

SINGING IN THE BRAIN: INVESTIGATING THE COGNITIVE BASIS OF EARWORMS

CALLULA KILLINGLY, PHILIPPE LACHEREZ, &
RENATA MEUTER
*Queensland University of Technology, Brisbane,
Australia*

MUSIC THAT GETS “STUCK” IN THE HEAD IS COMMONLY conceptualized as an intrusive “thought”; however, we argue that this experience is better characterized as automatic mental singing without an accompanying sense of agency. In two experiments, a dual-task paradigm was employed, in which participants undertook a phonological task once while hearing music, and then again in silence following its presentation. We predicted that the music would be maintained in working memory, interfering with the task. Experiment 1 ($N = 30$) used songs predicted to be more or less catchy; half of the sample heard truncated versions. Performance was indeed poorer following catchier songs, particularly if the songs were unfinished. Moreover, the effect was stronger for songs rated higher in terms of the desire to sing along. Experiment 2 ($N = 50$) replicated the effect using songs with which the participants felt compelled to sing along. Additionally, results from a lexical decision task indicated that many participants’ keystrokes synchronized with the tempo of the song just heard. Together, these findings suggest that an earworm results from an unconscious desire to sing along to a familiar song.

Received: April 29, 2020, accepted March 1, 2021.

Key words: earworm, involuntary musical imagery, phonological loop, working memory, subvocal rehearsal

MUSICAL IMAGERY IS THE MENTAL EXPERIENCE of music in the absence of any external auditory stimuli (Bailes, 2007). While imagining music in this way is done voluntarily, sometimes conscious control over musical imagery is lost, as reflected in the phenomenon of having music “stuck” in the head. This experience is referred to as an *earworm*, a form of involuntary musical imagery (Beaman & Williams, 2010; Halpern & Bartlett, 2011; Williams, 2015), characterized by the repetition of a short fragment of music in the mind, occurring without the

conscious initiation or maintenance of the observer. Earworms are a common occurrence for many people; in one large survey study, 89.2% of participants reported getting a song stuck in their head on a weekly basis (Liikkanen, 2011).

Despite their ubiquity, little is known about how and why earworms occur. An earworm is a phenomenon that serves no apparent purpose to the observer; it is moreover an experience that is neither consciously solicited, nor easily eradicated. Paradoxically, people often perceive their earworms to be “intrusive” (Hyman et al., 2015), yet while they do not ascribe to themselves any ownership of this experience, it is clearly one that is initiated and maintained internally. In this sense, the earworm poses a challenge to people’s naïve views of their own volition in their cognitive processes. Hence a number of recent studies have attempted to derive a theoretical basis for involuntary musical imagery, by studying both the characteristics of earworms and those who tend to experience them (Beaman & Williams, 2010; Finkel, Jilka, Williamson, Stewart, & Müllensiefen, 2010; Halpern & Bartlett, 2011; Jakubowski, Finkel, Stewart, & Müllensiefen, 2017; Liikkanen, 2011; Müllensiefen et al., 2014).

There is debate as to whether any song can be experienced as an earworm or whether there are particular predisposing characteristics of “sticky” tunes (Finkel & Müllensiefen, 2012; Jakubowski et al., 2017; Williamson et al., 2012). In general, while there are certain songs renowned for their intrusive qualities, earworms tend to vary among individuals (Hyman et al., 2013). Moreover, despite the common belief that earworms are annoying or unwanted (Beaman & Williams, 2010; Cunningham, Downie, & Bainbridge, 2005), research indicates that most songs people report getting stuck in their head are those that they enjoy (Halpern & Bartlett, 2011; Hyman et al., 2013; Williamson, Liikkanen, Jakubowski, & Stewart, 2014) and to which they commonly listen (Halpern & Bartlett, 2011). Research finds that factors such as recent and repeated presentation of music are influential in the development of an earworm (Byron & Fowles, 2015; Floridou, Williamson, & Stewart, 2017; Halpern & Bartlett, 2011; Liikkanen, 2012). Together, these findings indicate that simple exposure to music can render a song more likely to get stuck in the head.

Similarly, it is possible that an earworm may arise simply due to the manner in which a song is heard. Some have theorized that a song is more likely to get stuck in the head if it is partially heard than if it is heard in its entirety (Hyman et al., 2013), potentially because people are compelled to engage in “imagined vocalization” in order to complete the song (McCullough-Campbell & Margulis, 2015). In general, people might be more likely to recall an unfinished task than one that is complete, a phenomenon known as the Zeigarnik effect: when interrupted partway through a task, one feels a sense of incompleteness, resulting in an increased desire to finish the task, and a greater prominence of the task in memory (Zeigarnik, 1938). Accordingly, an earworm may develop after the frustrated attempt to hear a song to its completion. In a series of in-class experiments, Hyman et al. (2013) found that if the final song presented continued to play in participants’ heads at the end of the class, it was reported as an earworm more frequently in a follow-up survey 24 hours later, interpreted as partial support for the Zeigarnik effect. However, an experimental study by McCullough-Campbell and Margulis (2015) found no difference between finished and incomplete songs in subsequent self-reported earworms. Hence the existence of the Zeigarnik effect for unfinished songs has received mixed findings in past research.

Other research suggests that there may be underlying musical features in a song that make it more likely to elicit an earworm. For instance, songs that people report as earworms tend to be those which are easier to sing. Typical earworm songs have been described as having “simplicistic” and “repetitious” melodies (Kellaris, 2001, 2003), features that would render a song more singable. In fact, these features (simple, unembellished “slogans”) are characteristic of songs that inspire singing along in general (Pawley & Müllensiefen, 2012).

One study found that the most important predictors of reported earworms were notes that were longer and closer in pitch, from which it was inferred that songs that evoke earworms are easier to sing (Finkel & Müllensiefen, 2012; Williamson & Müllensiefen, 2012). Using similar methodology, another recent study found that different features characterized the songs reported as earworms (Jakubowski et al., 2017). Specifically, in comparing self-reported earworms to songs matched on popularity and style, results indicated that songs that elicit earworms were faster in tempo and tended to have melodic contours characteristic of Western music. Despite the difference in predictive features produced by these two studies, Jakubowski et al. (2017) suggested that both findings might be indicative of earworm songs being easier to sing.

It is perhaps unsurprising that, in arising from more singable songs, earworms have been associated with various singing behaviours, such as self-rated singing ability and the extent to which people report singing along to music generally (Floridou, Williamson, & Müllensiefen, 2012; Flouridou, Williamson, Stewart, & Müllensiefen, 2015; McCullough-Campbell & Margulis, 2015; Müllensiefen et al., 2014). A common response to having a song stuck in the head is to sing it aloud (Müllensiefen et al., 2014), and a recent experimental study demonstrated that participants who sang or hummed along with vocal songs were more likely to experience them as earworms (McCullough-Campbell & Margulis, 2015).

Given these findings, we proposed that earworms are songs with which one feels compelled to sing along, and that this desire results in an automatic process of singing the song in one’s head. Rather than conceptualizing the earworm as a passive thought that intrudes on one’s consciousness, it may be that earworms are an activity in which one engages in response to an unconscious desire to sing along with certain songs.

An open question is whether *internal singing*—in the sense of singing a song to oneself silently—involves subvocal articulation. Subvocal articulation (or subvocalization) can be thought of as an inner monologue that engages motor mechanisms involved in speech production (Locke & Fehr, 1970) and is continually recruited in a wide variety of mental tasks (Perrone-Bertolotti, Rapin, Lachaux, Baciú, & Lævenbruck, 2014). In particular, subvocal articulation is known to be involved in short-term working memory (Baddeley, 1992; Baddeley & Hitch, 1974), and is employed for tasks involving the retention of information in sequential order, such as recalling a list of random words or digits.

According to the Baddeley and Hitch working memory model (1974), immediate memory for information requiring verbal coding is facilitated by a limited-capacity *phonological loop* that employs subvocal articulation to rehearse phonological information and prevent it from decaying. While the phonological loop was initially proposed to represent exclusively speech-based information (Baddeley & Hitch, 1974), there exists debate as to whether other auditory information is maintained via this cognitive resource (Jones & Macken, 1993; Salamé & Baddeley, 1989), with some research suggesting that tasks involving voluntary musical imagery may also recruit phonological working memory resources (Aleman & van’t Wout, 2004; Nees, Corrini, & Leong, & Harris, 2017; Smith, Wilson, & Reisberg, 1995), or may employ a common rehearsal

mechanism (Williamson, Baddeley, & Hitch, 2010). Additionally, some studies have shown that voluntary musical imagery tends to elicit patterns of neural activity similar to that which is observed during phonological tasks (Gaab, Gaser, Zaehl, Jancke, & Schlaug, 2003; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Koelsch et al., 2009) and activities requiring vocal-motor planning (Brown & Martinez, 2007).

Earworm episodes characteristically encompass both the melodic and lyric content of a song, and these songs are predominantly (although not always) well-known to the observer (Beaman & Williams, 2010; Halpern & Bartlett, 2011; Hyman et al., 2013, 2015; McNally-Gagnon, 2016). Generally, an episode involves just a short fragment of the song, most frequently the chorus (Beaman & Williams, 2010; Liikkanen, 2012; McCullough-Campbell & Margulis, 2015), which loops repeatedly in the mind (Bailes, 2007; Halpern & Bartlett, 2011; Hyman et al., 2013). Hence in the case of an earworm, we conjecture that a repetitive fragment of song can become “stuck” in the phonological loop, such that one feels compelled to continually rehearse the line in working memory. When the individual becomes consciously aware of this process, it is perceived as an earworm; yet the extent of this awareness may depend upon individual differences or the perceived intrusiveness of the song (Hyman et al., 2015).

If a song “stuck” in the head co-opts phonological working memory resources, one would expect this to be evidenced as a deficit in tasks that represent this cognitive resource. One of the main pieces of evidence in favor of the Baddeley and Hitch working memory model comes from the dual-task paradigm, in which it has been consistently shown that engaging in two concurrent phonological tasks will disrupt performance (Baddeley, 1992). Two recent studies assessed the extent to which concurrent subvocalization would reduce the self-reported experience of earworms. Hyman and colleagues (2013) found that earworms were experienced for a shorter time by participants who completed anagrams (a task thought to rely on verbal working memory), in comparison to Sudoku (a visuo-spatial task). This finding was taken to indicate that earworms rely on subvocal rehearsal via the phonological loop, as undertaking a verbal task appeared to suppress the earworm.

Further, Beaman and colleagues (2015) found that interference to articulatory motor programming (upon which subvocalization is thought to depend) resulted in reduced frequency of both voluntary and involuntary musical imagery. Specifically, a series of experiments demonstrated that after listening to a song, chewing gum reduced the extent of self-reported earworms and

self-reported thoughts related to the song. Finally, although not directly assessing the effects of subvocalization, Floridou et al. (2017) investigated the effects of concurrent cognitive load on earworms, finding that a sustained attention dot task involving a medium or high concurrent load tended to reduce the self-reported experience, and the researchers noted that these conditions entailed the use of phonological working memory, as participants were required to subvocally count the number of dots of a certain color (medium cognitive load), or count backwards in threes when they saw dots of a certain color (high cognitive load).

The present series of experiments used a dual-task paradigm to investigate the extent of interference caused by earworms in a concurrent phonological task. The prediction was that the music that gets stuck in people’s heads is that with which they are subconsciously motivated to sing along, and thus to which they automatically sing along subvocally. If earworms are caused (or maintained) by people singing to themselves, this should be observable by inferior performance on a concurrent phonological task, particularly if the automatic singing engaged subvocal articulatory processes.

Experiment 1

The first experiment tested the above prediction using well-known songs, which were predicted to be catchy or not catchy, on the basis of a pilot study. We also explored the possibility that unfinished songs are more likely to get stuck in people’s heads than finished ones, in which case we would expect inferior performance on the task following presentation of a truncated song.

METHOD

Participants

Thirty participants (20 female) were recruited in person and via email. Participants consisted of psychology students at the Queensland University of Technology, friends, and associates. Ages ranged from 20 to 49 years ($M = 27.1$, $SD = 7.38$).

Materials

Songs. To determine the songs to be used, a pilot study was conducted among 12 naïve participants. Each participant listened to a selection of 10–12 songs out of a possible 22. Songs were compiled on the basis of previous literature pertaining to earworms (Beaman & Williams, 2010; Finkel et al., 2010; Hyman et al., 2013; Williamson et al., 2012), and comprised those we believed were either easy or difficult to sing, as determined by musical dimensions such as large jumps in

pitch, vocal ornamentation, and vocal range. The selection of songs was representative of a variety of musical genres (e.g., pop, rock, alternative), a range of musical tempi, and included popular songs from numerous time points over the past century. All songs had lyrics (i.e., no instrumental music).

Pilot participants rated songs on a set of seven-point scales with regards to: 1) how easy the songs were to sing along with, 2) the extent to which they felt compelled to sing along with the song upon hearing it, 3) how well they knew the song, and 4) how likely it was that the song would get stuck in their head. On the basis of these ratings, 20 songs were chosen for use in the study, half of which were rated highest on each of the dimensions (denoted “catchy” songs), and the remaining 10 which were rated lowest on each of the dimensions (“non-catchy”). Further details of the pilot study can be found in Supplementary Materials (see mp.ucpress.edu). In the main experiment, all songs were played through CD Sonic AE120E speakers, at a volume ranging between 72–79 dB SPL.

Serial Recall Task. Participants undertook a number of serial recall trials in each testing block. In each trial, participants were presented with seven random digits appearing simultaneously on a computer screen. They were instructed to remember these digits in the order of presentation and not to repeat the numbers aloud. They were given 10 seconds to view the digits, after which an onscreen instruction appeared, asking them to type the digits in the presented order. The program used for this task was written in PsychoPy (v2.7.1; Peirce, 2007). Visual stimuli were presented on a Samsung Syncmaster S19B420 monitor with a screen resolution of 1440 x 900 pixels.

Experience of Songs During Study Questionnaire. As a corollary to the experimental data, and to confirm that participants did subjectively experience some earworms during the study, participants also completed a questionnaire in which they were provided with a list of songs heard throughout each session, and asked whether they had experienced any of these songs as earworms during or following testing (indicated in a “yes” or “no” response). Earworms were defined as the experience of having a piece of music “stuck” in the head, typically a short segment of a song or melody, not intentionally initiated by the individual. Participants were asked whether they knew each song and to provide a rating (on a 7-point scale) of how well they knew the song, the extent to which they liked it, and the extent to which they felt compelled to sing along with it.

Procedure

Research was conducted with ethical clearance from the University Human Research Ethics Committee of QUT.

Participants were provided with an information sheet detailing their involvement in the study, in which they were informed that the research concerned the effects of music on cognitive performance in general; consequently, they did not know that we were specifically investigating earworms. They provided written consent of their willingness to participate and were informed that they could withdraw at any point.

Participants were tested individually in two sessions conducted approximately one week apart. Each session comprised eleven short blocks of tests, in which participants undertook continuous trials of the serial recall task. The first block was completed in silence to establish a performance baseline. In the following ten blocks, participants completed the trials either while listening to music or in silence, alternating such that each music block had a silent block occurring immediately afterwards.

During one session participants were presented with five songs pre-classified as catchy and in another with five non-catchy songs. These songs were randomly selected from the 10 catchy and 10 less catchy songs determined in the pilot study. The order of conditions (catchy and non-catchy) and the order of songs were both randomized between participants. For half of the participants, songs were truncated in the final chorus during an unresolved cadence. In each block, participants completed as many trials of the task as they were able within the timeframe, averaging 12.51 per block. For the blocks in which music was played, the duration of the block was the length of the song. For silent blocks, the timeframe mimicked the averaged song length (3 minutes and 47 seconds). At the conclusion of the experiment (following the second session), participants completed a short questionnaire regarding their experience of the songs during the experiment.

RESULTS AND DISCUSSION

The method of eliciting earworms through recent exposure to a song was largely effective, with almost two-thirds (63.3%) of participants reporting that they got at least one song stuck in their head during or between sessions. Table 1 provides participants’ ratings of the songs they heard, and Table 2 provides mean ratings for the different conditions.

Song Ratings

To account for the repeated measures, linear mixed effects models were employed to investigate the relationship between responses to the songs, combining item-level (song) and participant-level (individual differences) information (Baayen, Davidson, & Bates,

TABLE 1. Descriptive Statistics for Ratings of Songs

Condition	Song	Familiarity		Enjoyment		Singalong	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Catchy songs	Happy (Pharell Williams)	6.47	1.01	5.47	1.70	4.94	2.14
	You're the Voice (John Farnham)	5.28	2.02	4.22	1.83	4.11	2.45
	I Knew You Were Trouble (Taylor Swift)	5.32	2.19	4.58	1.98	5.11	2.23
	Help (The Beatles)	4.23	2.24	4.77	1.48	3.46	2.30
	YMCA (Village People)	6.05	1.27	4.16	1.77	5.37	1.83
	I'm a Believer (The Monkees)	6.09	1.04	5.18	1.60	4.73	1.85
	Waterloo (ABBA)	5.44	1.42	4.72	1.87	4.56	2.18
	Paradise (Coldplay)	5.80	1.55	5.80	1.69	5.10	2.23
Non-catchy songs	Twist and Shout (The Beatles)	5.71	1.44	4.86	1.66	5.07	2.13
	Crazy Little Thing Called Love (Queen)	5.00	2.00	4.73	1.62	3.91	2.43
	Breakfast at Tiffany's (Deep Blue Something)	5.06	1.98	5.69	1.66	5.13	2.25
	Back in Black (AC/DC)	4.18	1.98	4.00	2.00	3.00	2.24
	Unchained Melody (Righteous Brothers)	5.00	1.78	4.62	1.45	3.92	2.14
	Wuthering Heights (Kate Bush)	2.50	1.78	2.75	1.82	2.25	1.82
	Roxanne (The Police)	5.62	2.06	5.62	1.76	4.85	2.27
	Sultans of Swing (Dire Straits)	3.00	2.22	4.08	2.11	2.42	2.07
	Chandelier (Sia)	5.68	1.60	5.58	1.74	4.89	1.88
	House of the Rising Sun (The Animals)	3.78	1.93	4.67	1.68	3.39	2.06
	Walk This Way (Aerosmith)	4.36	1.15	4.21	1.81	3.57	1.65
	Don't Dream it's Over (Crowded House)	5.13	1.67	5.13	1.63	5.06	1.81

Note. *M* = Mean, *SD* = Standard Deviation. Songs were rated on a scale of 1 to 7.

TABLE 2. Descriptive Statistics for Song Ratings in Each Condition

	Condition			
	Catchy		Non-catchy	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Familiarity rating (1 to 7)	5.56	1.73	4.50	2.03
Enjoyment rating (1 to 7)	4.78	1.76	4.70	1.91
Desire to sing along rating (1 - 7)	4.67	2.19	3.92	2.22
Earworm (% songs reported)	22.00	29.41	15.33	23.89

Note. *M* = Mean, *SD* = Standard Deviation.

¹Percentage of songs reported as an earworm (out of the 5 presented per condition), averaged across participants.

2008). As predicted, there were significant relationships among the factors (see Table 3). Importantly, the desire to sing along was positively related to whether the song evoked an earworm, providing essential support for the underpinning hypothesis of this study. There were positive relationships between whether a song was reported as an earworm and people's enjoyment of and familiarity with the song. Further, familiarity and enjoyment were strongly positively related to participants' desire to sing along.

Generally, there was a distinct bias toward reporting the final song heard (in the second session) as an earworm (39% of the time), $\chi^2(1) = 7.72, p = .005$. Participants may have been particularly likely to nominate

TABLE 3. Bivariate LME Effect Sizes (Conditional Pseudo R^2)¹ and Significance Tests among Song Ratings

	Desire to sing along	Familiarity	Enjoyment	Earworm
Desire to sing along	1	.63**	.49**	.36**
Familiarity		1	.38**	.28**
Enjoyment			1	.23**
Earworm				1

¹ Calculated as per Nakagawa and Schielzeth (2013)

Note. ** $p < .001$

the song that was still in their head, as it was the song to which they were most recently exposed.

Serial Recall Accuracy

Performance on the serial recall task was scored as the proportion of digits produced in the correct order in each trial (relative order scoring; see Drewnowski & Murdock, 1980). Linear mixed models were employed for main analyses, with accompanying conditional pseudo- R^2 effect sizes, calculated per Nakagawa and Schielzeth (2013).

We first examined whether serial recall performance differed depending on whether catchy or non-catchy music was presented, in a linear mixed model including the predictors of Song Condition (2 levels; catchy or

non-catchy) and Time (2 levels; during or following presentation). This analysis ($R^2 = .30$) yielded a significant main effect of Song Condition (indicating poorer performance for catchy songs), $F(1, 7428.1) = 117.26$, $p < .001$, and a significant main effect of Time (indicating poorer performance during music presentation), $F(1, 7428.3) = 198.42$, $p < .001$, but no interaction, $F(1, 7428.1) < 0.10$, $p = .965$. The absence of a significant interaction indicates that catchy songs produced significantly greater phonological interference compared to non-catchy, and this effect was evident both during and following their presentation.

We then investigated in another linear mixed model whether performance varied following exposure to the different song conditions in comparison to baseline (3 levels; baseline, silent blocks following catchy music, silent blocks following non-catchy music), and found there was a significant effect, $F(2, 4781) = 15.44$, $p < .001$, $R^2 = .33$. Simple contrasts comparing baseline to each song condition (with no alpha adjustment) indicated that performance in the silent blocks following catchy songs was significantly poorer than baseline, ($p = .006$), which provides evidence for the presence of the earworm in phonological working memory. However, performance in the silent blocks following the non-catchy songs did not significantly differ from baseline ($p = .132$).

Figure 1 displays mean performance before, during, and after presentation of catchy and non-catchy music. This figure shows that performance during and following catchy music clearly decreased in comparison to baseline. In contrast, while the non-catchy music also affected serial recall performance while it played, accuracy was similar to baseline following its presentation. Figure 2 shows the performance during and following the individual songs.

As evident in Figure 2, this effect of performance dropping below baseline was strongly exhibited for five of the catchy songs, while the remaining five were above baseline. Given the large number of songs and lack of *a priori* hypotheses with regard to individual songs, a complete pairwise analysis of the different songs was not possible; yet these results suggest that some songs may be more effective triggers than others.

A linear mixed model was employed to assess whether leaving a song unfinished would moderate the effect of song catchiness in the silent block after the music ceased. The predictors were Song Condition (2 levels; catchy or non-catchy) and Truncation (2 levels; truncated or played in full; between groups). The analysis ($R^2 = .36$) yielded a significant main effect of Song Condition (indicating catchier music produced more interference), $F(1, 3978.1) = 81.87$, $p < .001$, a nonsignificant effect for

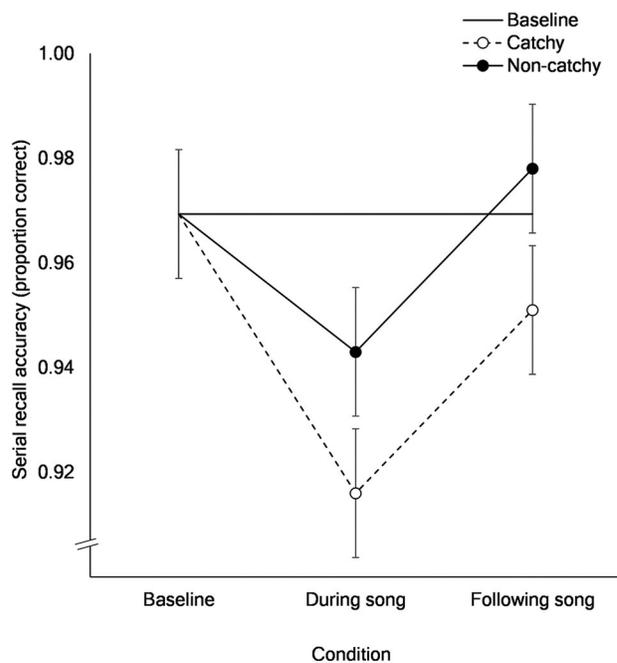


FIGURE 1. Serial recall performance (proportion of digits correct) before, during, and following presentation of catchy vs. non-catchy songs. Error bars represent ± 1 SE.

Truncation, $F(1, 28) = 0.30$, $p = .589$, and an interaction between them, $F(1, 3978.1) = 76.30$, $p < .001$. To follow up this interaction (Figure 3), we investigated the simple effect of Song Condition separately for truncated and complete songs. We found that performance was significantly poorer following presentation of catchy truncated songs compared to their non-catchy counterparts ($p < .001$). However, there was no effect of catchiness for the complete songs ($p = .825$).

We investigated whether participants' perceptions of the songs influenced their serial recall performance in the silent blocks following presentation, in a linear mixed model including their ratings of the songs as continuous predictors (familiarity, enjoyment, the desire to sing along). In this analysis ($R^2 = .32$), participants' self-reported desire to sing along with the songs, $F(1, 4005.5) = 31.25$, $p < .001$, was significantly negatively related to recall accuracy during silent conditions after song presentation, providing further support for the study's underpinning hypothesis. Participants' enjoyment of the songs, $F(1, 3997.1) = 20.58$, $p < .001$, was positively related to accuracy. Familiarity was not significantly related to serial recall in the silent blocks, $F(1, 4005.9) = 0.04$, $p = .852$.

In summary, the first experiment investigated the hypothesis that an earworm episode manifests as inner

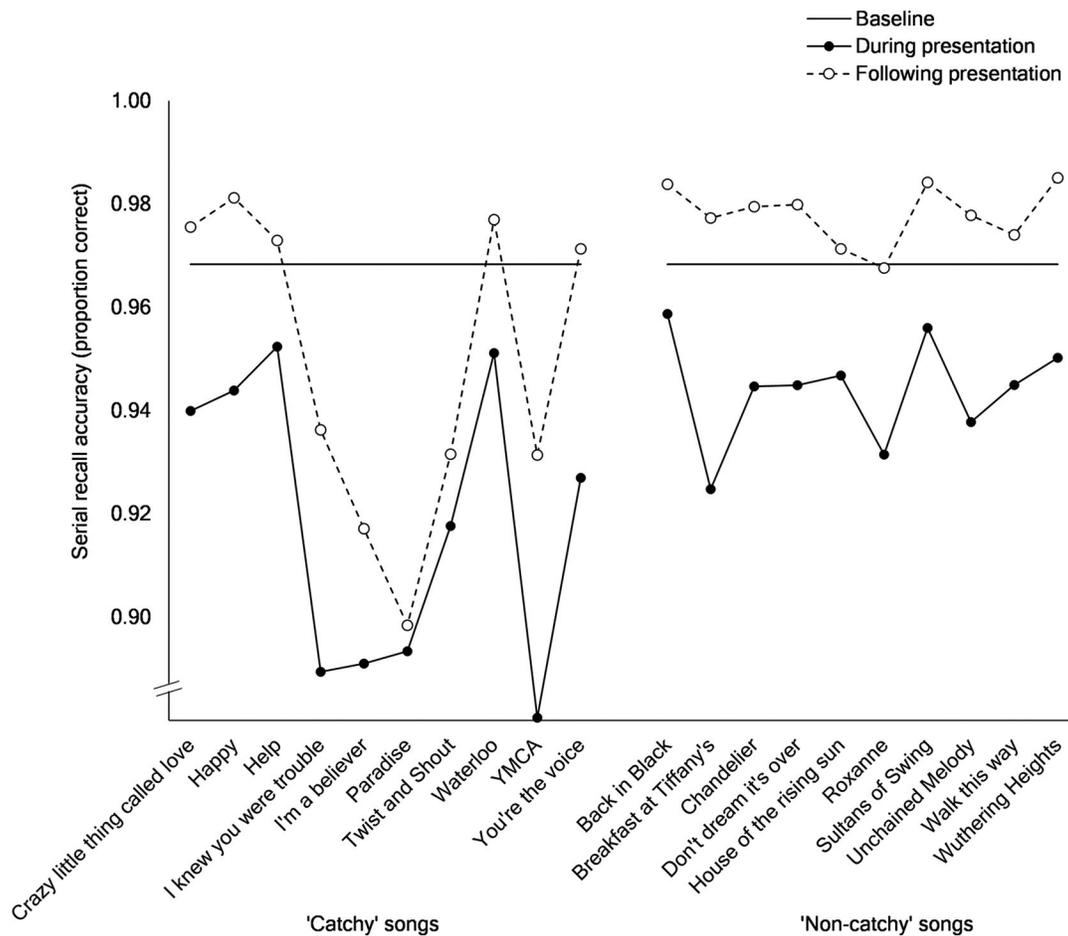


FIGURE 2. Serial recall performance (proportion of digits correct) before, during, and following presentation of songs. The ten songs classified as "catchy" are presented on the left side; the ten songs classified as "non-catchy" presented on the right.

singing, maintained by subvocal articulation. Specifically, we found that following exposure to catchy songs (as independently rated by pilot participants), interference was produced in a serial recall task, particularly for those songs that had not been played to completion. Further, we found that the extent of interference was influenced by the participants' self-rated enjoyment and the desire to sing along with particular songs. Hence this finding supports the prediction that phonological working memory is engaged when one has a song stuck in the head and established the effective use of a dual-task paradigm to objectively investigate people's earworms without risk of demand characteristics.

Experiment 2

The first experiment demonstrated that songs that are more catchy tend to remain active in phonological

working memory. Notably, this effect was more pronounced for some songs than others, and for catchy songs that were truncated. While this experiment showed that some songs might be generally more "sticky," it seems that individual preferences also contribute to the effect. Indeed, findings showed that the likelihood of a song being reported as an earworm was highly correlated with participants' ratings across the dimensions of enjoyment, familiarity, and the desire to sing along. To replicate and extend these findings, the second experiment explored the utility of a music selection method in which study participants pre-rated music on these characteristics, to identify songs most likely to evoke an earworm in a given participant.

This study adopted the same earworm induction paradigm developed in Experiment 1. As a corollary to the serial recall task, we also employed a lexical decision task as a secondary measure of phonological

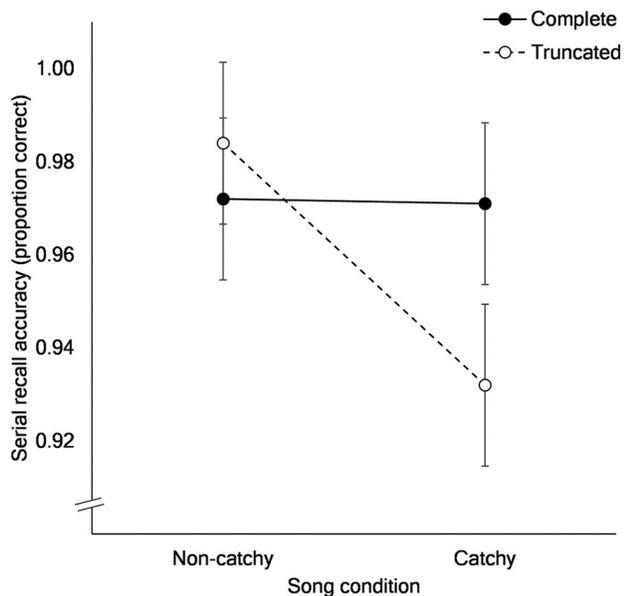


FIGURE 3. Song Condition x Truncation interaction for serial recall performance (proportion of digits correct) following presentation of music. Error bars represent $\pm 1 SE$.

interference. Since lexical decision making, like serial recall, relies on phonological working memory (Leininger, 2014; Lukatela, Frost, & Turvey, 1998), we hypothesized that interference caused by the presence of an earworm would slow people's response times in the lexical decision task. In addition, since lexical decisions can typically be made very quickly, we hypothesized that many more trials could be performed in the music and silent blocks, providing a more sensitive measure of interference.

In the present experiment, we predicted that phonological recoding during the lexical decision task would be impaired by concurrent subvocalisation of the songs, particularly for songs with which participants were compelled to sing along. Hence for both tasks, we hypothesized that performance would be impaired following presentation of songs rated higher on the desire to sing along.

METHOD

Participants

Participants were 50 undergraduate students (42 female) from Queensland University of Technology, recruited primarily from first year psychology students. Their ages ranged between 17 and 53 ($M = 21.58$, $SD = 8.68$).

Materials

Songs. Thirty potential catchy and non-catchy songs were compiled, similarly to Experiment 1. Specifically,

well-known vocal songs were chosen from a variety of artists and genres (e.g., pop and rock). The songs ranged in tempo, and were considered either easy or difficult to sing along with, based on melodic range and simplicity. Among these, we retained songs from Experiment 1 that had been repeatedly nominated as earworms or that had engendered a relatively high amount of phonological interference following presentation.

Serial Recall Task. The serial recall task was like that used in Experiment 1. However, digits were presented one at a time rather than simultaneously, with a presentation rate of one digit per second, and no interstimulus time interval. This was done to increase variance in individual performance accuracy, as performance was generally very high in the first experiment, as is typical for simultaneous presentation of stimuli with longer presentation times (Mackworth, 1962).

Lexical Decision Task. This task involved deciding whether a string of letters appearing on the computer screen represented a real English word or a nonword. In each trial, the stimulus would appear in the centre of the screen for 300 ms. Participants were asked to respond as quickly as possible with one of two keypresses (distinguished by colored stickers) to indicate if it was a word or nonword. The program advanced automatically to the next trial once a keypress was registered or if the participant did not respond within 300 ms (in which case the response was recorded as an error). Participants completed as many trials as they could within the timeframe, averaging 204 trials per block. Wordlists used for the blocks contained 224 stimuli, consisting of 109 nonwords and 115 English words (an equal proportion of high, medium, or low frequency); no stimuli were repeated for any participant, and stimuli were randomly presented in each block. All stimuli were obtained from the English Lexicon Project (Balota et al., 2007).

Procedure

Research was conducted with ethical clearance from the University Human Research Ethics Committee of QUT. In the first session, participants undertook a baseline on the serial recall and lexical decision tasks, and then listened to 20-second snippets of all 30 songs, presented in random order via the survey platform Qualtrics® (Qualtrics, Provo, UT). For each segment, participants rated their familiarity, enjoyment, and desire to sing along, on a scale of 0–100 using a slider. We identified the five songs rated highest in terms of desire to sing along by each participant, and the five with the lowest rating. Songs with which the participant was not familiar (a rating of below 15%) were excluded.

TABLE 4. Descriptive Statistics for Ratings of Songs

Song	Familiarity		Enjoyment		Desire to sing along	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Roxanne	74.89	27.86	57.67	34.82	44.33	41.18
You're the Voice	82.50	15.93	61.30	28.43	64.20	34.39
Let it Go	88.16	16.85	58.08	32.25	73.60	36.51
YMCA	91.20	13.63	67.90	28.11	76.30	36.09
Brown Eyed Girl	77.00	25.46	75.56	26.96	67.89	36.32
Imagine	74.11	26.25	51.61	32.96	40.72	40.56
Good Vibrations	75.29	20.70	32.29	27.66	23.88	23.79
Back in Black	67.25	19.11	51.58	26.33	31.79	27.57
500 Miles	87.30	17.04	74.39	27.24	81.61	27.75
All by Myself	84.22	20.25	59.30	25.35	58.61	33.98
Rolling in the Deep	90.21	17.93	78.00	22.01	83.79	19.52
Perfect	62.55	26.25	55.36	32.80	53.73	34.61
Breakfast at Tiffany's	59.56	21.21	38.89	29.24	29.11	30.06
Unchained Melody	58.53	23.41	60.80	28.68	38.87	35.10
Smells Like Teen Spirit	70.40	26.34	53.65	30.88	46.30	36.23
Walk this Way	69.20	18.76	43.00	18.78	26.60	21.66
Crazy Little Thing Called Love	67.50	23.99	59.83	22.67	47.67	28.76
Waterloo	87.64	19.61	69.36	32.17	63.86	37.43
I Knew You Were Trouble	94.14	8.66	69.82	33.31	72.45	34.30
Thinking Out Loud	90.25	14.36	84.04	19.71	79.00	23.91
Poker Face	83.89	17.78	52.22	30.95	53.89	35.23
I Will Always Love You	87.11	16.67	68.05	29.55	58.89	35.41
Still Haven't Found . . .	53.91	22.78	46.09	26.29	33.00	29.11
Are You Gonna Be My Girl	78.25	18.65	63.19	34.40	51.19	43.26
Africa	80.07	23.88	77.13	25.68	67.07	37.30
Beat It	92.52	12.66	78.95	21.97	72.10	32.92
MMMBop	70.71	26.59	51.64	22.84	36.14	29.60
Wonderwall	82.10	23.40	66.10	33.14	64.80	38.69
Don't Worry Be Happy	75.22	27.84	68.06	33.02	63.83	38.03
Uptown Girl	81.12	25.44	76.82	29.95	75.29	36.59

Note. *M* = Mean, *SD* = Standard Deviation. Songs were rated on a scale of 0 to 100.

Following this, participants undertook ten paired music-silent blocks that alternated between the serial recall and lexical decision task, such that the same task was conducted in the music block and silent block which immediately followed. The order of the two tasks was randomized among participants. The second session followed the same procedure, except without song ratings at the outset. Instead, like Experiment 1, participants were asked if they had any songs stuck in their head at the end of the second session. In one session, the highly rated songs were used, and for the other session, the lowest rated songs were used; the order of sessions was randomized.

RESULTS AND DISCUSSION

Once again, the method of earworm induction was highly successful, with all but one participant reporting at least one song as an earworm during the experiment. As with the first experiment, there was a tendency to report the most recently heard song as an earworm

(70% of the time), $\chi^2(1) = 5.09, p = .024$. Table 4 presents the average ratings of songs heard during the experiment and indicates that the songs were generally well-known. Table 5 presents mean ratings in each condition.

Relationships among the desire to sing along, familiarity, enjoyment, and self-reported earworms are displayed in Table 6. Notably, the effect sizes (conditional pseudo- R^2) are similar to that observed in Experiment 1, with moderate positive relationships among ratings of the desire to sing along, familiarity, and enjoyment of a song.

Serial Recall Accuracy

Data comprised participants' serial recall response accuracy, for each individual trial undertaken. As with Experiment 1, we calculated response accuracy as the average number of digits reproduced in the correct order (analysed with linear mixed models). The measure of baseline accuracy used for analysis was the

TABLE 5. Descriptive Statistics for Song Ratings in Each Condition

	Desire to sing along			
	High		Low	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Familiarity rating (/100)	93.73	9.99	65.53	22.50
Enjoyment rating (/100)	85.07	16.51	40.44	23.66
Desire to sing along rating (/100)	90.20	13.96	25.84	21.34
Earworm (% songs reported) ¹	83.20	25.35	42.40	28.47

Note. *M* = Mean, *SD* = Standard Deviation.

¹Percentage of songs reported as an earworm (out of the 5 presented per condition), averaged across participants.

TABLE 6. Bivariate Linear Mixed Effects Sizes¹ and Significance Tests Among Song Ratings

	Desire to sing	Familiarity	Enjoyment	Earworm
Desire to sing	1	.56**	.75**	.29**
Familiarity		1	.51**	.26**
Enjoyment			1	.28**
Earworm				1

¹Conditional pseudo R^2 effect sizes calculated as per Nakagawa and Schielzeth (2013)

Note. ** $p < .001$

participants' baseline performance from the commencement of the second session, as it was felt that the first baseline session might be artificially lower due to relative inexperience with the task. Baseline accuracy was 88.1% ($SD = 14.45\%$). In general, performance on the task was lower in comparison to Experiment 1, as expected with a sequential rather than simultaneous presentation of digits.

A linear mixed model analysis was conducted using the predictors of Song Condition (2 levels; high vs. low rated songs) and Time (2 levels; during and following presentation). Ratings of Enjoyment and Familiarity were included as covariates. We found a significant effect of Song Condition, such that performance was lower for the songs rated higher in terms of desire to sing along, $F(1, 7113.0) = 9.29, p = .002$, model $R^2 = .27$. There was also a significant effect of Time, indicating that performance was lower during presentation compared to following, $F(1, 7128.6) = 567.06, p < .001$; yet the interaction was nonsignificant, $F(1, 7128.4) = 1.76, p = .184$. Enjoyment was significantly related to performance, $F(1, 7005.4) = 7.15, p = .008$, while Familiarity was not related, $F(1, 7069.1) = 0.02, p = .892$.

Figure 4 displays the estimated marginal means, showing first that during song presentation, performance was

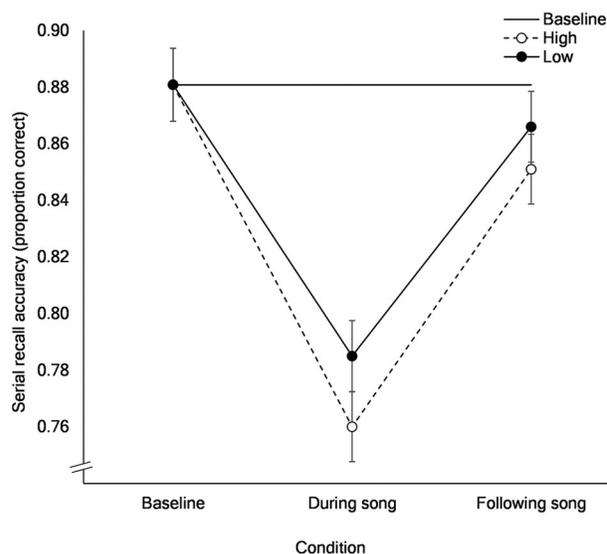


FIGURE 4. Serial recall performance (proportion of digits correct) before, during, and following presentation, split according to how much people generally wanted to sing along. Error bars represent ± 1 SE.

worse for the songs that were rated higher in terms of the desire to sing along. Following presentation, a similar pattern is observed, and performance clearly drops below baseline, with a more pronounced effect following those songs which elicit the compulsion to sing along.

We again investigated whether baseline performance differed from performance in the silent blocks following music presentation (3 levels; baseline, blocks following songs rated high, blocks following songs rated low). There was an overall significant effect, $F(2, 4348) = 6.73, p = .001, R^2 = .27$. Simple contrasts comparing baseline to each song condition (no alpha adjustment) indicated that there was a significant decrease in accuracy between baseline and silent conditions following music presentation, for songs that had been rated higher in terms of desire to sing along ($p < .001$), providing evidence that recently heard music was occupying working memory resources. For songs rated low in terms of desire to sing along, there was also a difference between baseline and performance following song presentation ($p = .002$), albeit to a lesser degree.

Lexical Decision Task

Data consisted of a reaction time and accuracy score (dichotomous) for each trial in each block. Reaction times that were over 3 or below -3 standard deviations away from the participant's average performance were considered outliers and excluded.

TABLE 7. Lexical Decision Response Accuracy (Proportion of Trials Correct) and Reaction Time (in Milliseconds) on Accurate Trials

		Condition (desire to sing along)			
		High		Low	
		M	SD	M	SD
Accuracy	Baseline	.897	.089	.897	.089
	During song presentation	.885	.060	.895	.057
	Following song presentation	.887	.057	.890	.071
Reaction time	Baseline	714	239	714	239
	During song presentation	698	.237	693	223
	Following song presentation	696	.229	691	217

Linear mixed model analyses were conducted for both response accuracy and reaction time, again for the factors Song Condition and Time, and including covariates of Enjoyment and Familiarity. While accuracy decreased slightly during and following songs rated higher on the desire to sing along (see Table 7), the difference between songs in these categories was non-significant, $F(1, 437.3) = 1.50, p = .222$. Time was also non-significant, $F(1, 427.0) < 0.10, p = .874$, and there was no interaction, $F(1, 427.0) = 0.84, p = .360$. However, significant variance was explained by both Familiarity, $F(1, 440.4) = 14.04, p < .001$, and Enjoyment, $F(1, 443.5) = 11.43, p = .001$. There was no difference between baseline performance and blocks following the high- or low-rated songs, $F(2, 545.8) = 1.78, p = .170$.

Additionally, response times on accurate trials were slightly longer for the songs rated high on desirability to sing along, yet this effect of Song Condition was not significant, $F(1, 85938.3) < 0.10, p = .897$, nor was there an effect of Time, $F(1, 85914.8) = 0.17, p = .679$, or an interaction, $F(1, 85914.4) < 0.10, p = .908$. There were no significant effects of Enjoyment, $F(1, 85799.5) = 0.24, p = .624$, or Familiarity, $F(1, 85891.9) < 0.10, p = .890$.

However, although we did not observe the effect of Song Condition in these results, visual inspection of the data showed that reaction times varied depending on the actual song being presented, and there was a strong carryover effect into the subsequent silent block (see Figure 5). A linear mixed model, including Song and Time as predictors, indicated a main effect of Song, $F(29, 99435.1) = 21.80, p < .001$, model $R^2 = .19$, a non-significant effect of Time, $F(1, 99398.2) = 1.06, p = .302$, and an interaction, $F(29, 99398.2) = 6.76, p < .001$. To

follow up this interaction, we examined the simple effect of Song at each level of Time, which showed that these reaction times significantly differed both while the song played, $F(29, 99438.3) = 18.20, p < .001$, and afterwards, $F(29, 99438.1) = 11.51, p < .001$.

When examining individual cases, it was evident that many participants' average reaction time in each block varied depending on the tempo of the song concurrently presented, and this shift in tempo continued following presentation of the song in the silent block. These data seemed to indicate that some people were responding synchronously to the beat of the music, a phenomenon observed by the experimenter during some participants' sessions.

To further investigate this response pattern, participants were classified *post hoc* into two groups: those who exhibited a tempo-dependent response pattern (22 participants), and those whose reactions were relatively stable across the different blocks (28 participants). Those in the former group were included if their response times were negatively correlated with song tempi, such that faster songs (with a higher tempo) were associated with quicker responses (lower reaction times). For those in the latter group, no such relationship was observed. Song tempo, defined in beats per minute (BPM), was split into four quartiles – the first representing songs with slower tempi 66–91 BPM, the second with tempi 93–117 BPM, the third with tempi 118–136 BPM, and the fourth with the fastest tempi 138–154 BPM.

For the first subset of participants, song tempo had a significant effect on response times, both during, $F(3, 20972.3) = 23.31, p < .001, R^2 = .14$, and following presentation of songs, $F(3, 21675) = 5.20, p = .001, R^2 = .14$. As seen in Figure 6, reaction times appeared to be faster on average during and following blocks where songs were faster in tempo, and were slower for songs that were slower. These data therefore indicate that some people tended to synchronise their response time to the beat of the music, and this continued even after the song finished, once more suggesting that the song persisted in the mind of the person. For the remainder of participants, tempo was not significantly associated with response times, during, $F(3, 27971.1) = 0.18, p = .911$, or following performance, $F(3, 28845.9) = 0.53, p = .663$.

These findings support the hypothesis that earworms occupy working memory resources, and that this effect is more pronounced for songs that motivate one to sing along. Further support for this hypothesis stems from the finding that song tempo impacted reaction time in the lexical decision task for a number of participants, as

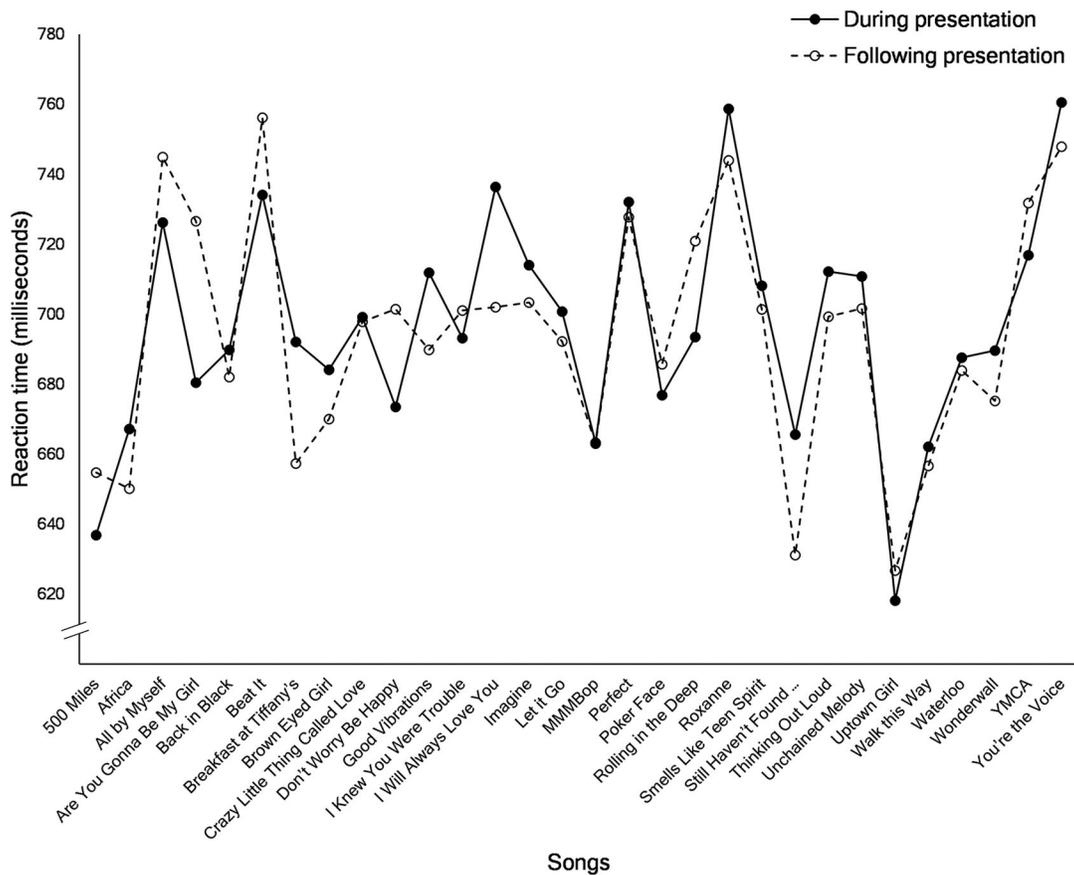


FIGURE 5. Average reaction times during (black markers) and following (white markers) songs.

they appeared to respond in time with the beat of the song, even after it stopped playing. This particular result strongly suggests a continued presence of the song in the mind of those participants, corroborating the hypothesis that a recently heard song will continue to be rehearsed in the mind as an earworm.

General Discussion

The aim of the present study was to investigate the hypothesis that earworms, rather than being an intrusive percept or imaginary sound, are instead a manifestation of an active process, in which a person subvocally sings the earworm to themselves, although they may not be aware that they are doing this. This prediction was supported on a number of levels. Experiment 1 demonstrated that catchy songs (predicted to elicit earworms) generate significantly more phonological interference compared to non-catchy songs, both while the song plays, and then after presentation. Hence the first experiment supports the prediction that phonological

working memory is engaged when one has a song stuck in the head, suggesting that one is singing along with the song mentally. Moreover, this effect was greater for those catchy songs which had been partially played, providing support for a Zeigarnik effect.

Experiment 2 replicated and extended these findings, by demonstrating the earworm effect for songs self-rated by the individuals in terms of their desire to sing along. In addition, results from a lexical decision task showed that many participants continued to respond in time with the song after it finished, again providing support for the notion that people continue to subvocally rehearse songs after they finish playing. Therefore, it seems there are certain songs with which one is more likely to mentally sing along (while the song plays), and that this process may occur automatically by means of subvocal articulation. Further, these songs tend to remain in people's heads after they have been heard as evidenced by interference in a phonological task.

Together, these findings provide support for the involvement of the phonological loop in the experience

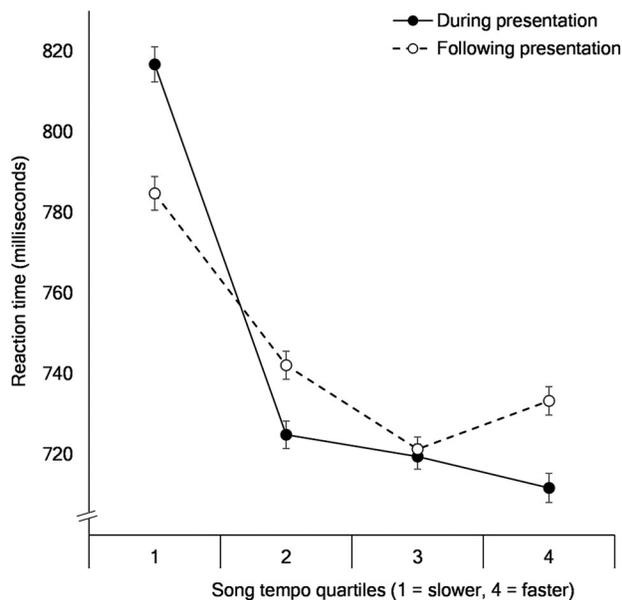


FIGURE 6. Relationship between response time and music tempo for the subset of participants ($n = 20$) who exhibited a tempo-dependent response pattern. Data are shown both during (black markers) and following (white markers) song presentation. Error bars represent ± 1 SE.

of earworms, a connection that has been suggested in two previous studies (Beaman et al., 2015; Hyman et al., 2013). Moreover, this has important implications for how the earworm is conceptualized as a cognitive phenomenon. While some researchers have considered earworms as a kind of intrusive thought (Hyman et al., 2015), the present findings indicate that earworms manifest in automatic inner singing. This may explain why individual differences in everyday sing-along behavior are related to the experience of earworms (McCullough-Campbell & Margulis, 2015; Müllensiefen et al., 2014; Williamson et al., 2014).

Findings from both experiments indicated that the desire to sing along with a song is highly related to the earworm effect produced. Specifically, individuals' ratings of how much they wished to sing along were able to predict the extent of interference caused, during and following presentation. Additionally, some songs in the present study had high mean ratings for the desire to sing along yet were not always rated highly in terms of enjoyment; hence, it may be that the songs which motivate mental singing along are not necessarily liked by the individual. Beaman (2018) proposed that while a positive response to a song can elicit internal rehearsal of the music, a negative response might equally result in an earworm, as the attempts to suppress the song would

only serve to reinstate it as an earworm (Wegner, 1994). It is also worth noting that in the second experiment, almost every song was represented in both experimental conditions, indicating that a song that compels one person to sing along might yet exert the opposite effect on another individual. Hence the extent of a song's stickiness cannot be wholly ascribed to its inherent features; instead, earworms seem to arise partly as a result of one's response to the song, and this may explain why the earworm tends to be such an idiosyncratic phenomenon.

A corollary of the hypothesis stems from the unexpected finding that some participants continued to respond in line with a song's tempo following its presentation. This result suggests that people may become unconsciously entrained to the music in their head, as manifested in their motor responses. Although this response was not anticipated, it somewhat parallels recent research investigating a "pulse continuity phenomenon," whereby people become entrained to the beat of a song and continue to tap along to the beat after the music fades out (Kopiez, Platz, Muller, & Wolf, 2013).

Additionally, this finding adds to a growing body of literature indicating a relationship between musical imagery and sensorimotor entrainment (Jakubowski, Farrugia, & Stewart, 2016; Manning & Schutz, 2013; Pecenka & Keller, 2009). For instance, research shows that motor engagement (tapping along) during an episode of deliberately imagined music was significantly more effective than a condition without motor engagement (adjusting a click-track) in improving temporal accuracy (Jakubowski et al., 2016). Research also finds there to be a high degree of accuracy in how individuals retain and reproduce temporal characteristics of certain songs, whether these are deliberately recalled or during an episode of involuntary musical imagery (Jakubowski, Bashir, Farrugia, & Stewart, 2018). Findings from the present study indicate a need to further investigate the extent to which involuntary musical imagery can influence motor and cognitive processes, and suggest that this could potentially be achieved by using similar reaction time tasks in the dual-task paradigm.

Experiment 1 provided support for a Zeigarnik effect in increasing the likelihood of earworms occurring, as hearing a song truncated partway through resulted in significantly inferior performance in the subsequent silent condition. Notably, this occurred only for the "catchy" songs. Hence this finding indicates that those catchy songs with which one is compelled to sing along are even more likely to stay "stuck" in the head if they are not heard to completion, presumably due to a sense

of “unfinished business,” and the compulsion to finish the song mentally. In previous research, this effect has gained some indirect support (Hyman et al., 2013), also with well-known music; however McCullough-Campbell and Margulis (2015) found no effect of song truncation. It is interesting to note that the researchers had presumed this effect of increased earworm frequency would occur because people would engage in “imagined vocalization” after the song had been truncated (McCullough-Campbell & Margulis, 2015). The present study indicates that the important factor in continued mental singing of a song is not only whether the song finished or not, but whether the individual is motivated to sing along, albeit unconsciously.

In the present research, participants completed the phonological tasks both during and following presentation of the music, and performance was particularly disrupted by concurrent music presentation. Previous research has documented the well-known distracting effects of background music on serial recall performance (Alley & Greene, 2008; Pring & Walker, 1994; Salamé & Baddeley, 1989), particularly for songs that are more enjoyable to the individual (Perham & Sykora, 2012). The present findings extend this body of research, demonstrating that songs that compel the individual to sing along can exacerbate this distracting effect, potentially as the observer is compelled to sing along while the music plays, as well as following the song’s presentation.

The present study highlights the strength of employing an experimental method, such as the dual-task paradigm, to examine the earworm phenomenon in a direct and objective fashion. In each experiment, it was necessary that participants did not complete questionnaires related to earworms until the end of the experiment (following the second session), as deliberately informing them of the research aims may have engendered demand characteristics; for instance, participants may have experienced earworms simply by thinking about the experience of earworms (Beaman & Williams, 2013). The focus of our study was to establish a measure of earworms, and this was predicated on the notion that earworms can be (at least according to previous self-reports) phenomena that occur to people automatically and without intentional effort on their part; that is, involuntarily. It is possible that in some instances, participants were deliberately singing along to the music, either during or after the presentation of the music. While we feel it is unlikely this would have been the main mechanism involved in the phonological

interference observed, our present findings cannot rule out this possibility, and it would be beneficial for future research employing this paradigm to probe participants about their experiences of both involuntary and voluntary instances of musical imagery following the experiment.

An important aspect of this paradigm is that participants always undertook the baseline serial recall block at the outset of the experimental session, prior to being presented with any music. While this was necessary to the design of the experiments, it does raise the question of whether performance accuracy in subsequent blocks was influenced by fatigue effects. However, we observe that this conjecture is not reflected in the data. As is evident in Figure 2, which displays performance before, during, and following individual songs, performance in the silent condition hovers around the baseline value, only going below that value for certain songs. We therefore consider it unlikely that these overall effects are attributable to task fatigue.

The present study uniquely contributes to the body of literature regarding earworms, in demonstrating that the dual-task paradigm is an effective tool with which to study this phenomenon. Most importantly, the present findings indicate that earworms are facilitated by an automatic process of subvocal articulation, such that one is actually singing the song in their head, although they may not be consciously aware that they are actively maintaining the melody in this way. Hence these findings may have implications for the way in which “intrusive” mental phenomena are conceptualized and suggest that familiar music that inspires people to sing along is likely to become lodged in working memory. Future research could use this methodology to study earworms for instrumental music, and ascertain whether these are maintained in the same way as vocal music earworms. Additionally, the paradigm could be employed to investigate more systematically the musical features which are implicated in this phenomenon, which could be of potential benefit to those in advertising and marketing.

Author Note

Correspondence concerning this article should be addressed to Callula Killingly, Queensland University of Technology (QUT), Faculty of Creative Industries, Education and Social Justice, School of Early Childhood and Inclusive Education, Level 4, E Block, Kelvin Grove Campus, Brisbane, Queensland, 4059, Australia. E-mail: c.killingly@qut.edu.au

References

- ALEMAN, A., & VAN'T WOUT, M. (2004). Subvocalization in auditory-verbal imagery: Just a form of motor imagery? *Cognitive Processing*, 5, 228–231. <https://doi.org/10.1007/s10339-004-0034-y>
- ALLEY, T. R., & GREENE, M. E. (2008). The relative and perceived impact of irrelevant speech, vocal music and non-vocal music on working memory. *Current Psychology*, 27(4), 277–289. <https://doi.org/10.1007/s12144-008-9040-z>
- BAAYEN, R. H., DAVIDSON, D. J., & BATES, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- BADDELEY, A. D. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- BADDELEY, A. D., & HITCH, G. (1974). Working memory. In H. B. Gordon (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- BAILES, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music*, 35(4), 555–570. <https://doi.org/10.1177/0305735607077834>
- BALOTA, D. A., YAP, M. J., CORTESE, M. J., HUTCHISON, K. A., KESSLER, B., LOFTIS, B., ET AL. (2007). The English lexicon project. *Behavior Research Methods*, 39(3), 445–459.
- BEAMAN, C. P. (2018). The literary and recent scientific history of the earworm: A review and theoretical framework. *Auditory Perception and Cognition*, 1(1–2), 42–65. <https://doi.org/10.1080/25742442.2018.1533735>
- BEAMAN, C. P., POWELL, K., & RAPLEY, E. (2015). Want to block earworms from conscious awareness? B(u)y gum! *Quarterly Journal of Experimental Psychology*, 68(6), 1049–1057. <https://doi.org/10.1080/17470218.2015.1034142>
- BEAMAN, C. P., & WILLIAMS, T. I. (2010). Earworms ('stuck song syndrome'): Towards a natural history of intrusive thoughts. *British Journal of Psychology*, 101, 637–653. <https://doi.org/10.1348/000712609X479636>
- BEAMAN, C. P., & WILLIAMS, T. I. (2013). Individual differences in mental control predict involuntary musical imagery. *Musicae Scientiae*, 17(4), 398–409. <https://doi.org/10.1177/1029864913492530>
- BROWN, S., & MARTINEZ, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain and Cognition*, 63(1), 59–69. <https://doi.org/10.1016/j.bandc.2006.08.006>
- BYRON, T. P., & FOWLES, L. C. (2015). Repetition and recency increases involuntary musical imagery of previously unfamiliar songs. *Psychology of Music*, 43(3), 375–389. <https://doi.org/10.1177/0305735613511506>
- CUNNINGHAM, S. J., DOWNIE, J. S., & BAINBRIDGE, D. (2005). "The pain, the pain": Modelling music information behavior and the songs we hate. *6th International Conference on Music Information Retrieval* (pp. 474–477). London, UK: ICMIR.
- DREWNOWSKI, A., & MURDOCK, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory*, 6(3), 319–332. <https://doi.org/10.1037/0278-7393.6.3.319>
- FINKEL, S., JILKA, S. R., WILLIAMSON, V. J., STEWART, L., & MÜLLENSIEFEN, D. (2010). Involuntary musical imagery: Investigating musical features that predict earworms. *Third International Conference of Students of Systematic Musicology (SysMus10)*. Cambridge, UK: SysMus10.
- FINKEL, S., & MÜLLENSIEFEN, D. (2012). Involuntary musical imagery and musical structure – Do we get earworms only for certain tunes? In E. Cambouropoulos, C. Tsougras, K. Mavromatis, & K. Pasiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition (ICMPC) and 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (p. 301). Thessaloniki, Greece: ESCOM.
- FLORIDOU, G. A., WILLIAMSON, V. J., & MÜLLENSIEFEN, D. (2012). Contracting earworms: The roles of personality and musicality. In E. Cambouropoulos, C. Tsougras, K. Mavromatis, & K. Pasiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 302–310). Thessaloniki, Greece: ICMPC/ESCOM.
- FLORIDOU, G. A., WILLIAMSON, V. J., & STEWART, L. (2017). A novel indirect method for capturing involuntary musical imagery under varying cognitive load. *Quarterly Journal of Experimental Psychology*, 70(11), 2189–2199. <https://doi.org/10.1080/17470218.2016.1227860>
- FLORIDOU, G. A., WILLIAMSON, V. J., STEWART, L., & MÜLLENSIEFEN, D. (2015). The Involuntary Musical Imagery Scale (IMIS). *Psychomusicology: Music, Mind, and Brain*, 25(1), 28–36.
- GAAB, N., GASER, C., ZAEHLE, T., JANCKE, L., & SCHLAUG, G. (2003). Functional anatomy of pitch memory - An fMRI study with sparse temporal sampling. *NeuroImage*, 19(4), 1417–1426. [https://doi.org/10.1016/S1053-8119\(03\)00224-6](https://doi.org/10.1016/S1053-8119(03)00224-6)
- HALPERN, A. R., & BARTLETT, J. C. (2011). The persistence of musical memories: A descriptive study of earworms. *Music Perception*, 28(4), 425–432. <https://doi.org/10.1525/mp.2011.28.4.425>

- HICKOK, G., BUCHSBAUM, B., HUMPHRIES, C., & MUFTULER, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in Area Spt. *Journal of Cognitive Neuroscience*, 15(5), 673–682. <https://doi.org/10.1162/089892903322307393>
- HYMAN, I. E., BURLAND, N. K., DUSKIN, H. M., COOK, M. C., ROY, C. M., McGRATH, J. C., & ROUNDHILL, R. F. (2013). Going gaga: Investigating, creating, and manipulating the song stuck in my head. *Applied Cognitive Psychology*, 27, 204–215. <https://doi.org/10.1002/acp.2897>
- HYMAN, I. E., CUTSHAW, K. I., HALL, C. M., SNYDERS, M. E., MASTERS, S. A., AU, V. S. K., & GRAHAM, J. M. (2015). Involuntary to intrusive: Using involuntary musical imagery to explore individual differences and the nature of intrusive thoughts. *Psychomusicology*, 25(1), 14–27. <https://doi.org/10.1037/pmu0000075>
- JAKUBOWSKI, K., BASHIR, Z., FARRUGIA, N., & STEWART, L. (2018). Involuntary and voluntary recall of musical memories: A comparison of temporal accuracy and emotional responses. *Memory and Cognition*, 1–16. <https://doi.org/10.3758/s13421-018-0792-x>
- JAKUBOWSKI, K., FARRUGIA, N., & STEWART, L. (2016). Probing imagined tempo for music: Effects of motor engagement and musical experience. *Psychology of Music*, 44(6), 1274–1288. <https://doi.org/10.1177/0305735615625791>
- JAKUBOWSKI, K., FINKEL, S., STEWART, L., & MÜLLENSIEFEN, D. (2017). Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, 11(2), 122–135. <https://doi.org/10.1037/aca0000090>
- JONES, D. M., & MACKEN, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 369–381. <https://doi.org/10.1037/0278-7393.19.2.369>
- KELLARIS, J. J. (2003). Dissecting earworms: Further evidence on the song-stuck-in-your-head phenomenon. *Proceedings of the Society for Consumer Psychology Winter 2003 Conference*, 220–222.
- KELLARIS, J. J. (2001). Identifying properties of tunes that get stuck in your head: Toward a theory of cognitive itch. In M. L. Cronley and D. Nayakankuppam, (Eds.). *Proceedings of the Society for Consumer Psychology Winter 2001 Conference* (pp. 66–67). New Orleans, LA.
- KOELSCH, S., SCHULZE, K., SAMMLER, D., FRITZ, T., MÜLLER, K., & GRUBER, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, 30(3), 859–873. <https://doi.org/10.1002/hbm.20550>
- KOPIEZ, R., PLATZ, F., MULLER, S., & WOLF, A. (2013). When the pulse of the song goes on: Fade-out in popular music and the pulse continuity phenomenon. *Psychology of Music*, 43(3), 359–374. <https://doi.org/10.1177/0305735613511505>
- LEINENGER, M. (2014). Phonological coding during reading. *Psychological Bulletin*, 140(6), 1534–1555. <https://doi.org/10.1037/a0037830>
- LIKKANEN, L. A. (2011). Musical activities predispose to involuntary musical imagery. *Psychology of Music*, 40(2), 236–256. <https://doi.org/10.1177/0305735611406578>
- LIKKANEN, L. A. (2012). Inducing involuntary musical imagery: An experimental study. *Musicae Scientiae*, 16, 217–234. <https://doi.org/10.1177/1029864912440770>
- LOCKE, J. L., & FEHR, F. S. (1970). Subvocal rehearsal as a form of speech. *Journal of Verbal Learning and Verbal Behavior*, 9(5), 495–498. [https://doi.org/10.1016/S0022-5371\(70\)80092-5](https://doi.org/10.1016/S0022-5371(70)80092-5)
- LUKATELA, G., FROST, S. J., & TURVEY, M. T. (1998). Phonological priming by masked nonword primes in the Lexical Decision Task. *Journal of Memory and Language*, 39(4), 666–683. <https://doi.org/10.1006/jmla.1998.2599>
- MACKWORTH, J. F. (1962). The effect of display time upon the recall of digits. *Canadian Journal of Psychology*, 16(246), 48–54. <https://doi.org/10.1037/h0083241>
- MANNING, F., & SCHUTZ, M. (2013). “Moving to the beat” improves timing perception. *Psychonomic Bulletin and Review*, 20(6), 1133–1139. <https://doi.org/10.3758/s13423-013-0439-7>
- MCCULLOUGH-CAMPBELL, S., & MARGULIS, E. H. (2015). Catching an earworm through movement. *Journal of New Music Research*, 44(4), 347–358. <https://doi.org/10.1080/09298215.2015.1084331>
- MCNALLY-GAGNON, A. (2016). *Imagerie musicale involontaire: Caractéristiques phénoménologiques et mnésiques*. (PhD thesis). University of Montreal, Canada.
- MÜLLENSIEFEN, D., FRY, J., JONES, R., JILKA, S., STEWART, L., & WILLIAMSON, V. J. (2014). Individual differences predict patterns in spontaneous involuntary musical imagery. *Music Perception*, 31(4), 323–338. <https://doi.org/10.1525/MP.2014.31.4.323>
- NAKAGAWA, S., & SCHIELZETH., H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- NEES, M. A., CORRINI, E., LEONG, P., & HARRIS, J. (2017). Maintenance of memory for melodies: Articulation or attentional refreshing? *Psychonomic Bulletin and Review*, 24, 1964–1970. <https://doi.org/10.3758/s13423-017-1269-9>
- PAWLEY, A. R., & MÜLLENSIEFEN, D. (2012). The science of singing along: A quantitative field study on sing-along behaviour in the north of England. *Music Perception*, 30, 129–146.
- PECENKA, N., & KELLER, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, 1169, 282–286. <https://doi.org/10.1111/j.1749-6632.2009.04785.x>
- PEIRCE, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>

- PERHAM, N., & SYKORA, M. (2012). Disliked music can be better for performance than liked music. *Applied Cognitive Psychology*, 26(4), 550–555. <https://doi.org/10.1002/acp.2826>
- PERRONE-BERTOLOTTI, M., RAPIN, L., LACHAUX, J. P., BACIU, M., & LÆVENBRUCK, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*, 261, 220–239. <https://doi.org/10.1016/j.bbr.2013.12.034>
- PRING, L., & WALKER, J. (1994). The effects of unvocalized music on short-term memory. *Current Psychology*, 13(2), 165–171. <https://doi.org/10.1007/BF02686799>
- SALAMÉ, P., & BADDELEY, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A*, 41, 107–122. <https://doi.org/10.1080/14640748908402355>
- SMITH, J. D., WILSON, M., & REISBERG, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia*, 33, 1433–1454.
- WEGNER, D. M. (1994). Ironic processes of mental control. *Psychological Review*, 101(1), 34–52. <https://doi.org/10.1037/0033-295X.101.1.34>
- WILLIAMS, T. I. (2015). The classification of involuntary musical imagery: The case for earworms. *Psychology of Music, Mind and Brain*, 25(1), 5–13. <https://doi.org/10.1037/pmu0000082>
- WILLIAMSON, V. J., BADDELEY, A. D., & HITCH, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory and Cognition*, 38(2), 163–175.
- WILLIAMSON, V. J., JILKA, S. R., FRY, J., FINKEL, S., MÜLLENSIEFEN, D., & STEWART, L. (2012). How do “earworms” start? Classifying the everyday circumstances of Involuntary Musical Imagery. *Psychology of Music*, 40(3), 259–284. <https://doi.org/10.1177/0305735611418553>
- WILLIAMSON, V. J., LIKKANEN, L. A., JAKUBOWSKI, K., & STEWART, L. (2014). Sticky tunes: How do people react to involuntary musical imagery? *PLoS ONE*, 9(1), e86170. <https://doi.org/10.1371/journal.pone.0086170>
- WILLIAMSON, V. J., & MÜLLENSIEFEN, D. (2012). Earworms from three angles: Situational antecedents, personality predisposition and the quest for a musical formula. In E. Cambouropoulos, C. Tsougras, K. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*, Thessaloniki, Greece: ICMPC/ESCOM.
- ZEIGARNIK, B. (1938). On finished and unfinished tasks. *A source book of Gestalt psychology*, 14(1), 300–314. <https://doi.org/10.1037/11496-025>