

## MOTOWN, DISCO, AND DRUMMING: AN EXPLORATION OF THE RELATIONSHIP BETWEEN BEAT SALIENCE, MELODIC STRUCTURE, AND PERCEIVED TEMPO

---

JUSTIN LONDON  
*Carleton College*

BIRGITTA BURGER & MARC THOMPSON  
*University of Jyväskylä, Finland, Jyväskylä, Finland*

MOLLY HILDRETH  
*Carleton College*

JOHANNA WILSON  
*University of Jyväskylä, Finland, Jyväskylä, Finland*

NICK SCHALLY  
*Carleton College*

PETRI TOIVIAINEN  
*University of Jyväskylä, Finland, Jyväskylä, Finland*

**IN A STUDY OF TEMPO PERCEPTION, LONDON, Burger, Thompson, and Toiviainen (2016) presented participants with digitally “tempo-shifted” R&B songs (i.e., sped up or slowed down without otherwise altering their pitch or timbre). They found that while participants’ relative tempo judgments of original versus altered versions were correct, they no longer corresponded to the beat rate of each stimulus. Here we report on three experiments that further probe the relation(s) between beat rate, tempo-shifting, beat salience, melodic structure, and perceived tempo. Experiment 1 is a replication of London et al. (2016) using the original stimuli. Experiment 2 replaces the Motown stimuli with disco music, which has higher beat salience. Experiment 3 uses looped drum patterns, eliminating pitch and other cues from the stimuli and maximizing beat salience. The effect of London et al. (2016) was replicated in Experiment 1, present to a lesser degree in Experiment 2, and absent in Experiment 3. Experiments 2 and 3 also found that participants were able to make tempo judgments in accordance with BPM rates for stimuli that were not tempo-shifted. The roles of beat salience, melodic structure, and memory for tempo are discussed, and the TAE as an example of perceptual sharpening is considered.**

*Received: January 15, 2018, accepted June 16, 2019.*

**Key words:** rhythm, beat salience, tempo, absolute versus relative tempo judgment, perceptual sharpening

---

**T**HE TEMPO OF A PIECE OF MUSIC IS RARELY, IF ever, ambiguous. After only a few notes or drum strokes we have a keen sense of whether the music is fast, moderate, or slow, and in that same brief span of time we have entrained to the music’s beat (Large, Herrera, & Velasco, 2015; Large & Palmer, 2002; Phillips-Silver et al., 2011; Tierney & Kraus, 2015). Yet the cues for tempo are neither simple nor straightforward. Tempo assessments involve more than just a determination of the rate of the primary pulse or beat, as salient cues include event density (Drake, Gros, & Penel, 1999; Elowsson & Friberg, 2013), metrical structure (London, 2011; Madison & Paulin, 2010), loudness, register, and timbre (Boltz, 2011; Eitan & Granot, 2009), and the listener’s own motor activity (Drake et al., 1999; London, 2011). Furthermore, cross-modal cues can also affect perceived tempo, as we have shown in a previous paper (London, Burger, Thompson, & Toiviainen, 2016). In that paper we also obtained an unexpected result. In that study, participants were presented with 10-second clips of Motown R&B songs in a variety of audio and/or video contexts, and they were asked to make a tempo judgment of each clip using a seven-point Likert-type scale. To forestall simple associations between a given song and a tempo judgment, participants were presented with both original and digitally “tempo-shifted” versions of each stimulus. Tempo-shifting varies the speed of an audio file without altering its pitch and timbral characteristics, at least when done to a moderate degree (in our case,  $\pm 5\%$ ). Stimuli were presented at three core BPM levels (105, 115, and 130 BPM), along with their tempo-shifted variants, creating a stimulus set that spanned the range from 100 to 135 BPM in 5 BPM increments.

As a baseline condition, we had our participants rate the audio stimuli without the paired videos. Given that absolute tempo memory for familiar songs can be quite

accurate (Bergeson & Trehub, 2006; Levitin & Cook, 1996), and that listeners are able to discern original versus tempo-shifted versions of songs (Honing, 2006, 2007), we presumed that participants would be able to sort out the original versus tempo-shifted versions of each stimulus without difficulty, and that their ratings would correspond to the BPM rates of the stimuli, as we had endeavored to keep other tempo cues relatively constant amongst the stimuli (see Method section). This did not occur. Our participants were able to correctly sort out the original versus tempo-shifted versions of each song (i.e., making veridical *relative* tempo judgments), even though the stimuli were presented in different random orders for each participant. But their tempo ratings did not exhibit a simple correspondence to the BPM rates of the stimuli. Instead, our participants consistently inflated the ratings of sped-up versions of a song and correspondingly deflated the ratings of slowed-down versions. For example, the average ratings of sped-up songs from the slowest tempo group (originally 105 BPM, now 110 BPM) were higher than average ratings of songs from the middle tempo group at their original tempo level of 115 BPM. Our participants thus failed to make veridical *absolute* tempo judgments, and we described this as the “Tempo Anchoring Effect” (TAE), as the perceived tempo of each song is “anchored” around the BPM rate of the original version. The TAE, more generally, is an over or underestimation of the absolute speed/tempo of musical patterns that are recognizable as “sped-up” or “slowed-down” versions (respectively) of patterns that have been previously heard. In London et al. (2016) we hypothesized that in order for the TAE to occur, one must have stimuli that are both memorable and musically distinctive, such that one can recognize faster versus slower versions of the same stimulus/piece of music. As a corollary, we hypothesized that the TAE will not arise in the context of rhythmically impoverished stimuli such as metronome clicks or even fairly complex stimuli that lack a distinct musical/sonic identity. More broadly, the TAE gives additional evidence that tempo judgments are based upon more than the BPM rate of the music, and thus in addition to other musical cues such as event density, loudness, register, and contour (as noted above), our memory of a specific recording or performance can also affect our tempo perception when we re-encounter the music anew.

Here we report on three experiments that systematically investigated the TAE and its relation to beat-based cues for musical tempo. In the first, our primary aim was to replicate our original findings. We presented a different group of participants with our original

Motown stimuli, but we also included a simple rock drumming pattern (presented at all BPM rates present in the stimulus set) to confirm participants’ ability to make tempo judgments of simple rhythmic stimuli that correspond to their BPM rates. In the second experiment a set of disco songs were used as stimuli, as their characteristic “four on the floor” rhythms have exceptionally high beat salience, and presumably BPM rate would be maximally salient as a tempo cue in this style of music. Tempo-shifted versions of disco songs were presented at three core BPM levels, analogous to the presentation of tempo-shifted stimuli in Experiment 1. In the second experiment we also presented participants with a set of Disco songs that were not tempo-shifted, to confirm their ability to make absolute tempo judgments that correspond to the beat rate when the TAE is not present. For the third experiment, we created a set of drum-pattern stimuli to be analogs to the stimuli used in Experiments 1 and 2. While the drum stimuli obviously lack melodic or harmonic cues, we were able to create a set of distinctive stimuli within the various BPM sub-ranges that were percussive analogs to the stimuli used in Experiments 1 and 2.

Our primary hypothesis is that for tempo-shifted stimuli, as beat salience increases and as the salience of other cues (melody, harmony, lyrics, etc.) decreases, the TAE will decrease, and participant tempo ratings will correspond more closely to the BPM rate of the stimuli. The stimuli chosen in these three experiments are designed to systematically increase beat salience (moving from Motown to Disco and Drumming) and decrease other cues (from Motown and Disco to Drumming). Our secondary hypothesis is that when tempo-shifting is not involved, participant tempo ratings will align with the BPM rate, all other cues being equal.

### Experiment 1: Replication of the TAE

In this experiment we sought to replicate the TAE, reusing our Motown stimuli with a different group of participants. We also included a generic rock drumming pattern, presented across the entire BPM range used in the experiment, to determine (a) if the rating scale range was fine-grained enough to distinguish the BPM rates used in the Motown stimuli, and (b) if our participants’ ratings of simple rhythmic stimuli would correspond to their BPM rates.

#### METHOD

*Participants.* Twenty-one participants (13 female) were recruited from the Carleton and Northfield, MN community for Experiment 1. The average age was 22.8

TABLE 1. Musical Stimuli Used in Experiment 1

Artist	Title	BPM	R&B Chart Ranking	Spotify Plays (x10 <sup>6</sup> )	Notes/Beat	Flux
Temptations	Get Ready	134.5	#1 (1966)	34.091	1.50	3.444
Supremes	Where Did Our Love Go?	133	#1 (1964)	29.643	1.13	1.953
Supremes	Stop, In the Name of Love	117	#2 (1964)	37.191	1.38	2.888
Wilson Pickett	The Midnight Hour	113	#1 (1965)	22.051	1.54	2.678
Stevie Wonder	Signed, Sealed, Delivered	105.5	#1 (1970)	144.603	1.29	2.912
Temptations	My Girl	103	#1 (1964)	214.891	1.38	2.932

Note: BPM rates subsequently corrected to 130, 115, and 105 core BPM levels; Chart rankings are from Billboard Charts Archive (<https://www.billboard.com/archive/charts>); Flux is in the 100–200 Hz sub-band.

years ( $SD = 7.9$  years, mostly due to two older participants ages 37 and 53 years; all other participants were between ages of 18–23 years). Six participants had more than 10 years of music training, eight had 5–10 years of training, and seven had less than 5 years of training. Ten participants were familiar with five or six of the songs used as stimuli in the experiment, while six participants were familiar with two or fewer songs. Participants were not directly compensated for their participation, but were entered into a drawing for a gift card from a local coffee shop.

*Stimuli.* The sources of the stimuli used in Experiment 1 are given in Table 1.

To ensure a balanced set of stimuli at each tempo level, a score-based analysis indexed the number of notes of the vocal melody, bass, and percussion parts in each bar. Based upon these analyses, an aggregate rhythmic density score was calculated for each song to ensure matched pairs at each tempo level (see Table 1). As a companion measure, we also analyzed the degree of spectral flux in each of our stimuli. Spectral flux is the moment-to-moment change in energy in particular frequency ranges of the auditory spectrum/acoustical signal, and in particular, low frequency flux has been shown to be related to rhythmic features in music (Burger, Ahokas, Keipi, & Toiviainen, 2013), as well as to characteristics of music-induced movement (Burger et al., 2013). The low-frequency spectral flux for each song was calculated by choosing an octave-wide frequency range between 100 and 200 Hz and calculating the sub-band flux using the MatLab-based MIRtoolbox “mirflux” function (Lartillot & Toiviainen, 2007) by taking the Euclidean distances of the power spectra for each two consecutive frames of the signal, using a frame length of 25 ms and an overlap of 50% between successive frames and then averaging the resulting time series of flux values obtained over the entire length of each stimulus. The average spectral flux values for the non-tempo-shifted stimuli are indicated in the “Flux” column of Table 1. Pair of songs at each core BPM level were

chosen whose average flux values were commensurate across the different BPM levels.<sup>1</sup>

All stimuli were in simple duple meter (i.e., 4/4) with a light amount of “swing” (i.e., no overt triplet subdivision), and all are well established members of the American 1960’s R&B canon, as indicated by their chart positions.<sup>2</sup> To assess how familiar participants might be with these Motown songs, we also looked at the number of times a song had been played on Spotify, one of the most widely used music streaming services in the world. To put these figures in context, the most-streamed songs have more than 1 billion streams/listens; Ed Sheeran’s “The Shape of You” tops the list with 1.87 billion streams.<sup>3</sup>

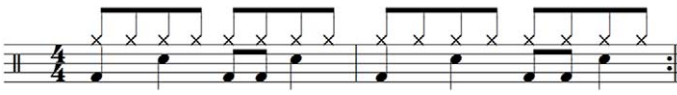
Stimuli were initially chosen based upon their original BPM rates, as we sought three distinct BPM levels within the range of 100–130 BPM and we wanted to avoid aggressively tempo-shifting the stimuli, as tempo-shifts beyond 5–7% can introduce noticeable audio artifacts. The range of 100–130 BPM is where pulse cues are maximally salient (Parncutt, 1994) and where tempo judgments are most accurate (Fraisse, 1984; Grondin, 2010; Madison & Paulin, 2010; Repp, 2005). Objective BPM measurements of the stimulus source files were determined by averaging the results of two independent raters who tapped along to each song using the “Quick Tempo” iPhone App (ver. 1.1) and were also checked using the MIRtoolbox “mirtempo” function; agreement between the raters and the software was within 1 BPM.

In Experiment 1 there were two songs at each Core BPM Level. Songs were lightly tempo-shifted to produce BPM rates precisely at each core BPM level (105, 115,

<sup>1</sup> Flux rates reported here differ from those reported in London et al. (2016), as here the values were normalized according to their RMS volume to allow for comparison across the three stimulus sets/experiments used here.

<sup>2</sup> As obtained from the Billboard Magazine historical “Hot 100” Chart archives: <https://www.billboard.com/archive/charts>

<sup>3</sup> As of May 2018, according to [https://en.wikipedia.org/wiki/List\\_of\\_most\\_streamed\\_songs\\_on\\_Spotify#Artists\\_with\\_the\\_most\\_songs\\_in\\_the\\_Spotify\\_Global\\_100](https://en.wikipedia.org/wiki/List_of_most_streamed_songs_on_Spotify#Artists_with_the_most_songs_in_the_Spotify_Global_100)

(a) 

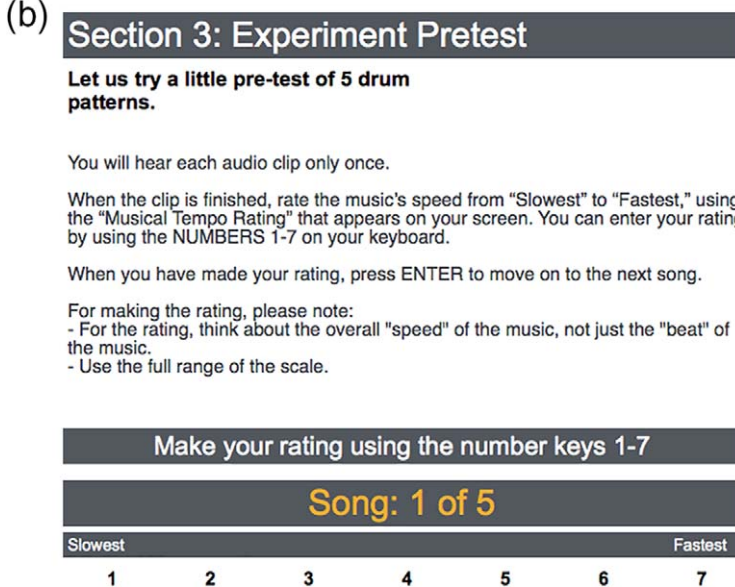
(b) 

FIGURE 1. (a) Generic rock drumming pattern used in pre-test and in Experiment 1; lowest notes = kick drum; middle notes = snare; highest notes (with crosses for note heads) = hi-hat. (b) Screenshot of the stimulus presentation and response interface used in all three experiments.

and 130 BPM), and then again to produce versions that were 5 BPM faster or slower, yielding a 3 x 3 factorial design (CoreBPM x Shift), with two stimuli at each level, for a total of 18 song stimuli. All stimuli were 10 seconds in duration and began on the first significant downbeat following the introductory portion of each song. In addition to the song stimuli, 10-second loops of a standard rock drumming pattern (see Figure 1a below) were presented at 100–135 BPM in 5 BPM increments, hence 8 additional stimuli, for a total of 26 stimuli.

*Apparatus and procedure.* The experimental task in each trial involved making a tempo rating of an individual stimulus using a 7-point Likert-type scale. Each experimental session lasted approximately 20 minutes and consisted of an introduction and pretest, followed by the main trials. Stimuli were presented in a random order for each participant, with the randomization constrained so that different versions of the same stimulus were not presented consecutively. Participants heard each stimulus once.

The introduction and pretest began with a set of demonstration songs at the low and high ends of the tempo range, as well as tempo-shifted versions of the same Motown R&B song to familiarize participants with the

type of stimuli used in the experiment (demo songs were not used in the experiment). Participants were thus not naive to the fact that they would be presented with original as well as tempo-shifted versions of the stimuli; this forestalls any simple association between a particular song and a single tempo rating. The pretest then presented participants with a simple rock drumming pattern (kick drum, snare, and hi-hat sounds; see Figure 1a) to precisely indicate the range of tempos used in the experiment (100–135 BPM) as well as to familiarize them with the response interface and tempo rating procedure. Figure 1b is a screenshot of the stimulus presentation and response interface used in both the pre-test and the main trials. Note that on every trial participants are reminded to focus on the overall speed of each stimulus (rather than simply its BPM rate), to use the full range of the 7-point scale, and that 1 = *slow* and 7 = *fast*.

Stimuli were presented to participants in quiet rooms on MacBook Pro laptop computers (13-inch screen size, 2.65GHz Intel Core i5, with 3 or 4GB RAM, running OS 10.10.3), using an updated version of the Max/MSP patch used in London et al. (2016). Participants listened via Sennheiser HD280 Pro headphones, which provided

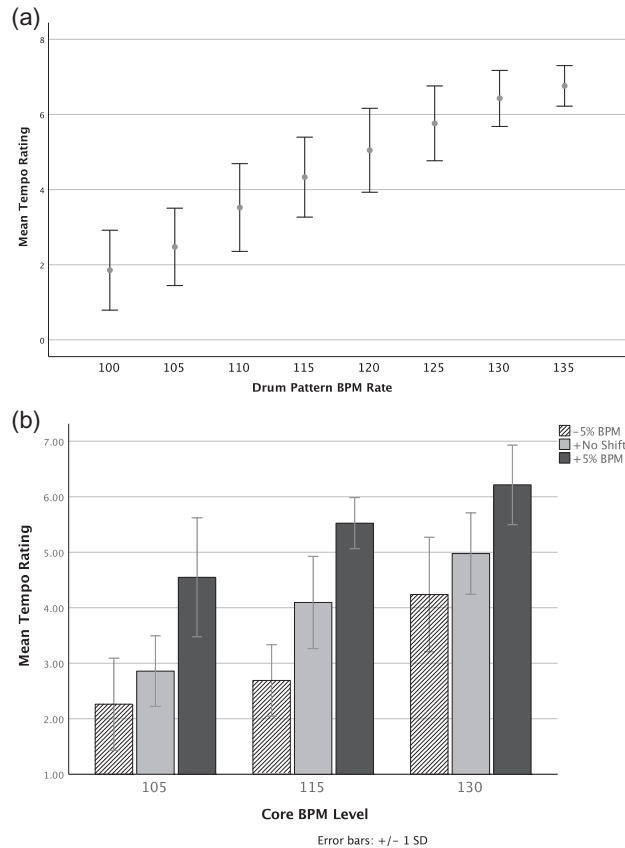


FIGURE 2. (a) Average tempo ratings of rock drumming stimuli in Experiment 1. (b) Average participant tempo ratings ( $y$ -axis) for Motown stimuli in Experiment 1, grouped by core BPM levels. Here and in all subsequent figures error bars indicate 1 standard deviation.

additional attenuation of ambient noise, with the headphone volume adjusted to a comfortable listening level. Participants were able to provide a response only after the entire stimulus had been presented. After making their response, participants then cued the next stimulus. Between stimuli there was a variable 4-5 second delay to minimize carry-over effects of auditory beat entrainment (London, 2012; van Noorden & Moelants, 1999), during which participants heard random environmental sounds—surf, seagulls, etc. Once the pre-test was complete, the experimenter left the room to avoid any biasing of the participant's responses. After all experimental trials were completed, participants' background data were collected using a computer-based questionnaire.

## RESULTS

*Drum pattern tempo ratings.* Participant tempo ratings for the drum stimuli interleaved amongst the song stimuli are given in Figure 2a.

A repeated-measures ANOVA found a significant effect for BPM rate,  $F(7, 140) = 80.57, p < .001, \eta_p^2 = .801$ . Even though the rock drum pattern stimuli appeared within the context of rating original and tempo-shifted versions of real songs, participants had no trouble producing absolute tempo ratings that corresponded to the different BPM rates of the drum stimuli (Pearson's  $r = .99, p < .001$ ).

*Song tempo ratings.* Figure 2b shows participant tempo ratings at each BPM level; the TAE is again clearly apparent. A 3 x 3 repeated-measures ANOVA (Core BPM x Shift) found main effects for Core BPM,  $F(2, 82) = 71.88, p < .001, \eta_p^2 = .637$ , and Shift,  $F(2, 40) = 197.04, p < .001, \eta_p^2 = .828$ . There was a small but significant interaction between BPM and Shift,  $F(4, 164) = 3.48, p = .009, \eta_p^2 = .073$ , as the effect of tempo-shifting was greater at 115 BPM than at the other two Core BPM levels (see Figure 2).

In summary, Experiment 1 replicated the TAE, but also found that participants were able to make veridical tempo ratings of drum stimuli interleaved amongst the Motown song stimuli. This indicates that the TAE is not due to either (a) participants' inability to make differentiated absolute tempo judgments within the range of stimulus presentation, and relatedly (b) that the rating scale used in the experiment was/is an adequate metric for making the requisite discriminations within the range of tempos used in the experiment.

## Experiment 2: Disco Stimuli

Having replicated the TAE in Experiment 1, Experiment 2 was conducted to see if the TAE would occur with a different set of stimuli. Disco stimuli were chosen as replacements for the Motown R&B stimuli as they were a different popular musical style that was stylistically similar to the original Motown stimuli, yet (a) would be relatively unfamiliar to most participants, and (b) would have a high degree of beat salience, as a strong beat presence is a hallmark of Disco music. Two sets of Disco stimuli were used in Experiment 2. In Experiment 2a we presented participants with original and tempo-shifted versions of stimuli at three core BPM levels, along with generic drum patterns across the BPM range used in the experiment, i.e., a Disco analog to Experiment 1. Disco songs also feature extended instrumental introductions, and so in Experiment 2a we also included intro versus vocal sections of each stimulus, to see if the TAE was related to cues specific to the vocal part of the musical texture. In Experiment 2b participants were presented with stimuli that were not tempo-shifted and which spanned the 100-130 BPM range as a further

TABLE 2. Source Stimuli Used in Experiments 2a and 2b

Song Name	Artist	BPM Level	Flux	Notes/ Second	Spotify (x1000)
Experiment 2a					
Get Dancin'	Disco Tex & The Sex-O-Lettes	105	6.570	5.890	34
Stayin' Alive	The Bee Gees		6.170	5.683	222,446
Forget Me Nots	Patrice Rushen	115	6.936	5.345	12,000
He's the Greatest Dancer	Sister Sledge		9.624	7.980	23,033
Disco Nights (Rock-Freak)	G. Q.	125	8.164	6.774	1,500
Cruisin the Streets	The Boystown Gang		7.325	6.346	21
Experiment 2b					
Get Dancin'	Disco Tex & The Sex-O-Lettes	105	6.570	5.890	34
Stayin' Alive	The Bee Gees		6.170	5.683	222,446
Keep it Coming Love	K.C. & The Sunshine Band	110	7.008	6.266	1,800
Dance With Me	Peter Brown		8.885	6.999	288
Forget Me Nots	Patrice Rushen	115	6.936	5.345	12,000
He's the Greatest Dancer	Sister Sledge		9.624	7.980	23,033
Bad Girls	Donna Summer	120	8.973	7.345	16,000
I.O.U.	Freez		8.581	7.423	2,000
Disco Nights (Rock-Freak)	G. Q.	125	8.164	6.774	1,500
Cruisin the Streets	The Boystown Gang		7.325	6.346	21
Get Off	Foxy	130	10.650	8.115	1,240
Instant Replay	Dan Hartman		8.519	7.165	2,058

Note: Stimuli in Experiment 2a were time stretched -5, 0, and +5%. Flux is the sub-band between 100–200 Hz.

assessment of the nature of their tempo judgments when relative tempo comparisons were not possible.

#### METHOD

**Participants.** Thirty-six participants (22 female) were recruited from the University of Jyväskylä community via e-mail and social media advertisement. Their ages ranged from 20–40 years (mean = 26.3 years,  $SD = 4.43$  years, median age = 25 years). Fifteen participants had received no music training, 11 had received 1–4 years of training, and 10 had received 5 or more years of training. All participants were unfamiliar with most of the twelve songs used in the experiment, having previously heard 4 or fewer of the songs, save for one participant who indicated a familiarity with 5–8 songs. Twenty-three participants were Finnish; other nationalities (one each) were Turkish, Russian, Greek, Dutch, Australian, Indian, Ethiopian, Moroccan, Vietnamese, South African, Chilean, Costa Rican, and Ukrainian. Informed written consent was obtained from each participant prior to their participation, and each participant was compensated with a voucher for a movie ticket (worth  $\approx 10.00$  €).

**Stimuli.** Source files for the stimuli used in Experiment 2 were taken from *The Disco Box* (Rhino Records 1999, ASIN: B00000HZEM), a compilation of 80 greater and lesser-known disco songs from the mid-1970s to the mid-1980s. Additional songs were taken from *Bee Gees Greatest* (Rhino Records 2007, ASIN B000TQZ7NA). As

can be seen in Table 2, the number of Spotify streams (as of May 2018) for our stimuli ranged from 21,000 (“Cruisin’ The Streets”) to 16 million (“Bad Girls”). Thus while a few disco songs are somewhat well known (e.g., those by the Bee Gees or Donna Summer), they are not nearly as well-known as the Motown stimuli used in Experiment 1, and most were relatively unfamiliar for the majority of our participants, as confirmed by our participant survey (reported above).

To select real musical stimuli that were nonetheless as similar as possible in terms of tempo cues aside from BPM rate, as in Experiment 1 the acoustic features of potential stimuli were assessed using the MIR Toolbox for their initial BPM rate, sub-band spectral flux, and notes per second (event density). Stimuli chosen for use in the experiment all had sub-band flux values in a narrow range, and an average of 4–6 notes/second (see Table 2). Stimuli were also chosen based upon their original BPM rates as assessed by MIR toolbox and confirmed by two researchers who tapped along to the beat rate using the “Quick Tempo” iPhone App (ver. 1.1), as we again wanted to minimize artifacts introduced due to time shifting. In a slight contrast to Experiment 1, and in pursuit of a more uniform set of BPM levels, we selected stimuli whose original BPM rates were at or near 105, 115, and 125 (as opposed to 130) BPM. This yielded a set of stimuli spanning the range from 100 to 130 BPM, with overlaps at 110 and 120 BPM. Tempo-shifted versions of stimuli used in Experiment 2a were produced

using the “Tempo-shift” plugin in Avid Pro Tools (ver. 12.2.1), with time shifting calibrated to produce precise increments of  $\pm 5$  BPM.

In Experiment 2a there were two songs at each Core BPM Level, with stimuli taken from the instrumental introduction (“intro”) and from the verse or chorus (“vocal”) of each song. Songs were lightly tempo-shifted to produce BPM rates precisely at each core BPM level (105, 115, and 125 BPM), and again to produce versions that were 5 BPM faster or slower. Thus Experiment 2a used a  $3 \times 3 \times 2$  factorial design (CoreBPM  $\times$  Shift  $\times$  Intro/Vocal), with two stimuli at each level, for a total of 36 song stimuli. In addition, and as in Experiment 1, a rock drum pattern was presented at 100–130 BPM (in 5 BPM increments), yielding 7 additional stimuli, for a total of 43 stimuli. Each participant heard each stimulus once.

In Experiment 2b, songs were presented at each of six distinct BPM rates from 105–130 BPM in 5 BPM increments, with two songs at each BPM level. Each song was only used at one BPM level, and hence there were 12 distinct stimuli in Experiment 2b (see Table 2). Here the aim is to assess the ability of participants to make tempo ratings when each stimulus is presented only at a single BPM rate (i.e., when no tempo-shifting is involved). As in Experiment 2a, we chose stimuli whose original BPM rates were close to the target BPM rate for the stimuli, and then tempo-shifted them so their BPM rate precisely matched the target rate. We were unable to find suitable stimuli at or near 100 BPM, and as we did not want to use more aggressively tempo-shifted stimuli, we omitted this BPM level from the stimulus set in Experiment 2b.

*Apparatus and procedure.* The pretest and presentation of stimuli used the same Max/MSP interface as in Experiment 1, save that Disco songs were used instead of R&B songs to illustrate fast vs. slow songs, as well as tempo-shifted versions of the same song. The same set of rock drum patterns used in Experiment 1 was used here to demonstrate the user interface and illustrate the range of BPM levels used in the experiment.

Experiments 2a and 2b were done in a single session with a short break between them; order of experiments was counterbalanced among participants. Data were also gathered for a third experiment (similar in design and length to 2b); those results are not reported here. Stimuli were presented in different random orders for each participant, with randomization controlled so they did not hear tempo-shifted versions of the same stimulus presented consecutively. Stimuli were presented in a quiet room on a 13-inch MacBook Pro (2.6 GHz Intel Core i5, High Sierra version 10.13.5) using the customized Max Patch described in Experiment 1 (Max/MSP

version 7.3.5). Participants listened using a pair of Audio Technica ATH-M50x over-the-ear headphones that provided additional noise attenuation. As in the first experiment, 4–5 seconds of random environmental sounds—surf, seagulls, etc.—were presented between each trial. The entire experiment took 40–50 minutes to complete.

## RESULTS

*Participant factors.* A few modest yet statistically significant results regarding the music training of our participants were found in Experiments 2a and 2b. For the drum stimuli included among the disco songs in Experiment 2a (but not the primary stimuli), a mixed model ANOVA, with music training as the between-participants variable, and BPM level as the within-participant variable found a significant effect of Training,  $F(2, 33) = 5.70, p = .007, \eta_p^2 = .257$ , as participants with 5+ years of music training gave faster ratings for the drum stimuli. In Experiment 2b a similar mixed model ANOVA again yielded a significant result for Training,  $F(2, 33) = 4.22, p = .023, \eta_p^2 = .204$ , as participants with no music training rated the primary stimuli slower than those with music training.

*Experiment 2a.* The tempo ratings for the drum stimuli in Experiment 2a are given in Figure 3a.

A repeated-measures ANOVA found a significant effect for BPM Level,  $F(6, 210) = 154.48, p < .001, \eta_p^2 = .815$ , unsurprisingly. Post hoc pairwise *t*-tests for each BPM level were all significant ( $p < .006$ ), save for the two fastest BPM rates (125 vs. 130 BPM), where the difference was *ns*, indicative of some compression at the top of the rating scale.

A  $3 \times 3 \times 2$  repeated-measures ANOVA (CoreBPM  $\times$  Shift  $\times$  Intro/Vocal) found main effects for CoreBPM,  $F(1.75, 124.09) = 215.54, p < .001, \eta_p^2 = .752$  (Greenhouse-Geisser correction applied), and Shift,  $F(1.78, 126.68) = 334.42, p < .001, \eta_p^2 = .825$  (Greenhouse-Geisser correction applied), as was expected. There was a modest effect of Intro vs. Vocal,  $F(1, 71) = 4.76, p = .032, \eta_p^2 = .063$ . Overall vocals were rated slightly faster than intros (grand mean of 4.37 vs. 4.19); the primary cause may be greater event density due to vocal line on top of other parts of the accompaniment, as the accompaniment served as the primary musical material of the introduction. Figure 3b gives the results of the variables of interest relative to the TAE, CoreBPM, and Shift. The TAE is clearly evident, though to a lesser degree than in Experiment 1 (see below for further discussion). Several interactions in Experiment 2a were statistically significant. There was an interaction between Core BPM and Intro/Vocal,  $F(1.86,$

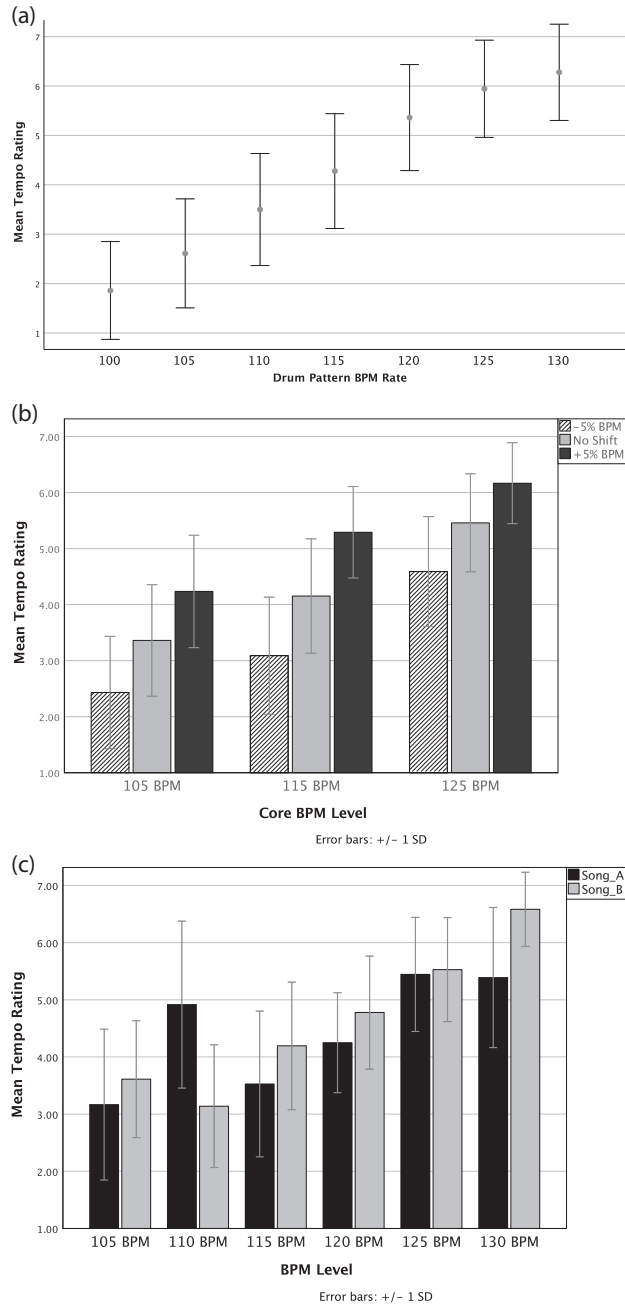


FIGURE 3. (a) Average tempo ratings of rock drumming stimuli in Experiment 2a. (b) Average participant tempo ratings (y-axis) for time-shifted disco stimuli in Experiment 2a, grouped by core BPM levels (N.B., Intro vs. Vocal ratings averaged at each stimulus level). (c) Average participant tempo ratings (y-axis) for non-shifted disco stimuli in Experiment 2b, grouped by BPM level.

131.68) = 12.30,  $p < .001$ ,  $\eta_p^2 = .148$  (Greenhouse-Geisser correction applied), as at the 105 BPM level vocals were rated faster than intros (mean ratings of 3.53 vs.

3.16), while intros were rated faster than vocals at the 115 BPM level (4.28 vs. 4.08, respectively). There was also an interaction between Core BPM and Shift,  $F(3.44, 244.54) = 2.79$ ,  $p = .027$ ,  $\eta_p^2 = .038$  (Greenhouse-Geisser correction applied), as differences in average tempo ratings (-5% vs. no shift vs. +5%) were greater at the 115 BPM level, than at the 105 or 125 BPM levels.

*Experiment 2b.* In Experiment 2b pairs of musical stimuli were presented at each of the BPM rates used in Experiment 2a, save for the slowest BPM rate (see Table 2). Mean participant ratings for each stimulus are given in Figure 3c. The anomalous song at 110 BPM is “Keep It Coming Love,” and as can be seen in Table 2, this song had the highest sub-band spectral flux value of any of the stimuli. A 6 x 2 repeated-measures ANOVA (BPM Level x Song) found a main effect for BPM Level,  $F(3.90, 136.49) = 77.35$ ,  $p < .001$ ,  $\eta_p^2 = .688$  (Greenhouse-Geisser correction applied), and a nearly significant effect of Song,  $F(1, 35) = 4.04$ ,  $p = .052$ ,  $\eta_p^2 = .104$ . As one would expect from Figure 3c, there was a significant interaction between BPM Level and Song,  $F(3.96, 138.68) = 21.460$ ,  $p < .001$ ,  $\eta_p^2 = .380$ , due to the particular ratings at the 110 and 130 BPM levels. When ratings for songs at each BPM level are combined, post hoc pairwise comparisons show differences between BPM levels are all significant ( $p < .008$ ), save for 105 vs. 115 ( $p = .072$ ), 110 vs. 115 ( $p = 1.00$ ), and 110 and 120 ( $p = .085$ ). Thus while there were some song-specific factors that affected participant ratings (e.g., “Keep It Coming Love”), participants were able, for the most part, to rate the tempos of the songs in Experiment 2b in accordance with their BPM rates.

### Experiment 3: Drum Stimuli

In Experiment 3 a set of drum patterns were used as stimuli. Drum patterns, by their very nature, lack most of the musical cues present in the stimuli used in Experiments 1 and 2 (melody, harmony, lyrics). To create stimuli in Experiment 3 that were nonetheless analogs to the stimuli used in Experiments 1 and 2, two sets of stimuli were prepared that differed in complexity, assessed in terms of their offbeatness/syncopation and note density. Two levels of complexity were included to assess (a) if complexity itself was a factor in judgments of tempo, and (b) to ensure that we had created stimuli of sufficient musical richness to be analogs to the stimuli used in Experiments 1 and 2.

#### METHOD

*Participants.* Twenty-eight participants (12 female) were recruited from the Northfield, MN area via



e-mail, posters, social media, and personal contacts. Their ages ranged from 18–69 years (mean = 35.6 years,  $SD = 18.1$  years, median = 23.5 years). Twenty-three participants had received some sort of music training as a child, and 16 of those currently practice an instrument or sing at least once a week; these 16 we operationally defined as “musicians.” Out of these 16 musician participants, six played the drums as one of their instruments. One participant self-identified as British, one as Tibetan, one as Canadian, and three as Chinese; the remaining 22 were US/American. Participants were not compensated but were instead entered in a drawing to win a \$20 gift card to a local coffee shop.

*Stimuli.* The stimuli were chosen to function as analogs to the tempo-shifted stimuli used in Experiments 1 and 2, and were presented at the same core BPM levels. Stimuli were drawn from an initial list of 52 two-bar patterns: 35 from the stimuli used by Witek, Clarke, Wallentin, Kringelbach, and Vuust (2014), three from a list of the “most sampled drum breaks,”<sup>4</sup> and 14 created by author JL, based upon patterns in Witek’s set of stimuli. Patterns consisted of three percussive layers: kick drum, snare, and hi-hat. The hi-hat played constant eighth notes in each pattern, whereas the kick and snare varied. A total of eight patterns, all in 4/4 meter, were then selected for use in Experiment 3 based on the complexity criteria given below; four were drawn from Witek et al. (2014), while the other four were created by authors JL and NS (see Appendix).

Stimulus complexity was assessed using Toussaint’s (2013) “weighted offbeatness” (WO) measure. Depending upon its metric position, a note or drumstroke can be regarded as more or less “offbeat” (or conversely, more or less “onbeat”). The offbeatness of each note in a standard 4/4 measure—the meter used for all of the stimuli in the current experiment—can range from 0 (for a note on the downbeat of the measure) to 4 (for a note on the offbeat on the lowest level of sixteenth-note subdivision). The WO for a given rhythm is simply the sum of the offbeatness values of each element in the rhythm. WO is correlated with the degree of syncopation in a rhythmic figure, but it is also naturally correlated with the number of notes in a pattern; patterns with a larger number of notes will have higher WO values. Patterns with a large number of notes will also have a large number of on-beat onsets, and thus may not sound particularly offbeat/syncopated. Therefore, we normalized Toussaint’s WO (“WOn”) simply by

TABLE 3. *WO, Note Counts (Hi-hat Excluded), WOn, and 100–200hz Sub-band Spectral Flux for 2-bar Versions of the Stimuli Used in Experiment 3*

Stimulus	WO	WOn	Note Count	Flux
JL_01	24.00	2.00	12	7.657
JL_02	27.00	2.08	13	8.462
witek_s03	16.00	1.60	10	10.307
witek_s04	18.00	1.80	10	9.451
JL_03	44.00	2.59	17	6.174
JL_04	46.00	2.56	18	7.336
witek_s18	43.00	2.69	16	8.835
witek_s28	40.00	2.50	16	8.613

dividing it by the number of notes in the pattern. Table 3 lists the WO, WOn, note count, and sub-band spectral flux for each two-bar stimulus used in the first part of Experiment 3 (spectral flux was measured in the manner described in Experiment 1).

As can be seen in Table 3, WOn values range from 1.60 to 2.08 for our simple stimuli (10–13 elements in each), versus 2.50 to 2.69 for our moderately complex stimuli (16–18 elements in each).

Stimuli were prepared in Logic X Pro (Version 10.4.1), using the “SoCal” set of drum samples as the basic sound elements for each stimulus. Stimuli were created in MIDI, with MIDI control of BPM rates; three versions of each stimulus (simple vs. complex) were prepared at each of the three core BPM levels (100, **105**, and 110 BPM; 110, **115**, and 120 BPM; 120, **125**, and 130 BPM; bold values mark the “Core BPM” rates). Thus as in Experiment 1, there were 18 stimuli (3 Core BPM Levels x 3 Shift Levels x 2 Complexity Levels). Note that while these versions were created in MIDI, we will refer to them as “tempo-shifted” at each core BPM level, for ease of comparison with Experiments 1 and 2. An additional pair of stimuli (simple vs. complex) were presented from 100 to 130 BPM, in 5 BPM increments, yielding 14 “All BPM” stimuli (7 BPM levels x 2 Complexity Levels); these stimuli were analogs of the drum stimuli presented in Experiments 1 and 2a, hence 32 total stimuli. To forestall any effects that might be due to a single drum pattern being presented across all tempo levels, two stimulus sets (Block A vs. Block B) were prepared, with alternative versions of the “All BPM” stimuli used in different blocks. Half of the participants heard Block A, and half heard Block B for each experiment. The block design of the two parts of Experiment 3 is summarized in Table 4.

Performances were “deadpan” (i.e., without any expressive timing alterations). Volumes of each of the drum layers (kick-drum, snare, and hi-hat) were

<sup>4</sup> Taken from a list of the 10 most sampled drum breaks: <https://music.avclub.com/these-are-the-breaks-10-of-the-most-sampled-drum-beats-1798283974>

TABLE 4. Block Design of Stimuli Used in Experiment 3

Core BPM	Block A	Block B
105	Simple #1	Simple #1
	Complex #1	Complex #1
115	Simple #2	Simple #4
	Complex #2	Complex #4
125	Simple #3	Simple #3
	Complex #3	Complex #3
ALL	Simple #4	Simple #2
	Complex #4	Complex #2

Note: Core BPM includes  $\pm$  5BPM tempo-shifted versions of each stimulus.

matched by ear. Once each pattern was created (either one, two, or four bars in length), it was then extended using the loop function in Logic Pro X, (v. 10.2.2) to create segments that were between 12 and 16 seconds long, depending upon BPM rate. Stimuli were exported as .wav audio files. To avoid end-of-file artifacts as well as confounding cues, each stimulus ended with a 1.5–2.0 second fade-out, created in Audacity (Version 2.2.2).

*Apparatus and procedure.* The interface and presentation of stimuli used a variant of the Max/MSP patch used in Experiments 1 and 2, with the same introduction and pretest (using the same basic rock drum pattern), save that instead of example songs, example drum patterns were used to introduce the distinction between “simple” versus “complex” drum patterns, as well as the same pattern presented at different speeds (demo patterns were not used in the experiment). Stimuli were presented in random orders for each participant; randomization was controlled so that participants did not hear more than three consecutive stimuli at the same BPM rate. The experiment took 15–20 minutes to administer; during this experimental session participants provided data for a similar experiment whose results are not reported here. Stimuli were presented on a 13-inch MacBook Air (2.3 GHz Intel Core i7, High Sierra Version 10.13.5) using a customized Max Patch in Max/MSP (Version 7.3.5). All trials were conducted in quiet rooms, and participants listened using a pair of Beyerdynamic 770 over-the-ear headphones, which provided additional noise attenuation.

## RESULTS

No statistically significant differences were found between blocks A and B, so results were pooled for analysis. Data from Experiment 3 were separated into two groups for analysis: (a) tempo rankings of the stimuli presented at three distinct BPM sub-ranges (i.e., Core BPM groups), and (b) tempo rankings of stimuli

presented across the entire range of BPM levels used in the experiment.

For stimuli presented at the three core BPM levels, a 3 x 2 x 3 repeated-measures ANOVA (Core BPM x Complexity x Tempo Shift) found main effects for Core BPM,  $F(2, 54) = 236.29, p < .001, \eta_p^2 = .897$ , Complexity,  $F(1, 27) = 26.02, p < .001, \eta_p^2 = .491$ , and Shift,  $F(2, 54) = 114.55, p < .001, \eta_p^2 = .809$ . More complex patterns, which had more notes, as well as higher off-beatness, were given faster average tempo ratings across all BPM levels (at 105 BPM: 2.56 vs. 2.98; at 115 BPM: 4.14 vs. 4.68; at 125 BPM, 5.27 vs. 5.96). Figure 4a shows the results of Core BPM x Shift, combining both complexity levels. There was a small but statistically significant interaction between Core BPM and Shift,  $F(4, 108) = 2.58, p = .041, \eta_p^2 = .087$ , due to the *ns* difference between the -5 and 0 Shift levels at 105 BPM (note error bars in Figure 4a). Thus, aside from some compression at the slowest and fastest ends of the rating scale, participants rated stimulus tempos in accordance with their respective BPM rates.

The results for the stimuli presented across the entire BPM range are shown in Figure 4b, with the two complexity levels graphed separately. A 2 x 7 repeated-measures ANOVA (Complexity x BPM Level) found main effects for Complexity,  $F(1, 26) = 10.14, p = .004, \eta_p^2 = .281$ , and BPM Level,  $F(4.16, 108.21) = 183.77, p < .001, \eta_p^2 = .876$  (Greenhouse-Geisser correction applied); the interaction between Complexity and BPM Level was nonsignificant. Within each BPM level more complex patterns, which had more notes, as well as higher weighted off-beatness ratings, were given faster average tempo ratings, save at the two fastest BPM levels. More broadly, post hoc pairwise comparison *t*-tests for tempo ratings at across all BPM rates (Bonferroni correction applied) were all significant ( $p < .003$ ), save for 125 vs. 130 BPM, where the difference was nonsignificant. This is part and parcel of the ceiling effect we see regarding *ns* differences in tempo ratings for stimuli at 125 vs. 130 BPM in the previous two experiments.

## Discussion and Conclusion

Three experiments were conducted to study the TAE discovered by London et al. (2016). When presented with tempo-shifted versions of real music, listeners’ tempo judgments no longer correspond to the BPM rates of the stimuli, although they are able to correctly label faster versus slower versions of the same song using a Likert-type scale. Here we replicated the original experiment from London et al. (2016) using the same

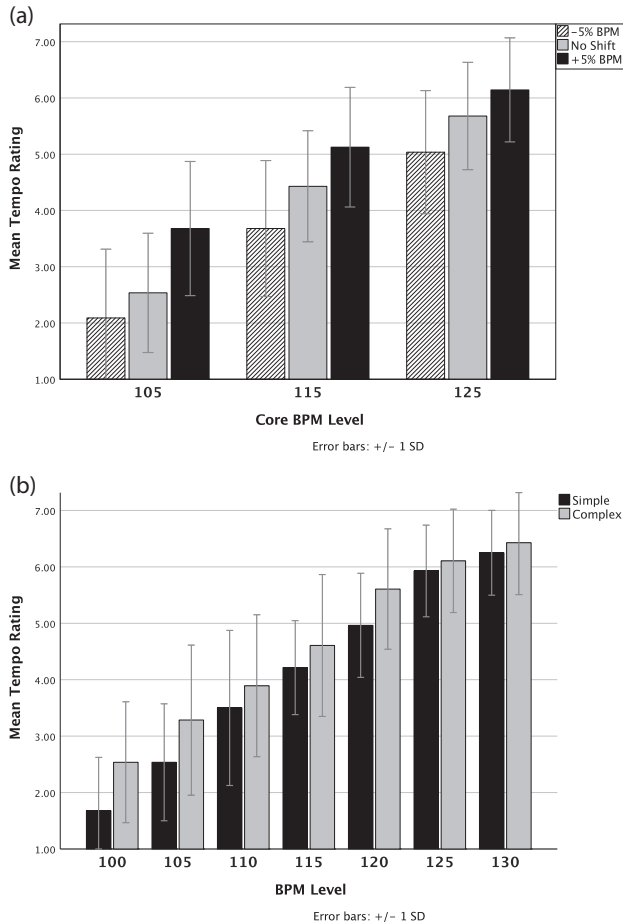


FIGURE 4. (a) Average tempo ratings of "time-shifted" percussive stimuli in Experiment 3, grouped by core BPM level. (b) Average tempo ratings of "all BPM" stimuli in Experiment 3.

Motown music stimuli (Experiment 1), and conducted two variant experiments, one with different music (Experiment 2) and one with purely percussive stimuli (Experiment 3). The TAE was replicated in Experiments 1 and 2, which used musical stimuli, but not in Experiment 3. Thus the TAE was not simply an artifact of the particular stimuli used in London et al. (2016), and seems systematically related to the musical characteristics of the stimuli involved.

Table 5 gives the Spearman's rho correlations between the grand average of all participant tempo ratings for each stimulus at a given core BPM x Shift level in all three experiments. As can be seen, the correlation between participant tempo ratings and the Absolute BPM rate increased from the Experiment 1 to Experiment 3, with the Drum stimuli exhibiting a nearly perfect correlation. Note also that for Experiments 1 and 2, their highest correlation was with each other, rather

TABLE 5. Spearman's Rho Correlations Between Grand Averaged BPM Ratings, Experiments 1, 2, and 3, and Absolute BPM Levels

	Exp 1 Motown	Exp 2 Disco	Exp 3 Drum	Absolute BPM
Exp 1 Motown				
Exp 2 Disco	0.983			
Exp 3 Drum	0.904	0.946		
Absolute BPM	0.870	0.941	0.996	

Note:  $p < .002$  in all cases; Absolute BPM adjusted for Experiment 1 (core BPM range 100–135)

than with either the ratings Experiment 3 or the absolute BPM rates of their stimuli. We applied Brandner's test (Brandner, 1933) to assess whether the correlations between the absolute BPM levels and participant tempo ratings (see bottom row of Table 5) displayed significant differences across the three experiments. Correlation was found to be significantly higher for Experiment 3 than for Experiment 2 ( $z = 4.28, p < .001$ , one-tailed), and significantly higher for Experiment 2 than for Experiment 1 ( $z = 1.75, p < .05$ ). This indicates that TAE was significantly stronger in Experiment 1 than in Experiment 2, and significantly stronger in Experiment 2 than in Experiment 3.

Our primary experimental hypotheses was thus confirmed, for as the salience of BPM rate versus other tempo cues increased, the TAE decreased, and was eliminated when other tempo cues (melodic, lyric) were removed. It is also worth noting that comparisons of the acoustic factors of Sub-Band Flux, Pulse Clarity, and Onsets/Second in the core BPM stimuli for all three experiments showed that while sub-band Flux increased across the three stimulus sets, the Disco stimuli had the highest average pulse clarity while the drum stimuli had the lowest number of events per second (see Table 6). Thus the increase in the salience of BPM as a cue for tempo rating depends not only on the presence of certain cues in the rhythmic/temporal domain (sub-band Flux, pulse clarity), but also upon the relative absence of cues in other musical parameters (melody, rhythm, harmony, and so forth), a decrease of what, for want of a better term, one might call "melodic salience."

In all three experiments participants were presented with a set of stimuli that were not tempo-shifted, spanning the range of BPM levels used in the experiment(s). Our secondary hypothesis was that tempo ratings for these stimuli would not exhibit the TAE. In Experiments 1 and 2a these were a simple rock drumming pattern presented across all tempo levels used in the experiment, in Experiment 2b there were a set of Disco songs, and in Experiment 3 they were drum patterns which

varied in complexity. In all three experiments participants' tempo ratings of these "all BPM" stimuli corresponded with their actual BPM rates. Our secondary hypothesis was thus also supported, as participants were able to make veridical tempo judgments of these various stimuli even though they were interleaved with the tempo-shifted musical stimuli. This also provides evidence that the TAE was not due to the limitations of the 7-point tempo scale used in the original and current experiments, as it was fine-grained enough to correctly rank all the different BPM levels of the "all BPM" stimuli in all three experiments.

Our findings regarding the TAE may inform a long-standing debate regarding the nature of absolute versus relative judgments of stimuli that vary along a single dimension. In their seminal article on the method of absolute judgment in psychophysics, Wever and Zener (1928) noted that any absolute judgment task presumes there is a relation between a particular stimulus and a larger series/set of stimuli against which it is judged (p. 469). They also note that our knowledge of that larger series is informed both by our broader life experiences with similar stimuli as well as our experience of the stimuli in a particular experimental context (pp. 472–473). This is what we have presumed when we speak of the tempo ratings in the three experiments reported on here as "absolute judgments." Our participants presumably had an internalized tempo scale based upon their lifelong encounters with music. Likewise, the pretest used in all three experiments clearly indicated the range of tempos used in the experiment (100–130 BPM), illustrated how the seven-point tempo rating scale mapped onto that range, and gave participants the opportunity to explore that range with simple drum stimuli that were unambiguous as to their tempo.

However, a number of researchers have challenged the very notion of absolute judgments. Lockhead has argued that seemingly absolute judgments are always made relative to context, memories, and channel capacity (e.g., Lockhead, 1992, 1995, 2004). Stewart, Brown, and Chater (2005) have made similar points, and have stressed the influence of sequential effects (i.e., trial-to-trial carryover effects) in making putatively absolute judgments. We were well aware of the problems of sequential effects, and that is why we randomized trials for each participant, constraining the randomization so that participants never heard the same song in successive trials. We also used "buffer" auditory material (recordings of birds, surf, etc.) between trials to minimize carryover effects as much as possible. Similarly, our use of a seven-point scale provided adequate channel capacity while allowing adequate discriminability

TABLE 6. Comparison of Stimulus Characteristics, Averaged Across all Core BPM Stimuli, Experiments 1, 2, and 3

	Exp 1 Motown	Exp 2 Disco	Exp 3 Drum
Sub-Band Flux	2.801	6.336	8.355
Pulse Clarity	0.443	0.721	0.607
Notes/Sec	4.992	5.105	3.031

Note: Sub-band Flux is in the 100–200 Hz range.

for the stimuli which were presented at seven absolute tempo/BPM levels.

Lockhead also raises a larger issue regarding absolute judgments, as he stresses the importance of making perceptual judgments relative to the invariant features of the perceptual array, as he argues that "object constancy is fundamental to perception and attribute scaling is not fundamental" (Lockhead, 2004, p. 267). This motivates his claim that people cannot abstract perceptual attributes (i.e., so-called secondary properties) from the objects that display them and the contexts in which those objects and their properties are perceived. We would respond by saying that music presents counter-evidence to this claim in interesting ways. First, music perception is not ecological, but, to use the term of Pierre Schaeffer, "acousmatic" (Schaeffer, 1966; see also Scruton, 1997). That is, the auditory objects of music are not sound sources per se, but patterns of sound abstracted from their sources. Thus, for example, when the sound of an oboe decreases in loudness, it does not signify that the oboist is moving away from the listener. Attributes such as loudness, tempo, and timbral brightness are expressive features of the musical sound that are modulated by the composer and performer for aesthetic effect. As such, musicians and music listeners have long referred to these attributes in terms of broadly shared absolute scales, and have developed terminologies (e.g., *largo*, *moderato*, *allegro*, *presto*) and technologies (e.g., early mensural notation and later use of metronome markings) to describe and communicate them. It is perfectly sensible to talk of several different pieces of music as having the same tempo or loudness—and indeed, this is precisely what our participants were able to do with the pieces presented at specific tempos in Experiment2b, as well as with the simple drum patterns presented across all tempos in Experiments 1 and 2a.

Yet our participants were not quite as able to relate tempo-shifted songs to the same absolute scale, giving rise to the TAE. We believe TAE is an example of *perceptual sharpening* (N.B. see also London, Thompson, Burger, Hildreth, & Toiviainen, 2019). Teufel, Dakin, and Fletcher (2018) reported that, for vision, "high-

level object representations interact with and sharpen early feature-detectors, optimizing their performance for the current perceptual context” (Teufel et al., 2018, p. 1). Kok, Jehee, and de Lange (2012) have discussed the effect of perceptual sharpening on visual expectation and perceptual sharpening, and Chennu et al. (2013) discuss analogous sharpening of auditory expectation and attention. The TAE, which arises in the particular context of tempo-shifted stimuli, is best understood as conflict between absolute and relative judgments—a conflict that only makes sense if both absolute and relative modes of judgment are available to the listener. In our experiments tempo judgments were framed as an absolute tempo-rating task within a 100–130 BPM “absolute series” (Wever & Zener, 1928, p. 471). When stimuli were presented singly (i.e., not time shifted), or across the entire range of the series, no conflict between absolute versus relative judgment occurred, and participants’ tempo ratings were strongly correlated with the BPM rates of the stimuli as indexed by the seven-point scale used in the experiments. The time-shifted stimuli, however, did present a conflict between absolute and relative judgments, as the robust memory for the tempo of these musically rich stimuli provided a basis for a relative judgment, but rather than more accurate responses relative to the scale, those responses were exaggerated—which is to say, sharpened. The TAE-as-a-form-of-perceptual-sharpening makes sense both in terms of our having robust, high-level object representations of music (as evidenced by the accuracy of our memories for pitch and tempo, e.g., Jakubowski, Farrugia, Halpern, Sankarpani, & Stewart, 2015; Levitin & Cook 1996), and in characterizing tempo as a low-level feature of our auditory perception. Drawing on the work of James and Stein (1961), Stewart, Brown, and Chater (2005) have shown that, relative to a central anchoring value (the grand mean of all of the observations), in a relative judgment task optimal criteria tend to be displaced outward toward the extremes. This relative shift of criteria away from the central value provides a mechanism for perceptual sharpening. In the case of the TAE, the internalized absolute scale provides the context for placement of the central value (the location of the grand mean), and when confronted with time-shifted stimuli, for which a relative tempo judgment can be made, the criterion shift occurs.

Conversely, the tempo-shifted percussive stimuli used in Experiment 3 did not give rise to the TAE, and this raises a question as to why. There are two possible explanations, and they are not mutually exclusive. One is that in the case of the percussive stimuli, the salience of the BPM rate is so high, given the absence of other

factors, that it becomes the dominant cue for tempo judgments. The other is that these patterns simply are not as musically rich and memorable as the Motown and Disco stimuli, and as a result, our participants were not able to make the high-level representations necessary for the TAE/perceptual sharpening to occur. This may also explain the decrease in the TAE from Experiment 1 to Experiment 2. First, as the Spotify data show (see Tables 1 and 2), the Motown stimuli in Experiment 1 are more widely known than the Disco stimuli of Experiment 2, and certainly both are more familiar than the artificial Drum stimuli of Experiment 3. Likewise, our participants had greater familiarity with the Motown stimuli in Experiment 1 than with the Disco stimuli in Experiment 2. Thus even if some of our participants had only an implicit knowledge of the songs and their musical styles in Experiments 1 and 2 (through their use in movies, commercials, and the like), that familiarity may have enhanced the formation of more detailed high-level representations of these stimuli, even if they were only formed within the context of the experiment itself.

Lastly, it may be that there is a “musicality” aspect to the TAE. Particular songs are associated with particular tempos, or optimal tempos (Boltz, 2017; Halpern, 1988; Jakubowski et al., 2015; Levitin & Cook, 1996), and this may have increased participants’ sensitivity to the altered tempos used in Experiments 1 and 2, though we note that the range of tempos used in the experiments was determined in large part upon the tempos that are characteristic of the R&B and Disco musical genres. To the extent that the stimuli used in Experiment 3 are musically “generic” (i.e., these are drumming patterns that are used across a range of popular musical styles and at a wide range of tempos), there is less sensitivity to their appearance(s) at different BPM rates.

Future studies of the TAE should investigate additional musical styles and different tempo ranges, to see how other “musical factors” (pitch contour and contour changes, pitch register, harmonic rhythm, etc.) may affect both beat salience versus melodic salience and memorability—for example, classical pieces may be less memorable for some participants than music in more familiar popular styles. Pieces of classical music are also not tied to particular recordings or versions, such that a particular performance (and its specific tempo) becomes canonical. Different experimental tasks may also help nuance the results obtained here. The use of a continuous measurement device (Schubert, 2001) would finesse the problem of fitting tempo ratings into a limited range, as well as the forced categorization of relative tempo judgments that may contribute to the

TAE. Likewise, the use of a standard-comparison tempo judgment task, especially for pairs of putatively equivalent stimuli (i.e., different melodies or rhythms at the same object BPM rate) would illuminate aspects of tempo equivalence. Finally, and perhaps most pressingly, the use of a stimulus familiarity test (i.e., can participants recognize if a stimulus has been previously used in the experiment) would provide data on the role of explicit memorability, especially for percussive stimuli, on the TAE.

As noted in the introduction, determinations of musical tempo involve far more than a simple assessment of the speed of the primary pulse or beat, as music presents a complex array of cues for speed and motion (Boltz, 2011; Drake et al., 1999; Elowsson & Friberg, 2013; London, 2011; Madison & Paulin, 2010). In addition to these cues, our prior musical experiences and memories—especially those of particular performances that are acquired through repeated listening to recordings—also provide a context for our judgments of musical tempo. The experiments presented here show that tempo judgments of rich musical stimuli can work in

ways that are starkly different from judgments of even relatively complex rhythmic patterns, and that the salience of beat rate/BPM can be affected by both stimulus structure and by experimental task. The TAE arises through a conflict between absolute versus relative frameworks for tempo judgments that involve the same music when heard at different tempos, but this conflict is likely to be present in most of our real-world musical experiences, informing our judgments of musical tempo to a greater or lesser degree.

### Author Note

Initial research was supported by a Finnish Core Fulbright Scholar grant to author JL, and by an Academy of Finland grant (project “Dynamics of Music Cognition,” project numbers 272250, 274037) to authors PT, BB, and MT.

Correspondence concerning this article should be addressed to Justin London, Department of Music, Carleton College, Northfield, MN 55057. E-mail: jlondon@carleton.edu

### References

- BERGESON, T. R., & TREHUB, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception*, 23, 345–360. DOI: 10.1525/mp.2006.23.4.345
- BOLTZ, M. G. (2011). Illusory tempo changes due to musical characteristics. *Music Perception*, 28, 367–386. DOI: 10.1525/mp.2011.28.4.367
- BOLTZ, M. G. (2017). Memory for vocal tempo and pitch. *Memory*, 25, 1309–1326.
- BRANDNER, F. A. (1933). A test of the significance of the difference of the correlation coefficients in normal bivariate samples. *Biometrika* 25(1/2), 102–109.
- BURGER, B., AHOKAS, R., KEIPI, A., & TOIVIAINEN, P. (2013). Relationships between spectral flux, perceived rhythmic strength, and the propensity to move. In R. Bresin & A. Askenfelt (Eds.), *Proceedings of the 10th Sound and Music Computing Conference*. (pp. 179–184). Stockholm, Sweden: Sosound and Music Computing Conference.
- CHENNU, S., NORIEKA, V., GUEORGUIEV, D., BLENKMANN, A., KOCHEN, S., IBÁÑEZ, A., ET AL. (2013). Expectation and attention in hierarchical auditory prediction. *The Journal of Neuroscience*, 33(27), 11194–11205.
- DRAKE, C., GROS, L., & PENEL, A. (1999). How fast is that music? The relation between physical and perceived temp. In S. W. Yi (Ed.), *Music, mind, and science* (pp. 190–203). Seoul, Korea: Seoul National University.
- EITAN, Z., & GRANOT R. Y. (2009). Primary versus secondary musical parameters and the classification of melodic motives. *Musicae Scientiae, Discussion Forum* 4B, 139–179.
- ELOWSSON, A., & FRIBERG, A. (2013). Modelling the perception of speed in music audio. In R. Bresin & A. Askenfelt (Eds.), *Proceedings of the Sound and Music Computing Conference* (pp. 735–741). Stockholm, Sweden: Sound and Music Computing Conference.
- FRAISSE, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, 35, 1–36.
- GRONDIN, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, and Psychophysics* 72(3), 561–582.
- HALPERN, A. R. (1988). Perceived and imagined tempos of familiar songs. *Music Perception* 6, 193–202.
- HONING, H. (2006). Evidence for tempo-specific timing in music using a web-based experimental setup. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 780–786.
- HONING, H. (2007). Is expressive timing relational invariant under tempo transformation? *Psychology of Music*, 35(2), 276–285.

- JAKUBOWSKI, K., FARRUGIA, N., HALPERN, A. R., SANKARPANDI, S. K., STEWART, L. (2015). The speed of our mental soundtracks: Tracking the tempo of involuntary musical imagery in everyday life. *Memory and Cognition*, 43(8), 1229–1242.
- JAMES, W., & STEIN, C. (1961). Estimation with quadratic loss. In J. Neyman (Ed.), *Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability* (Vol. 1, pp. 361–379). Berkeley, CA: University of California Press.
- KOK, P., JEHEE, J. F. M., & DE LANGE, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, 75, 265–270. <https://doi.org/10.1016/j.neuron.2012.04.034>
- LARGE, E. W., HERRERA, J. A., & VELASCO, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, 9, 159. DOI: <http://dx.doi.org/10.3389/fnsys.2015.00159>
- LARGE, E. W., & PALMER, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, 26, 1–37.
- LARTILLOT, O. & TOIVIAINEN, P. (2007). A Matlab Toolbox for musical feature extraction from audio. In S. Marchand (Organizer), *Proceedings of the 10th International Conference on Digital Audio Effects* (pp. 237–244). Bordeaux, France: DAFx.
- LEVITIN, D. J., & COOK, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception and Psychophysics*, 58(6), 927–935.
- LOCKHEAD, G. R. (1992). Psychophysical scaling: Judgments of attributes or objects? *Behavioral and Brain Sciences*, 15(3) 543–558.
- LOCKHEAD, G. R. (1995). Psychophysical scaling methods reveal and measure context effects. *Behavioral and Brain Sciences*, 18(3) 607–612.
- LOCKHEAD, G. R. (2004). Absolute judgments are relative: A reinterpretation of some psychophysical ideas. *Review of General Psychology*, 8(4), 265–272.
- LONDON, J. (2011). Tactus  $\neq$  tempo: Some dissociations between attentional focus, motor behavior, and tempo judgment. *Empirical Musicology Review*, 6(1), 43–55.
- LONDON, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Oxford, UK: Oxford University Press.
- LONDON, J., BURGER, B., THOMPSON, M., & TOIVIAINEN, P. (2016). Speed on the dance floor: Auditory and visual cues for musical tempo. *Acta Psychologica*, 164, 70–80. DOI: 10.1016/j.actpsy.2015.12.005
- LONDON, J., THOMPSON, M., BURGER, B., HILDRETH, M., & TOIVIAINEN, P. (2019). Tapping doesn't help: Self-motion and judgments of musical tempo. *Attention, Perception, and Psychophysics*, <https://doi.org/10.3758/s13414-019-01722-7>
- MADISON, G., & PAULIN, J. (2010). Ratings of speed in real music as a function of both original and manipulated beat tempo. *Journal of the Acoustical Society of America*, 128, 3032–3040.
- PARNCUTT, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11, 409–464.
- PHILLIPS-SILVER, J., TOIVIAINEN, P., GOSSELIN, N., PICHÉ, O., NOZARADAN, S., PALMER, C., & PERETZ, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia*, 49(5), 961–969. DOI: 10.1016/j.neuropsychologia.2011.02.002
- REPP, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review* 12(6), 969–992.
- SCHAEFFER, P. (1966). *Traité des objets musicaux*. Paris, France: Éditions Du Seuil.
- SCHUBERT, E. (2001). Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 393–414). New York: Oxford University Press.
- SCRUTON, R. (1997). *The aesthetics of music*. Oxford, UK: Oxford University Press.
- STEWART, N., BROWN, G. D. A., & CHATER, N. (2005). Absolute identification by relative judgment. *Psychological Review*, 112(4), 891–911.
- TEUFEL, C., DAKIN, S. C., & FLETCHER, P. C. (2018). Prior object-knowledge sharpens properties of early visual feature-detectors. *Nature Scientific Reports*, 8, 10853. <https://doi.org/10.1038/s41598-018-28845-5>
- TIERNEY, A., & KRAUS, N. (2015). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, 27(2), 400–408.
- TOUSSAINT, G. (2013). *The geometry of musical rhythm: What makes a 'good' rhythm good?* Boca Raton, FL: CRC Press.
- VAN NOORDEN, L., & MOELANTS, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, 28(1), 43–66.
- WEVER, E. G., & ZENER, K. E. (1928). The method of absolute judgment in psychophysics. *Psychological Review*, 35(6), 466–493. <http://dx.doi.org/10.1037/h0075311>
- WITEK, M. A. G., CLARKE, E. F., WALLENTIN, M., KRINGELBACH, M. L., & VUUST, P. (2014) Syncopation, body-movement and pleasure in groove music. *PLoS ONE* 9(4), e94446. <https://doi.org/10.1371/journal.pone.0094446>

## Appendix

### Tablature Representation of Stimuli used in Experiment 3

Stimuli consisted of hi-hat (HH), snare (S), and kick drum (K), as indicated on the left-hand side of each pattern. The hi-hat onsets are marked by “x”s, while snare and kick drum are marked by “o”s. Beat positions are indexed below each pattern.

#### Simple Patterns

JL\_01\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-----o-o-|---o-----o-o-||  
 K|o-----o-----|o-o-o-o-o-o-o-o-||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

JL\_02\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-o-o-o-o-o-|---o-o-o-o-o-o-||  
 K|o-----o-o-----|o-----o-o-----||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

witek\_s03\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-----o-o-|---o-----o-o-||  
 K|o-----o-----o-|o-----o-o-----||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

witek\_s10\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-----o-o-|---o-----o-o-||  
 K|o-----o-o-----|o-o-o-o-o-o-o-o-||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

#### Complex Patterns

JL\_03\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-o-o-o-o-o-|---o-o-o-o-o-o-||  
 K|o-o-----o-o-o-o-|o-o-o-o-o-o-o-o-||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

JL\_04\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-o-o-o-o-o-|---o-o-o-o-o-o-||  
 K|o-o-----o-o-o-o-|o-o-----o-o-o-o-||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

witek\_s18\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-----o-o-|o-o-o-o-o-o-o-o-||  
 K|o-----o-o-o-o-|---o-o-o-o-o-o-o-o-||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +

witek\_s28\_02  
 HH|x-x-x-x-x-x-x-x-|x-x-x-x-x-x-x-x-||  
 S|---o-----o-o-o-|---o-----o-o-o-||  
 K|oo-----o-o-o-o-|o-o-o-o-o-o-o-o-||  
 m 1 + 2 + 3 + 4 + 1 + 2 + 3 + 4 +