Probing imagined tempo for music: Effects of motor engagement and musical experience

Kelly Jakubowski[1], Nicolas Farrugia[1], and Lauren Stewart[1,2]

Abstract

Both musically trained and untrained adults can reproduce the tempo of familiar music with high precision. However, conflicting evidence exists as to how well representations of tempo are preserved within musical imagery. The present study investigated whether previous conflicting evidence might result from the use of different tasks to measure imagined tempo. Tempo judgments for familiar music were collected in a repeated-measures design using two imagined music tasks and one perceived music task. In one imagined music task participants tapped in time to the beat of the imagined music (Imagery (motor) task), while in the other they did not move in time with the music and instead adjusted a click track to the beat (Imagery (non-motor) task). Overall, performance was most accurate on the perceived music task, in which all musical cues were present. Performance on the Imagery (motor) task was also significantly more accurate than performance on the Imagery (non-motor) task. Training and active engagement with music positively predicted imagery task performance, whereas perceived music task performance was influenced by properties related to the song stimuli, such as familiarity and the original, recorded tempo. Results are discussed in relation to previous literature on auditory–motor interactions and musical expertise.

The precision with which musical tempo can be recalled is an area of great interest to musicians and researchers of musical memory. Professional musicians often exhibit remarkably stable timing profiles in performances of the same piece of music months or even years

[1]Goldsmiths, University of London, UK
[2]Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark

Corresponding author:
Kelly Jakubowski, Department of Psychology, Goldsmiths, University of London, New Cross Road, New Cross, SE14 6NW, London, UK.
Email: k.jakubowski@gold.ac.uk

apart (Clynes & Walker, 1982, 1986; Collier & Collier, 1994). Adults untrained in music can also reproduce the tempo of familiar tunes by singing both highly veridically, as compared to the original recording (Levitin & Cook, 1996), and highly consistently across time (Bailes & Barwick, 2011; Bergeson & Trehub, 2002). In tempo discrimination tasks for isochronous tone sequences, adult participants have displayed an average just noticeable difference (JND) of only 2.5%, and this JND was not significantly affected by musical training (Friberg & Sundberg, 1995).

How are such veridical and consistent tempo representations for music recalled? In order to successfully recall a melody, one must first represent the music within the "mind's ear." Musical imagery, the ability to "hear" music in one's head in the absence of a perceived external stimulus, has been an area of increasing recent research due to its crucial roles in action planning for music performance (Keller, 2012) and musical expectation generation (Janata, 2001). Musical imagery has often been studied in terms of its parallels to music perception (Hubbard, 2010; Zatorre & Halpern, 2005). For instance, many aspects of perceived music, including pitch, tempo, and timbre, are represented within musical imagery (Crowder, 1989; Halpern, 1988a, 1988b, 1989), and the areas of the brain recruited when people deliberately imagine music are highly similar to those recruited during music perception (Halpern, 2001; Zatorre, Halpern, Perry, Meyer, & Evans, 1996).

With regard to musical tempo in particular, several previous studies have affirmed that tempo information is represented within musical imagery (Halpern, 1988a; Weber & Brown, 1986). Halpern (1988b) asked participants to adjust the tempo of familiar songs that do not exist in canonical versions, such as "Happy Birthday", to the speed that sounded "right" to them in two conditions: (1) a perception condition, where they heard the music aloud; and (2) an imagery condition, in which they imagined a song and set a metronome to the beat of the imagined music. A high within-subjects correlation was found between chosen tempi in the perceived and imagined music conditions. In subsequent work, Halpern (1992) compared the tapped tempo for the same imagined song within the same participant across multiple trials from within the same testing session and between two sessions (2 to 5 days apart). Musically trained participants demonstrated high consistency in tapping speed both within and between sessions, whereas participants untrained in music showed high consistency within the same session but lower consistency between sessions.

While these studies provide evidence for the existence and general consistency of temporal aspects of musical imagery, some subsequent research suggests that representations of musical tempo within imagery may be less robust than other aspects of imagined music, such as pitch information. In a study by Janata and Paroo (2006), participants heard an ascending major scale and determined whether a final probe note was in tune (pitch task) or on time (time task). In an imagery condition, three or five notes of the scale were omitted and participants were asked to imagine these notes. Performance on the pitch task was highly similar in both the perception and imagery conditions, but performance in the timing domain deteriorated in the imagery as compared to the perception condition. A study by Weir, Williamson, and Müllensiefen (2015) used a similar imagery paradigm with familiar pop music as a stimulus. Performance on both the imagined pitch and timing tasks was above chance level, but participants performed more accurately in the pitch than the timing condition.

Thus, it appears there is some variation between findings within the literature as to how precisely tempo information is preserved within musical imagery. One potential source of this discrepancy could come from variations in the types of tasks used to measure imagined tempo. Janata (2012) conducted a comparison of three tasks for measuring imagined *pitch* representations and found that performance was more accurate in tasks that afforded a

greater degree of "sensory support," in terms of greater reinforcement from perceived stimuli. One might expect similar results in the imagined tempo domain, although no previous research has been conducted with this aim. Another factor that might increase discrepancies in performance on imagined tempo tasks is the influence of participant-level factors, such as musical expertise, or song-level factors, such as the range of song tempi used in the stimulus set. The present study aimed to clarify some of these unanswered questions by comparing multiple methods of probing imagined musical tempo, as well as by investigating individual differences between participants.

Evidence from several sources suggests that probing musical imagery using a task that engages the motor system might improve accuracy of tempo judgments over non-motor probing methods. Clynes and Walker (1982) asked professional musicians to perform, imagine, and conduct while imagining the same piece of music. They found that the duration of the conducted version of the piece was generally closer to the duration of a musician's actual performance than the duration of the purely imagined version. Manning and Schutz (2013) demonstrated that tapping during silent intervals improved performance on subsequent timing tasks (i.e., judging whether a final probe note is on time) relative to a task in which participants relied solely on auditory imagery. Correlational studies have shown a positive relationship between auditory imagery and sensorimotor synchronization abilities in musicians (Pecenka & Keller, 2009), and anecdotal reports from musical imagery experiments suggest that tapping along to imagined music supported more accurate imagery timing (Schaefer, 2014; Schaefer, Vleck, & Desain, 2011). The present study aimed to extend these findings by comparing accuracy of recall for familiar pop songs in two imagined musical tempo judgment tasks— one that engaged the motor system and one that did not—and one perceived music task within the same participants.

The two imagined tempo tasks employed in the present study were constructed to be highly similar to tasks used in previous literature. For the imagery task that did not require motor engagement with the imagery, a similar task was used to that of Halpern (1988b), who asked participants to set a metronome to the speed of an imagined song. The task in the present study used an isochronous click track, the speed of which could be changed in very small increments, such that participants' judgments were not constrained by the limited set of speeds on a typical metronome. For the imagery task that required motor engagement with the imagery, participants tapped along to the beat of the imagined music, as in Halpern (1992).

As mentioned above, individual differences may also play a key role in memory for musical tempo. In the present study, formal musical training and active engagement with music were assessed using the Goldsmiths Musical Sophistication Index (Müllensiefen, Gingras, Musil, & Stewart, 2014), as training in music has shown positive relationships to performance on some previous auditory imagery tasks (Aleman, Nieuwenstein, Böcker, & de Haan, 2000; Bishop, Bailes, & Dean, 2013, 2014; Halpern, 1992; Pecenka & Keller, 2009). Auditory imagery abilities are also known to naturally vary across the population (Aleman et al., 2000; Highben & Palmer, 2004). Therefore, it was predicted that participants who self-reported more vivid auditory imagery experiences in general or who were better able to control their auditory imagery would perform more accurately on the imagined tempo judgment tasks. These components of auditory imagery were assessed using the Bucknell Auditory Imagery Scale (Halpern, 2015). Finally, previous research indicates that music listeners tend to exhibit a "preferred" tempo for perceived music of around 100 to 120 beats per minute (bpm), such that metrical levels around this tempo are more salient to participants in listening experiments than other metrical interpretations (McKinney & Moelants, 2006; Van Noorden & Moelants, 1999). When asked to tap at a "comfortable rate" in tasks to measure spontaneous motor tempo, adult participants also

tend to produce a tapping rate around 100 to 120 bpm (McAuley, Jones, Holub, Johnston, & Miller, 2006; Moelants, 2002). In light of these findings that certain tempo ranges are more salient than others, the effect of the original, recorded tempo for each of the 12 songs used as stimuli in the present study on performance in the tempo tasks was also examined.

In summary, the main aims of the present study were to assess the accuracy of tempo judgments for familiar pop songs with two imagery tasks and one perceived music task, and to examine the impact of individual differences on performance on these tasks. It was hypothesized that an imagery task in which participants moved in time to the beat of the imagined music would result in more accurate tempo judgments than one that relied solely on auditory feedback, and that overall performance would be most accurate in a perceived music condition, where all musical cues were present. It was also hypothesized that performance on these tempo tasks would be influenced by participant-level factors such as musical training, musical engagement, and auditory imagery abilities, and song-level factors such as the original, recorded tempo and familiarity of each stimulus.

Method

Design

All participants completed three tasks in a repeated-measures design: two musical imagery tasks (hereafter referred to as the Imagery (motor) task and Imagery (non-motor) task) and one musical perception task (hereafter the Perceived Music task).

Participants

Participants were 25 university students (five male), ages 22 to 36 years ($M = 26.72$, $SD = 3.48$). Prior to the experiment, all participants were required to confirm that they were familiar with all 12 songs used as stimuli.

Materials

Musical stimuli. In order to select the pop songs that would be used as stimuli in the main study, a pilot study was conducted using 20 participants. None of these 20 volunteers participated in the main study, but were sampled from the same population as the main study participants (undergraduate and postgraduate students at Goldsmiths, University of London). A list of 145 pop songs was distributed, which was created in previous research to represent a wide range of chart-topping pop songs from within the past 50 years. Participants in the pilot study were asked to check a box next to each song that they were familiar with, such that they could recall a part of the song in their head. Twelve songs that were rated as being familiar to at least 75% of these participants were chosen as stimuli for the main experiment. These songs were also chosen to represent a wide range of recorded tempi from approximately 50 to 120 bpm (see Table 1).

For the purposes of the main experiment, a section from each of the pop songs of 16 beats in length was selected (see Appendix). In pilot work, these sections were deemed to be highly recognizable parts of the songs. The original tempo of each of these 16-beat musical excerpts was calculated from the recording of the song, using the sound editing software Audacity to locate the onset times of the first and last notes of the musical excerpt.

For the Perceived Music task, excerpts of the original recordings corresponding to these 16-beat sections were extracted using Audacity. Fade-ins and fade-outs of 0.25 seconds were

Table 1. Songs used as stimuli with original recorded tempi.

| Song	Artist	Recorded Tempo (bpm)
My Heart Will Go On	Celine Dion	49.8
Stairway to Heaven	Led Zeppelin	72.5
Hotel California	The Eagles	73.6
Let it Be	The Beatles	75.7
Imagine	John Lennon	76.0
Wonderwall	Oasis	86.6
Baby One More Time	Britney Spears	93.0
Last Christmas	Wham	104.8
Thriller	Michael Jackson	116.3
Billie Jean	Michael Jackson	117.1
Every Breath You Take	The Police	117.5
Like a Virgin	Madonna	120.2

applied to each excerpt. For the two Imagery tasks, participants did not hear any music aloud and were presented only the song title, artist, and lyrics of the excerpt.[1] The lyrics that coincided with the beat of the music were marked in bold and underlined to assist participants in finding the beat and to ensure all participants were tapping at the same metrical level. Beats that occurred between lyrics were marked as " ___ " (see Figure 1).

Tasks. To administer the three tasks in the experiment, applications were developed in Max/MSP. In the Imagery (non-motor) task, the participant adjusted the speed of a click track to the beat of an imagined song using a Griffin Powermate assignable USB controller. The Powermate is a circular dial that can be used to make tempo adjustments in small increments in real time. It has no external calibration marks, which helps to ensure independence of observations between trials. The tempo shift algorithm employed to change the tempo of the click track in the Imagery (non-motor) task, as well as the tempo of the music in the Perceived Music task, was ZTX (formerly DIRAC) time stretching technology.[2] ZTX is widely used in professional audio applications and the music industry to apply tempo changes without changing pitch or loudness.

In the Imagery (motor) task, participants tapped to the beat of an imagined song on the touchpad of a Mac laptop computer. Tap onset times were recorded by detecting a change in the status of the touchpad using the output of the fingerpinger object for Max/MSP.[3]

In the Perceived Music task, participants heard a section of a pop song aloud and adjusted the speed of the song in real time using the Griffin Powermate controller. Sennheiser HD 202 headphones were worn for both the Imagery (non-motor) and Perceived Music tasks.

Questionnaires. To examine the relationship between individual differences and task performance, the Musical Training and Active Engagement subscales of the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen et al., 2014) were administered to measure formal musical training (music lessons, instrumental practice, etc.) and engagement with music (concert attendance, listening habits, etc.) respectively. The Bucknell Auditory Imagery Scale (BAIS; Halpern, 2015) was used to assess self-reported auditory imagery abilities in terms of vividness of imagery (Vividness subscale) and control over the sounds within one's imagery (Control subscale). Participants also rated their familiarity with each of the 12 songs used as stimuli.
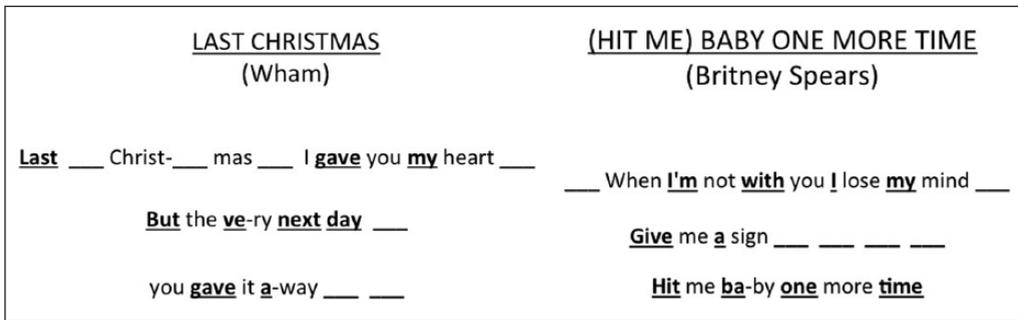
**Figure 1.** Beat marking technique for song lyrics.

Procedure

Each participant completed three blocks of tempo adjustment/production tasks. In all blocks, participants completed two practice trials of the task for that block to ensure they understood the procedure. Blocks 1 and 2 were Imagery tasks (motor and non-motor). The order of presentation of the two Imagery tasks was counterbalanced such that 13 participants completed the Imagery (motor) task in Block 1 and 12 completed the Imagery (non-motor) task in Block 1.

In the Imagery (non-motor) task, participants were presented with the title, artist, and lyrics to the 16-beat excerpt of each of the 12 pop songs on a computer screen (as in Figure 1). On each of the 12 trials, participants were asked to imagine the part of the song corresponding to the presented lyrics as closely as possible to the original recorded version of the song by the artist presented on the screen. Participants were then asked to set the speed of an isochronous click track to the beat of the song being imagined. The start speed of the click track was randomized and started at either 36 bpm or 180 bpm. These start speeds were chosen to lie well outside the range of the correct tempi for the pop songs so that the participants would make tempo adjustments for all or most trials. Participants were asked not to move to the imagined music during this task. Thus, although this task did require some small hand movements in order to adjust the click track to the beat of the music, participants were not moving in such a way that their body movements could become entrained to the beat of the imagined music. In contrast, the Imagery (motor) task required bodily entrainment to the beat of imagined music via tapping.

In the Imagery (motor) task, participants were presented with the same sections of the same 12 songs (as in Figure 1), in a different random order (all blocks were randomized for all participants). In this task, participants were asked to imagine the song as closely as possible to the original version by the artist presented on the screen and to tap all the way through the section of the song on the screen.

Block 3 was a Perceived Music task for all participants.[4] Participants heard the excerpt of music from each pop song aloud while being presented with the song title, artist, and lyrics. The tempo of the music started either quite fast or quite slow (60% or 150% of the original tempo), again to ensure participants made adjustments for all or most trials. The start tempi corresponded to the start tempi of the Imagery (non-motor) task, such that if a participant received a slow start speed for the click track (36 bpm) in the Imagery (non-motor) task for the song "Let it Be," he/she also received a slow start tempo for the actual music in the Perceived Music task for "Let it Be." The reason for this was that some previous work using a similar paradigm (Jakubowski, Halpern, Grierson, & Stewart, 2015) found a bias of
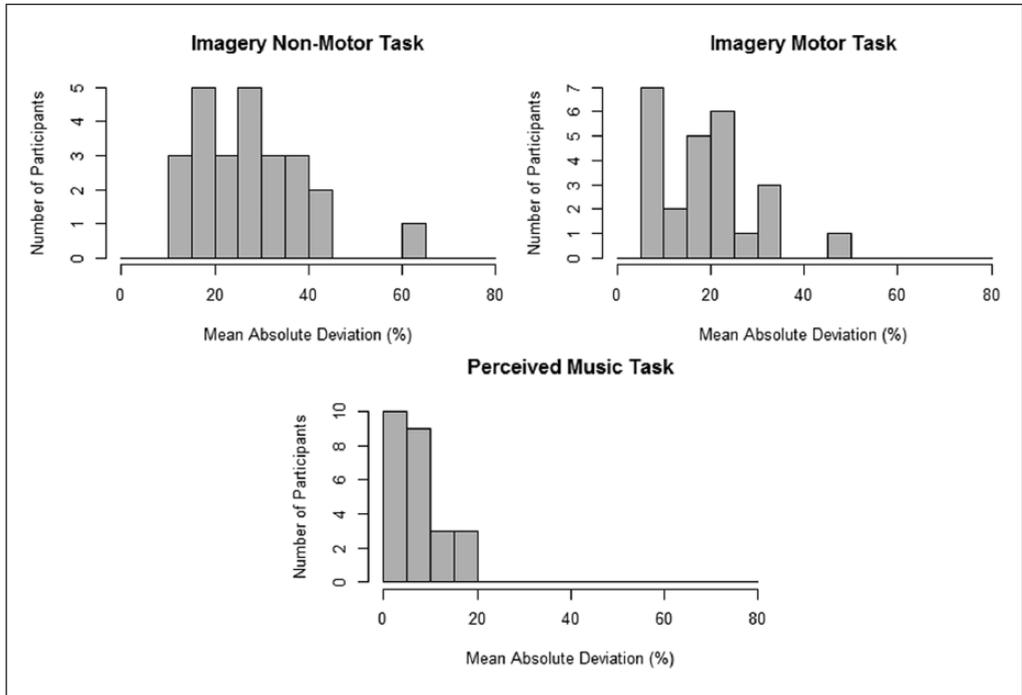
Figure 2. Mean deviations (as percentages) for the 25 participants for each task.

participants' tempo judgments toward the start tempo of the stimulus. As such, we aimed to keep the start tempo of the stimuli (click track or music) similar across the blocks to minimize any potential effects of such a bias.

Finally, each participant filled out the Gold-MSI and BAIS questionnaires and rated their familiarity with the songs used as stimuli on a scale of 1 to 5.

Results

Imagery (non-motor) task. The data for the Imagery (non-motor) task were filtered such that trials in which the ratio of a participant's chosen tempo to the original, recorded tempo was greater than 1.9 or less than 0.6 were discarded, following Halpern (1988b). This filtered out data that likely represented participants doubling or halving the tempo of a song when setting the speed of the imagined music. This resulted in discarding 20% of trials of the task.

For the remaining trials, the absolute deviation of a participant's chosen tempo from the original, recorded tempo was calculated as a percentage for each trial (0% deviation = participant chose the exact original tempo). The absolute deviations were averaged across all usable trials for each of the 25 participants (see Figure 2). The mean absolute deviation from the original tempo across all usable trials from all participants was 25.9% (SD = 17.3%). A ratio of the chosen to original tempo was also calculated for each trial (ratio of 1 = participant chose the exact original tempo). The mean ratio of the chosen to original tempo across all trials was 1.11 (SD = 0.25).

Imagery (Motor) Task. For the Imagery (motor) task, the tapping data were analyzed using a similar method to previous tapping research (Benoit et al., 2014). First, outliers (defined as

inter-tap intervals [ITIs] deviating by more than three times the interquartile range from the median ITI) and artefacts (ITIs of less than 100 ms) were removed from the tapping sequences. All trials with fewer than 8 total taps (2.3% of the data) were then discarded, as these trials provided too few taps to produce reliable tempo estimates.

For the remaining trials, the arithmetic mean ITI and mean coefficient of variation (CV; a normalized measure of tapping variability defined as the standard deviation of the ITI series divided by the mean ITI) were calculated for each trial. The ITI values were converted to bpm values for ease of comparison to the other tasks. Next, 28 trials (9.3% of the total data) were discarded on the basis of the ratio of the tapped tempo to the original, recorded tempo being greater than 1.9 or less than 0.6 (as in the Imagery (non-motor) task). Due to a technical problem in the data collection software, one trial of tapping for one participant was not recorded. All other trials were completed in all tasks.

As in the Imagery (non-motor) task, the absolute deviation from the original tempo and ratio of the tapped to original tempo were calculated for each trial. The mean absolute deviation from the original tempo across all usable trials was 18.4% (SD = 13.4%; see Figure 2). The mean ratio of the tapped to original tempo was 1.02 (SD = 0.20).

Perceived Music task. In the Perceived Music task, participants' chosen tempi deviated on average from the original song tempi by 8.1% (SD = 6.3%; see Figure 2). The mean ratio of chosen to original tempo across all trials was 0.98 (SD = 0.10).

Comparisons between tasks. Overall, performance was most accurate on the Perceived Music task, and least accurate on the Imagery (non-motor) task (see Table 2). The effect of task on performance (measured as deviation from the original tempo) was investigated using a linear mixed effects model. This mixed effects model allowed us to take into account the differences in variance across the three tasks. A significant difference in performance was found between all three tasks (Perceived Music vs. Imagery (motor): $t(777)$ = 8.63, $p < .001$; Perceived Music vs. Imagery (non-motor): $t(777)$ = 14.41, $p < .001$; Imagery (motor) vs. Imagery (non-motor): $t(777)$ = 5.86, $p < .001$; all comparisons Bonferroni corrected).

Next, correlations in performance between the three tasks were examined. Spearman correlations were calculated due to the presence of some outliers and positive skewness in the data. A significant positive correlation was found between performance on the two Imagery tasks, $\rho(23)$ = .58, $p = .003$, and a significant positive correlation was found between performance on the Perceived Music task and Imagery (non-motor) task, $\rho(23)$ = .40, $p = .048$. The correlation between the Perceived Music and Imagery (motor) tasks did not reach statistical significance, $\rho(23)$ = .32, $p = .12$.

Individual differences. Linear mixed effects models were fitted to investigate individual differences related to task performance. Separate models were constructed for each of the three tempo judgment tasks with musical training (GoldMSI-MT), active musical engagement (GoldMSI-AE), auditory imagery vividness (BAIS-V), auditory imagery control (BAIS-C), familiarity with each song (Familiarity), and the original, recorded tempo of each song (OrigTempo) as predictor variables. "Participant" was included as a random effect in the models to take account of the multiple trials recorded per participant on each task. The dependent variable of interest was the absolute deviation from the original tempo of a song.

After the full models were fitted with all predictor variables included, each of the three models was then fitted again with only the previously significant predictors included. This improved the fit of the model substantially in all cases over the full model (based on the Bayesian Information

Table 2. Mean absolute deviations for each task.

Task	Number of Usable Trials	Mean absolute deviation (SD)
Perceived Music	300	8.1% (6.3%)
Imagery (motor)	265	18.4% (13.4%)
Imagery (non-motor)	239	25.9% (17.3%)

Criterion (BIC); see Table 3). In the Imagery (non-motor) task, only active musical engagement was a significant predictor of task performance, such that higher engagement scores predicted more accurate task performance. In the Imagery (motor) task, musical training and a faster original, recorded tempo predicted more accurate performance. This effect of original song tempo on Imagery (motor) task performance was present despite the fact that the coefficient of variation of the participants' tapping was not significantly related to original song tempo in a linear mixed effects model, $t(239) = -1.15$, $p = .25$. This suggests participants were less accurate but not more variable in their tapped productions of slower tempi songs. Finally, songs that were more familiar to participants and had a faster original, recorded tempo were recalled more accurately in the Perceived Music task than less familiar and slower tempo songs.

Follow-up study. Finally, a follow-up study was conducted to investigate whether familiarity with other versions of the pop song stimuli may have affected the results of the main study. Although participants were clearly instructed during all tasks to imagine the version of the song by the artist whose name was presented on the computer screen (e.g., Britney Spears, Celine Dion), the possibility that participants were more familiar with a different version of the song and recalled this version instead cannot be entirely ruled out. As such, two samples of participants completed a follow-up study. The first was a subset of 16 of the 25 participants who had participated in the original study. The other was a new, independent sample of 33 participants who were selected to be from the same age demographic as the original study participants. The mean age of the follow-up study participants was 29.6 years (range = 22–42 years, *SD* = 4.6) and nine of these participants were male.

An online task was administered to both samples. On each trial of this task, participants were presented the title of one of the 12 pop songs and the lyrics to a section of that song (but were *not* presented with an artist name). The sections of the songs were the same sections that had been used in the original study. Participants were asked to imagine the section of the song playing in their heads and to advance to the next page of the task once they had the song in their heads. They were then presented a recording of the section of the song (the same recording used in the Perceived Music task of the original study) and were asked, "Was this the version of the song you thought of?" The response options to this question were "Yes," "No," and "I wasn't able to bring to mind any version of this song." If a participant answered "No" to the question, he/she was also encouraged to provide information on the version of the song that had come to mind. This task was completed for all 12 of the pop songs that had been used as stimuli in the original study. It should be noted that, as the follow-up study was conducted approximately 12 months after the original study, it is likely that the participants who had completed the original study did not remember exactly which versions of the pop songs had been used in the previous study.

In regard to the new sample of participants, a different version of the song than the version used in the main study was imagined on 4.5% of trials. In regard to the 16 participants who had completed the original study, a different version of one of the pop songs was imagined on

Table 3. Linear mixed effects models for performance on each tempo task (reduced models).

1. Imagery (non-motor) task.

	Estimate	SE	t-value	p-value
Intercept	0.546	0.097	5.62	< .0001
GoldMSI-AE	−0.007	0.002	−2.94	.007

Note: BIC of full model with all predictors = 2.27, BIC of reduced model = −58.89.

2. Imagery (motor) task.

	Estimate	SE	t-value	p-value
Intercept	0.466	0.052	8.95	< .0001
GoldMSI-MT	−0.004	0.001	−3.36	.003
OrigTempo	−0.002	0.0004	−4.46	< .0001

Note: BIC of full model with all predictors = −128.69, BIC of reduced model = −177.47.

3. Perceived Music task.

	Estimate	SE	t-value	p-value
Intercept	0.215	0.027	7.87	< .0001
Familiarity	−0.016	0.005	−3.63	.0003
OrigTempo	−0.0007	0.0002	−3.96	.0001

Note: BIC of full model with all predictors = −592.67, BIC of reduced model = −647.88.

Table 4. Mean absolute deviations for each task (subset of 24 participants).

Task	Number of usable trials	Mean absolute deviation (SD)
Perceived Music	288	8.2% (6.4%)
Imagery (motor)	253	18.8% (13.5%)
Imagery (non-motor)	227	26.4% (17.7%)

six trials (3.1% of trials). Of the six trials for which a different version of a song was imagined by participants from the original study, four of these were reported by the same participant. In light of this result, the main data for the study were re-analyzed while excluding this particular participant. This subset of data followed the same pattern of results as the original analysis (see Table 4). The differences in performance between the three tasks all remained significant (Perceived Music vs. Imagery (motor): $t(742) = 8.55$, $p < .001$; Perceived Music vs. Imagery (non-motor): $t(742) = 14.14$, $p < .001$; Imagery (motor) vs. Imagery (non-motor): $t(742) = 5.67$, $p < .001$; all comparisons Bonferroni corrected).

Discussion

The present study compared the accuracy of tempo judgments for familiar music across two imagined music tasks and one perceived music task. Overall, performance was most accurate in

the Perceived Music task. This result parallels findings reported by Janata (2012) in the imagined *pitch* domain that greater "sensory support" enhances performance in imagery tasks. The findings from the Perceived Music task are also comparable to those of Levitin and Cook (1996), who found that participants sang within 8% of the correct tempo on 72% of trials in a production task using familiar pop songs. In the present study, 65% of trials on the Perceived Music task were within 8% of the original tempo. The slightly lower percentage in the current study could be due to the use of self-selected music in Levitin and Cook's work, which might facilitate even more accurate recall by increasing familiarity.

Performance in the Imagery (motor) task was also significantly more accurate than performance in the Imagery (non-motor) task. These findings extend previous research suggesting a relationship between auditory imagery and sensorimotor synchronization (Clynes & Walker, 1982; Manning & Schutz, 2013; Pecenka & Keller, 2009) by demonstrating that the tempi of familiar melodies can be reproduced more closely to the correct (veridical) tempo during an imagery task that requires moving along to the beat of the music than one that does not. Recent research has highlighted the crucial role of close interactions between the auditory and motor systems in the production of music, as well as in purely perceiving a musical beat (Bangert et al., 2006; Grahn & Brett, 2007; Zatorre, Chen, & Penhune, 2007). The present study provides some evidence that the auditory component within these auditory–motor interactions might be able to be replaced by *imagined* sound, such that the motor output (tapping) interacts with auditory imagery to improve the accuracy of time-related judgments. Future research should investigate in detail how these findings fit within current models of auditory–motor coupling and sensorimotor feedback mechanisms (e.g., Warren, Wise, & Warren, 2005; Wolpert, Ghahramani, & Jordan, 1995; Zatorre et al., 2007).

Overall, performance on the Imagery (non-motor) task, in which participants adjusted the speed of a click track to the beat of imagined music, was the least accurate of the three tasks. Aside from the lack of motor entrainment, there are a handful of other factors that might help to explain this less precise performance. One difficulty that may have been experienced in this task is that the sound of the click track playing at the incorrect speed could have influenced or interfered with the tempo of participants' imagined music. This point is especially relevant if one considers that auditory perception and auditory imagery engage several of the same cognitive resources (e.g., Zatorre & Halpern, 2005). Another possible source of difficulty in completing this task stems from the fact that participants were required to make small hand movements to adjust the dial controlling the click track speed. As these movements were not made in time with the beat, they may have served as a source of motor interference with the beat of the music. Further research is needed to clarify the relatively poor performance on the Imagery (non-motor) task. However, as they stand, the present results suggest that the Imagery (motor) task may be a more suitable method for probing imagined musical tempo than the Imagery (non-motor) task in future studies.

The main findings of the study suggest a wide range of average accuracy between participants for each tempo judgment task. As such, there is much scope in exploring potential individual differences related to performance on each task. Performance on the Perceived Music task was significantly influenced by familiarity with the pop songs. The absence of an effect of musical training or engagement suggests that even everyday music listeners who are not formally trained or avid concert goers can make highly accurate judgments about the speed of music in a task where all the musical cues are present—as long as the music is familiar enough. These findings might also help to account for those of Levitin and Cook (1996) and Bergeson and Trehub (2002), in which non-musicians performed extremely accurately and consistently at reproducing

the tempo of songs that were indeed highly familiar (all songs in these studies were self-selected) in tasks in which participants received auditory feedback from their own singing.

Although those participants with less musical experience tended to perform as well as those with more experience when the music was heard aloud, it appears that experience does play a role in accuracy of tempo judgments in purely imagined music tasks. In the Imagery (motor) task, musical training was a significant positive predictor of task performance. One possible explanation for this, supported by several behavioral and neuroscientific studies, concerns potentially enhanced integration between the auditory and motor systems in trained musicians as compared to non-musicians (Bangert et al., 2006; Zatorre et al., 2007). This finding also serves to support previous research that has revealed an advantage of musical training in general in sensorimotor synchronization tasks, e.g., tapping to perceived music, (Chen, Penhune, & Zatorre; 2008; Repp, 1999) by extending this advantage to the imagined music domain. On the other hand, musical engagement (e.g., concert attendance, regular listening) was a significant positive predictor of task performance in the Imagery (non-motor) task. This finding suggests that, although musical experience appears to enhance performance on the Imagery (non-motor) task, this experience does not necessarily need to be in the form of structured, formal lessons.

A significant effect of the original, recorded tempo of the song stimulus was found on performance in both the Perceived Music and Imagery (motor) tasks, such that songs with a faster original tempo were produced at a more veridical tempo than slower songs. The presence of this effect in the Perceived Music task may be related to previous research suggesting a general preferred perceptual tempo in adults of around 100 to 120 bpm (McKinney & Moelants, 2006; Van Noorden & Moelants, 1999), while the results in the Imagery (motor) task may relate to similar findings that the most comfortable spontaneous tapping rates in adults also lie around 100 to 120 bpm (McAuley et al., 2006; Moelants, 2002), since the fastest original tempo for a stimulus used in the present study was approximately 120 bpm (see Table 1). Additionally, as tapping variability was not significantly related to original song tempo, this provides indirect evidence that *production* of imagined tempo (via tapping) was not hindered during recall of slower songs, but rather that participants experienced difficulty in accurately *recalling* the image at the correct tempo. Further studies should aim to distinguish these two components (recall and production) to provide a clearer understanding of the mechanisms underlying these two aspects of the Imagery (motor) task. Future work should also utilize a wider range of pop song stimuli, including songs with original tempi both above and below 120 bpm, to provide further support for these findings.

No significant effects were found of self-reported imagery abilities on performance on any of the tasks, as measured by the BAIS. This might be due to the fact that the BAIS measures quite general auditory imagery abilities for environmental, speech, and musical sounds. As such, ratings of the vividness or amount of control over one of these sounds might not be directly related to the highly specific task of making refined tempo judgments for imagined music. In future research it might be useful to include more specific self-report measures that assess solely musical imagery abilities—in particular temporal aspects of the experience—as well as objective tests of imagery abilities to complement self-report measures.

Finally, the follow-up study suggests that, although cover versions of the songs used in the main study might exist, the versions of the songs we used as stimuli were by far the most commonly recalled versions in a task in which *any* version of the song in question was allowed to be recalled. The finding that participants in the follow-up study imagined a different version of a song than the one we had used on less than 5% of trials, along with the fact that the main study participants were explicitly instructed to recall the version by the artist whose name was presented on the computer screen, provide evidence consistent with the argument that participants in the main study did call to mind the versions of the songs that we intended. It should

also be noted that it is possible for certain cases in which follow-up study participants imagined a different version of a song that a participant did actually know the version used in the main study, but that he/she was just able to recall a different version more readily when probed to think of the *first* version that came to mind. Thus, the already small percentage of trials in which participants imagined a different version of a song may actually be an overestimate of the degree to which the results of the main study were affected by participants imagining different versions of the stimuli.

In summary, the present study suggests that while accuracy in judging the tempo of familiar music is greater in a perception condition with all musical cues present than in imagined music conditions, engagement of the motor system, musical experience (both training and active engagement), and the original tempo of the song stimulus itself can significantly affect performance in imagined tempo tasks. These findings have key implications for researchers of musical memory and imagery, as well as for musical performers, who use imagery for tempo as a means of planning and adapting their performances in real time and rely on these images to be as accurate as possible.

Funding

Notes

1. Participants were also presented the song title, artist, and lyrics in the same format during the Perceived Music task to ensure that the three tasks were as similar in design as possible.
2. See http://www.zynaptiq.com/ztx/.
3. Available at: http://www.anyma.ch/2009/research/multitouch-external-for-maxmsp/.
4. The Perceived Music task was completed in Block 3 for all participants, as we aimed not to prime participants with the real, recorded music before either of the Imagery tasks.

References

Aleman, A., Nieuwenstein, M. R., Böcker, K. B. E., & de Hann, E. H. F. (2000). Music training and mental imagery ability. *Neuropsychologia*, *38*, 1664–1668.

Bailes, F., & Barwick, L. (2011). Absolute tempo in multiple performances of Aboriginal songs: Analyzing recordings of Djanba 12 and Djanba 14. *Music Perception*, *28*(5), 473–490.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., & Altenmuller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *Neuroimage*, *30*(3), 917–926.

Benoit, C. E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., & Kotz, S. A. (2014). Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Frontiers in Human Neuroscience*, *8*, 494.

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, *13*, 71–74.

Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical expertise and the ability to imagine loudness. *PLoS ONE*, *8*(2), e56052, doi:10.1371/journal.pone.0056052

Bishop, L., Bailes, F., & Dean, R. T. (2014). Performing musical dynamics: How crucial are musical imagery and auditory feedback for expert and novice musicians? *Music Perception*, *32*(1), 51–66.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory–motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239.

Clynes, M., & Walker, J. (1982). Neurobiologic functions of rhythm, time, and pulse in music. In M. Clynes (Ed.), *Music, mind, and brain: The neuropsychology of music* (pp. , 171–216). New York: Plenum.

Clynes, M., & Walker, J. (1986). Music as time's measure. *Music Perception*, *4*(1), 85–120.

Collier, G. L., & Collier, J. L. (1994). An exploration of the use of tempo in jazz. *Music Perception*, *11*(3), 219–242.

Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 472–478.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *The Journal of the Acoustical Society of America*, *98*(5), 2524–2531.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Halpern, A. R. (1988a). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology-Learning Memory and Cognition*, *14*(3), 434–443.

Halpern, A. R. (1988b). Perceived and imaged tempos of familiar songs. *Music Perception*, *6*, 193–202.

Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, *17*(5), 572–581.

Halpern, A. R. (1992). Musical aspects of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 1–27). Hillsdale, NJ: Erlbaum.

Halpern, A. R. (2001). Cerebral substrates of musical imagery. *Annals of the New York Academy of Sciences*, *930*, 179–192.

Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, & Brain*, *25*(1), 37–47.

Highben, Z., & Palmer, C. (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education*, *159*, 58–65.

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*, 302–329.

Jakubowski, K., Halpern, A. R., Grierson, M., & Stewart, L. (2015). The effect of exercise-induced arousal on preferred tempo for familiar melodies. *Psychonomic Bulletin & Review*, *22*(2), 559–565.

Janata, P. (2001). Neurophysiological mechanisms underlying auditory image formation in music. In R. I. Godøy & H. Jørgensen (Eds.), *Musical Imagery* (pp. 27–42). Lisse: Swets & Zeitlinger Publishers.

Janata, P. (2012). Acuity of mental representations of pitch. *Annals of the New York Academy of Sciences*, *1252*, 214–221.

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics*, *68*(5), 829–844.

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*, 206–213.

Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, *58*, 927–935.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review*, *20*(6), 1133–1139.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Live span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*, 348–367.

McKinney, M. F., & Moelants, D. (2006). Ambiguity in tempo perception: What draws listeners to different metrical levels? *Music Perception*, *24*(2), 155–166.

Moelants, D. (2002). Preferred tempo reconsidered. In *Proceedings of the 7th international conference on music perception and cognition* (pp. 580–583). Causal Productions, Adelaide.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PloS One*, *9*(2): e89642.

Pecenka, N., & Keller, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Neurosciences and Music III: Disorders and Plasticity*, *1169*, 282–286.

Repp, B. H. (1999). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior*, *31*(2), 145–164.

Schaefer, R. S. (2014). Images of time: Temporal aspects of auditory and movement imagination. *Frontiers in Psychology*, 5, 4.

Schaefer, R. S., Vlek, R. J., & Desain, P. (2011). Music perception and imagery in EEG: Alpha band effects of task and stimulus. *International Journal of Psychophysiology*, *82*(3), 254–259.

Van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*(1), 43–66.

Warren, J. E., Wise, R. J. S., & Warren, J. D. (2005). Sounds do-able: Auditory–motor transformations and the posterior temporal plane. *Trends in Neurosciences*, *28*, 636–643.

Weber, R. J., & Brown, S. (1986). Musical imagery. *Music Perception*, *3*, 411– 426.

Weir, G., Williamson, V. J., & Müllesiefen, D. (2015). Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind, & Brain*, *25*(1), 48–57.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*, 1880–1882.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558.

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and the auditory cortex. *Neuron*, *47*, 9–12.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, *8*(1), 29–46.

Appendix: Sections of the 12 pop songs used as stimuli in all three tasks

Song	Section used as stimulus
Baby One More Time	When I'm not with you I lose my mind, Give me a sign. Hit me baby one more time.
Billie Jean	Billie Jean is not my lover. She's just a girl who claims that I am the one.
Every Breath You Take	Every breath you take, every move you make.
Hotel California	Welcome to the Hotel California Such a lovely place, such a lovely place, such a lovely face.
Imagine	Imagine there's no heaven. It's easy if you try.
Last Christmas	Last Christmas I gave you my heart But the very next day you gave it away.
Let it Be	When I find myself in times of trouble Mother Mary comes to me, Speaking words of wisdom, let it be.
Like a Virgin	Like a virgin, touched for the very first time
My Heart Will Go On	Every night in my dreams, I see you, I feel you That is how I know you go on.
Stairway to Heaven	There's a lady who's sure all that glitters is gold, and she's buying a stairway to Heaven.
Thriller	Cause this is thriller, thriller night And no one's gonna save you from the beast about to strike.
Wonderwall	Today is gonna be the day that they're gonna throw it back to you. By now you should've somehow realised what you gotta do.