment 3

### Method

#### Participants

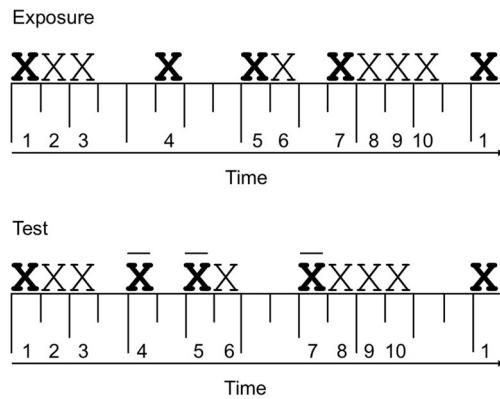
Twenty participants from the Western Sydney University took part in the experiment in exchange for course credit. The 16 female and four male participants had a mean age of 21.70 years ( $SD = 4.89$  years range = 18–38 years). All participants had self-reported normal hearing and a mean of .45 years of musical training ( $SD = .69$  years, median = 0).

#### Materials

**SRTT** The temporal structure of each sequence was a chaining of 10 IOIs, which was repeated 12 times in a block. As in Experiment 1, the shortest IOI was 600 ms and the longer IOIs were multiples of this base temporal unit (i.e., 1, 200 ms and 1,800 ms). The IOI chaining of the exposure sequence was 600-600-1,800-1,800-600-1,200-600-600-1,200, and the IOI chaining of the test sequence was 600-600-1,200-600-1,800-600-600-1,800 (see Fig. 4). The two patterns were identical in terms of grouping structure but the difference between the exposure and test sequences was the duration of the IOIs between the groups and the isolated events: IOIs of 1,200 ms were changed to 1,800 ms and IOIs of 1,800 ms were changed to 1,200 ms.

C scores were calculated on the exposure and test patterns, and they were both weakly metrical at 14 and 10, respectively. As each pattern had 16 beats, the C scores were weighted (i.e., multiplied by 1.33) to make them comparable to the longer 24 beat patterns in Experiments 1 and 2. Hence, the weighted C scores were 18.62 and 13.30, making the exposure pattern and test patterns closely matched to Pattern 2 ( $C = 19$ ) and Pattern

<sup>6</sup> One syllable position following a between-group IOI was “on time” (i.e., the first event in the pattern) and no positions following between-group IOIs were presented later than expected.



**Fig. 4** Temporal structures of exposure and test blocks. Over-lined font indicates positions presented earlier than expected at the Test blocks. “X” represents a syllable, digits indicate the serial positions of the syllables, bold font represents syllable positions that follow between-group IOIs. Vertical lines represent the hypothesized metrical structure (the longer the line, the stronger the perceived beat)

1 ( $C = 15$ ) in the previous experiments, respectively. Thus, the metrical strength of exposure and test patterns corresponded as closely as possible with the Pattern 2 condition in Experiments 1 and 2.

Syllables used were those from Experiments 1 and 2, and the same syllable order constraints were applied. The seven blocks (1 min and 56 s), and two practice blocks (58 s) were presented with the concurrent woodblock pulse (as in Experiment 2).

Although we expected learning of the temporal structure to be implicit, the focus of this experiment was to examine learning of the between-group IOIs in a weakly metrical structure when the temporal violations were primarily earlier than expected. Hence, participants were presented with the posttest questionnaire used in Experiments 1 and 2, but did not complete the familiarity and confidence ratings task.

#### Procedure

**SRTT and posttests** The procedure for the SRTT was identical to Experiment 2, except that responses were using keys 1, 2, and 3 on a USB numeric keypad (rather than on a computer keyboard). All participants were presented with the same temporal structure during exposure blocks and the same temporal violations to between-group IOIs at the test block.

### Results

#### SRTT

The criteria for identifying correct responses and including RT data were as in Experiments 1 and 2. Mean accuracy ( $M = 56.25\%$ ,  $SD = 12\%$ ) was significantly better than chance (33.3 %),  $t(19) = 8.60$ ,  $p = .00$ . For syllables following between-group IOIs, accuracy was 64.87 %<sup>2</sup>.

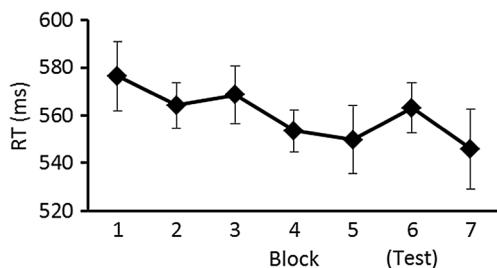
### Learning of between-group IOIs

**Exposure blocks** An ANOVA with block (Blocks 1–5) as a within-subjects factor was conducted. The main effect of block fell short of significance,  $F(4, 76) = 2.04, p = .09$ , partial  $\eta^2 = .01$ . However, the direct comparison between Block 1 and Block 5 revealed a significant decrease in RT,  $F(1, 19) = 4.77, p = .04$ , partial  $\eta^2 = .20$  (see Fig. 5).

**Test block** To investigate the effect of introducing changes to the between-group IOIs at the test block, an ANOVA with block as a within-subjects factor (mean of Blocks 5 and 7, Block 6) was conducted. This revealed an effect of block,  $F(1, 19) = 5.78, p = .03$ , partial  $\eta^2 = .23$  (see Fig. 5), with significantly increased RT at the test block in comparison to the adjacent exposure blocks. To further examine the effects of earlier-than-expected violations, RT to Serial Positions 4, 5, and 7 were averaged at Block 6 and at Blocks 5 and 7. In other words, the RT to the “on-time” syllable position was removed. A one-way ANOVA with block as a within-subjects factor (mean of Blocks 5 and 7, Block 6) confirmed a significant effect of block,  $F(1, 19) = 8.34, p = .01$ , partial  $\eta^2 = .31$ . No participants were able to describe the temporal structures of the exposure or test sequences.

### Discussion

Experiment 3 demonstrates with a new set of sequences that between-group IOIs in a weakly metrical structure can be learned. Specifically, the results revealed an RT increase at the test block when all but one position following between-group IOIs was earlier than expected in the global structure. Thus, the lack of RT increase for positions following between-group IOIs in the Pattern 2 condition of Experiments 1 and 2 was most likely due to the nature of the temporal violations (i.e., more late than early violations), in addition to the relative weakness of the meter.



**Fig. 5** Experiment 3: Correct RT to syllables immediately following between-group IOIs. Error bars show standard error of the mean

### General discussion

In three experiments, we demonstrated implicit learning of weakly metrical temporal structures in the presence of an uncorrelated and pseudorandom event structure. Importantly, this learning occurred implicitly—without the participants’ attention being drawn to the temporal structure. In other words, participants learned the temporal structure without an intention to do so, and without awareness that the temporal structure was the to-be-learned material. A key difference between the present study and earlier studies (e.g., Salidis, 2001) was that the temporal structures were based on a repeating sequence of IOIs, rather than response-stimulus intervals. The presentation of events based on IOIs eliminates the variability inherent in structures based rather on response-stimulus intervals (i.e., where the temporal presentation of events is somewhat determined by the RT to the previous event).

Although previous findings have suggested that temporal structure cannot be learned independently of event structure (Buchner & Steffens, 2001; Shin & Ivry, 2002), our research demonstrates otherwise. Syllables were presented in pseudo-random order so that no relationship between a particular syllable and a temporal interval could be formed. In other words, participants could not exploit the relationship between the timing of a motor-based structure and the upcoming syllable identity in order to facilitate responding (see Tillmann & Poulin-Charronnat, 2010, for a discussion of the distinction between motor learning and perceptual learning). Hence, our study has demonstrated a perceptual benefit driven by the learning of a sequence of temporal intervals only.

Our findings confirm and extend upon previous research showing that responses to auditory events are faster when they occur at expected compared to unexpected points in time. It is argued that attentional energy is drawn to these expected time points, facilitating processing and responding (e.g., Jones et al., 2002; Jones & Boltz, 1989; Jones & Yee, 1997; Penel & Jones, 2005; Tillmann & Lebrun-Guillaud, 2006). While our research supports this previous work, our additional contribution is the demonstration that listeners can develop temporal expectations in a single experimental session when the rhythms are unfamiliar, relatively complex (i.e., not isochronous) and weakly metrical. Although our temporal structures were weakly metrical, they did conform to an underlying beat structure. That is, all temporal intervals were multiples of a base interval (600 ms). This structure likely aided the learning of the complex, nonisochronous rhythms and facilitated the anticipation of the temporal occurrence of upcoming syllables (Ellis & Jones, 2010).

An important aspect of the current research was the use of test blocks that were highly controlled. There were a number of benefits of incorporating these blocks in the design: (a) Significant increases in RT at the test blocks indicated that RT decreases over exposure were not solely due to task

learning but rather to learning of the temporal structure, (b) We could examine the implicit learning of a particular temporal feature: the duration of intervals between groups of auditory events. This was achieved by manipulating between-group IOIs and controlling the group order and group sizes across exposure and test blocks. A further and crucial feature of our design was the counterbalancing of patterns across exposure and test (i.e., Pattern 1 and Pattern 2 conditions). This allowed us to ensure that RT increases at the test block were due to temporal violation and not general pattern difficulty (see also Destrebecqz & Cleeremans, 2001; Saffran et al., 1999, for counterbalanced use of patterns). If pattern difficulty drove slower RT at the test block in the Pattern 1 condition (Experiments 1 and 2), then we would expect RT to be slower to Pattern 2, compared to Pattern 1, over exposure blocks. This was not the case as our analyses revealed no interaction between pattern and block in the exposure phase. In addition, if Pattern 1 was less difficult than Pattern 2, then after exposure to Pattern 2, RT should considerably speed up at test block (now with the supposedly “easier” Pattern 1). As neither of these outcomes were observed in our experiments, an interpretation of our results based on pattern difficulty can be excluded. Support of an interpretation based on learning rather than pattern difficulty also arises from our finding that RT increased at the test block for controlled within-group intervals in the Pattern 1 condition (Experiments 1 and 2) and, importantly, in the Pattern 2 condition, when meter was strengthened with the woodblock pulse (Experiment 2). Had pattern difficulty driven our findings, then we would not have seen this influence of meter on global structure learning.

More evidence of global structure learning emerged from the early versus late analyses, and from the results of Experiment 3. Findings showed that global-level violations within complex temporal structures conform to similar behavioral principles as do local level violations within isochronous sequences (e.g., Capizzi et al., 2012; Correa et al., 2004; Correa, Lupiáñez, & Tudela, 2006; Tillmann & Lebrun-Guillaud, 2006). That is, RT is slower to events that occur earlier, compared to later than expected. While previous research demonstrated specific effects of early and late temporal violation after a short isochronous exposure pattern—a *local* violation—we found these same effects when the violations were early and late relative to the beat grid—a *global* violation. In local contexts, a later-than-expected event would provide listeners with additional time in which to prepare a response, thereby facilitating RT relative to an earlier-than-expected event (e.g., Penel & Jones, 2005; Schmuckler & Boltz, 1994). However, in a global context, the event occurs earlier or later than expected relative to the beat grid rather than to the immediately preceding event. Thus, an explanation based on temporal preparation is not sufficient. Rather, it may be that exposure to a rhythm (with its underlying metric structure) drives neural oscillations, and the systematic occurrence of events aligning

with oscillations at particular time points drives attention to these time points (Boltz, 1989; Schmuckler & Boltz, 1994). If a beat grid has been abstracted, and listeners expect events to fall on beats (via prior exposure), then a late event would most likely be anticipated to align with the next upcoming beat. As previous research has shown, RT to events occurring on perceived beats are faster than when occurring off beat (Jones, 2010). Furthermore, RT is facilitated to events that occur later than expected (relative to the immediately preceding temporal context) as listeners reorient their attention in anticipation of the upcoming event (Jones et al., 2002). In our study, although the later-than-expected events in the test sequence were relative to the exposure sequence (rather than the immediately preceding context), similar principles may apply. As a consequence, the expected learning effect (i.e., increased RT) was diminished for later- compared to earlier-than-expected events.

Following this, global violation effects (i.e., early vs. late) on positions following between-group IOIs do appear to be modulated by the abstraction of a beat. Indeed, the temporal violation of an event in reference to a beat grid inevitably requires the abstraction of such a grid. In support of this proposition, we found evidence of the effect of global temporal violation only when exposure was to the (a) metrically stronger of two rhythms (Pattern 1 condition in Experiments 1 & 2) and (b) metrically weaker Pattern 2 when it was strengthened with an accented pulse (Experiment 2). Experiment 3 further clarified the findings of the first two experiments by presenting an exposure pattern that (a) was matched as closely as possible to Pattern 2 in its metrical strength (presented with a woodblock pulse) and (b) allowed for primarily earlier-than-expected global violations at the test block. In this case, learning of the between-group IOIs was demonstrated. The present study is a step toward demonstrating global structure learning, and it highlights the behavioral effects of global structure violation. It calls attention to the importance of considering the effects of both local and global violations when investigating temporal structure learning.

In summary, most prior research investigating the perception and cognition of temporal structure has used either judgement tasks, requiring a response to a final tone in an isochronous sequence (e.g., Jones et al., 2002; Tillmann & Lebrun-Guillaud, 2006), or sequence discrimination tasks that present more complex temporal structures (Handel, 1998; Hébert & Cuddy, 2002; Ross & Houtsma, 1994). In the current set of studies, an implicit learning paradigm was employed to further understand temporal cognition. With the presentation of temporally complex stimuli that required a response to every event, implicit learning of the precise timing between groups of events and of the global temporal structure has been demonstrated. This temporal structure learning was independent of the event structure and was not dependent upon a motor sequence. Yet, it provided a perceptual advantage that enabled participants to make faster responses when their temporal

expectations were maintained but led to slower responses when these expectations were violated.

**Acknowledgements** This research was supported by an Australian Research Council Discovery Project grant (DP0986394) held by Catherine Stevens and Barbara Tillmann, and an Australian Postgraduate Award held by Josephine Terry.

## References

- Bausenhart, K. M., Rolke, B., & Ulrich, R. (2007). Knowing when to hear aids what to hear. *The Quarterly Journal of Experimental Psychology*, 60(12), 1610–1615.
- Bigand, E., & Poulin-Charonnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100, 100–130.
- Boltz, M. (1989). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception and Psychophysics*, 46, 9–17.
- Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics*, 53(6), 585–600.
- Boyer, M., Destrebecqz, A., & Cleeremans, A. (2005). Processing abstract sequence structure: Learning without knowing, or knowing without learning? *Psychological Research*, 69, 383–398.
- Brandon, M., Terry, J., Stevens, C., & Tillmann, B. (2012). Incidental learning of temporal structures conforming to a metrical framework. *Frontiers in Psychology/Frontiers in Auditory Cognitive Neuroscience*, 3, 294. doi:10.3389/fpsyg.2012.00294
- Buchner, A., & Steffens, M. C. (2001). Simultaneous learning of different regularities in sequence learning tasks: Limits and characteristics. *Psychological Research*, 65, 71–80.
- Capizzi, M., Sanabria, D., & Correa, Á. (2012). Dissociating controlled from automatic processing in temporal preparation. *Cognition*, 123, 293–302.
- Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, 2(10), 406–416.
- Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(1), 17–30.
- Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Methods, Instruments and Computers*, 25, 257–271.
- Correa, Á., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing: A review and new evidence from event-related potentials. *Brain Research*, 1076, 116–128.
- Correa, Á., Lupiáñez, J., Milliken, B., & Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Perception and Psychophysics*, 66(2), 264–278.
- Correa, Á., Lupiáñez, J., & Tudela, P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychological Bulletin & Review*, 12(2), 328–334.
- Correa, Á., Lupiáñez, J., & Tudela, P. (2006). The attentional mechanism of temporal orienting: Determinants and attributes. *Experimental Brain Research*, 169, 58–68.
- Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(1), 189–202.
- Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin and Review*, 8(2), 343–350.
- Deutsch, D. (1999). Grouping mechanisms in music. In D. Deutsch (Ed.), *Psychology of music* (pp. 299–348). San Diego, CA: Academic Press.
- Ellis, R. J., & Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. *Attention, Perception, and Psychophysics*, 72(8), 2273–2288.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York, NY: Academic Press.
- Handel, S. (1998). The interplay between metric and figural rhythmic organization. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5), 1546–1561.
- Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, 16(1), 48–55.
- Hébert, S., & Cuddy, L. L. (2002). Detection of metric structure in auditory figural patterns. *Perception and Psychophysics*, 64(6), 909–918.
- Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100(1), 33–72.
- Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, 41(6), 621–634.
- Jones, M. R. (2009). Musical time. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 81–92). Oxford, UK: Oxford University Press.
- Jones, M. R. (2010). Attending to sound patterns and the role of entrainment. In A. C. Nobre & J. T. Coull (Eds.), *Attention and time* (pp. 317–330). Oxford, UK: Oxford University Press.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491.
- Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, 13(4), 313–319.
- Jones, M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception and Performance*, 23(3), 693–709.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(2), 350–364.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32.
- Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 116–126.
- Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, 163, 226–238.
- Penel, A., & Jones, M. R. (2005). Speeded detection of a tone embedded in a quasi-isochronous sequence: Effects of a task-irrelevant temporal regularity. *Music Perception*, 22(3), 371–388.
- Perruchet, P. (2008). Implicit learning. In H. L. Roediger (Ed.), *Cognitive psychology of memory* (Vol. 2, pp. 597–621). Oxford, UK: Elsevier.
- Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2(4), 411–440.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge* (Vol. 19). New York, NY: Oxford University Press.
- Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 585–594.

- Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. *Psychonomic Bulletin and Review*, 14(3), 522–526.
- Ross, J., & Houtsma, A. J. M. (1994). Discrimination of auditory temporal patterns. *Perception & Psychophysics*, 56(1), 19–26.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52.
- Salidis, J. (2001). Nonconscious temporal cognition: Learning rhythms implicitly. *Memory and Cognition*, 29(8), 1111–1119.
- Sanabria, D., Capizzi, M., & Correa, Á. (2011). Rhythms that speed you up. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 236–244.
- Schmuckler, M. A., & Boltz, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception and Psychophysics*, 56, 312–325.
- Schultz, B., Stevens, C., Keller, P., & Tillmann, B. (2013). The implicit learning of metrical and nonmetrical temporal patterns. *The Quarterly Journal of Experimental Psychology*, 66(2), 360–380.
- Scott, R. B., & Dienes, Z. N. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5), 1264–1288.
- Shin, J. C., & Ivry, R. I. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 445–457.
- Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research*, 70, 345–358.
- Tillmann, B., & Poulin-Charronnat, B. (2010). Auditory expectations for newly acquired structures. *Quarterly Journal of Experimental Psychology*, 63(8), 1646–1664.
- Tillmann, B., Stevens, C., & Keller, P. E. (2011). Learning of timing patterns and the development of temporal expectations. *Psychological Research*, 75(3), 243–258.
- van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, 28(1), 43–66.