In press 2023, *Memory*

Can Synchronized Tones Facilitate Immediate Memory for Printed Lists?

Bailey Pannell[1], Dominic Guitard[2], Yu Li[1], and Nelson Cowan[1]

[1]University of Missouri

[2]Cardiff University

Author Notes

This research was supported by Grant R01-HD021338 from the National Institute of Child Health and Human Development to Nelson Cowan, and while working on this project Dominic Guitard was supported by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada. The research was based on the senior thesis of Bailey Brimer (now Pannell).

Address correspondence to Nelson Cowan, Department of Psychological Sciences, University of Missouri, McAlester Hall, Columbia, MO 65211 USA. Email: CowanN@missouri.edu.

Abstract

In verbal list recall, adding features redundant with the ones to be recalled theoretically could assist recall, by providing additional retrieval cues, or it could impede recall, by draining attention away from the features to be recalled. We examined young adults' immediate memory of lists of printed digits when these lists were sometimes accompanied by synchronized, concurrent tones, one per digit. Unlike most previous irrelevant-sound effects, the tones were not asynchronous with the printed items, which can corrupt the episodic record, and did not repeat within a list. Memory of the melody might bring to mind the associated digits like lyrics in a song. Sometimes there were instructions to sing the digits covertly in the tone pitches. In three experiments, there was no evidence that these methods enhanced memory. Instead, there appeared to be a distraction effect from the synchronized tones, as in the irrelevant sound effect with asynchronized tones.

Keywords: *Working memory, short-term memory, learning strategy, tones, verbal list memory*

Can Synchronized Tones Facilitate Immediate Memory for Printed Lists?

The last author of this article ([blinded for submission]) is old enough to remember Wilson Pickett's 1965 hit song in which he pleads for his love interest to call him at 634-5789. Yet, this author got the number wrong when searching for the song on the web. Many have seen videos of cultures in which melodies are used to help teach children the alphabet, text, or simple rules for behavior, and this has often been advocated (e.g., Almutairi, & Shukri, 2016; Fisher & McDonald, 2001; Noll, 2003; Orlova, 2003; Yeni & Amelia, 2020). Ferreri and Verga (2016) extensively reviewed the effects of music on verbal learning and memory and found that a variety of factors influenced whether it was helpful, neutral, or even harmful.

This phenomenon of possibly learning verbal material through melody gives rise to a more basic question that we studied. Specifically, can varied tones associated with items within a verbal list facilitate immediate recall of the list items? Immediate verbal, serial recall could be a precursor to verbal learning (e.g., Chen & Cowan, 2005). Arguments could be made, however, either against the possibility that the tones would help, or for it.

The information against the usefulness of tone cues comes from the literature on irrelevant sounds on immediate recall, but they may not include the situation we have in mind. Salamé and Baddeley (1982) found that streams of irrelevant speech impeded memory for concurrently-presented, printed verbal lists. In the irrelevant sound effect discovered by Dylan Jones and colleagues (e.g., Tremblay et al., 2000), a series of changing tones can interfere with memory for printed verbal items. Unlike the present work, however, in the irrelevant sound effect studies, the visual and acoustic presentations are generally not synchronized. In the object-oriented episodic record account of the irrelevant sound effect (Jones & Tremblay, 2000), each tone creates an entry in the episodic record that conflicts with the order of events described by the printed verbal items. If, however, the verbal items and tones were presented concurrently, each pairing of a verbal item with a tone might merge into an integrated, multimodal event, so that the tones would not be expected to cause interference.

Another basis for expecting a detrimental effect of the tones is an effect based on distraction. In particular, attention might be more difficult to keep on the verbal items in the presence of tones (on the attention cost of tones see Blain et al. 2022), an attention effect that has been suggested as the basis of multimodal interference in short-term or working memory tasks (Ayres & Sweller, 2014; Morey & Cowan, 2004; Sweller, 2011; for varying definitions of short-term and working memory see Cowan, 2017). This distraction could cut down on encoding of the verbal items or the use of attention to refresh them in memory (e.g., Barrouillet et al., 2011). Also, to the extent that there may be covert rehearsal of tones (e.g., Peynircioğlu, 1995), it could interfere with the covert rehearsal of the verbal items, causing a kind of articulatory suppression effect (e.g., Schendel & Palmer, 2007).

In contrast to the potential detrimental effects, it is possible that tones synchronized with printed digits will assist recall. A helpful effect of distinctive tones might be expected according to the notion that the retrieval process will be more successful if there are more distinctive cues that can facilitate retrieval. For example, according to the Feature Model (Nairne, 1990; Surprenant & Neath, 2009), the retrieval process should be aided by the notion that retrieval of the tones could serve as a cue to retrieval of the associated digits, at least sometimes. Memory performance on an item is improved when it is distinct from other items in the trial (e.g. von Restorff, 1933; Jacoby & Craik, 1979; Guérard et al., 2008, 2010) so it would be expected that supplying a unique tone for each printed digit could increase the distinctiveness of each item relative to the others. We investigate this possibility by including a control condition in which all of the tones are the same (and thus not distinct), to look for a detrimental effect of non-distinctiveness much like the phonological similarity effect in recall (e.g., Conrad, 1964).

Beyond the mere presentation of distinctive features in a tone series, there could be some additional benefit of having participants imagine singing the digits in the melody formed by the tones. In the production effect, items are retained better if the participant must produce these items during encoding rather than just receiving them. Quinlan et al. (2019) found a production effect for singing words to be remembered that was larger than a production effect for speaking the items, when singing, speaking, and remaining silent were instructions for different lists within each participant. In contrast, no such production effect was obtained when the instructions were between-participants, suggesting that the effect came from a boost in distinctiveness of the sung words compared to the other types. If we find a helpful effect of instructions to imagine singing the digits in the pitches of the accompanying tones compared to a no-instruction control condition, it can serve as evidence favoring a similar mechanism. In this case it would be not only the perceptual input that includes more features distinguishing items in a list from one another, but also the motor planning, which should be important according to accounts in which perceptual and gestural processes determine memorability (e.g., Hughes et al., 2009, 2011, 2016). Another way the tones could help is by strengthening and sharpening the perception of a regular rhythm for attending to the digits. The rhythm of attending is important according to a dynamic attending view (Jones & Boltz, 1989; Large & Jones, 1999), a view that has been supported by considerable evidence (Povel & Essens, 1985; Purnell-Webb & Speelman, 2008; Silverman, 2012; Tillmann & Dowling, 2007). The presence of a regular rhythm during working memory tasks has resulted in improved memory performance compared to a silent condition

(Plancher et al., 2018; Fanuel et al., 2018; 2020). There are good reasons to believe that the temporal precision of sounds in immediate memory is greater than the temporal precision of visual presentations (Penney, 1989), which should mean that the attending rhythm is more precise for our tones than for our printed digits.

Recent research gives some clues to the mechanisms of irrelevant tones in serial recall of visually-presented lists. Some recent work describes a dual mechanism in which irrelevant sounds both cause distraction and interrupt the episodic record of the visual stimulus order (e.g., Marsh et al., 2020; Röer et al., 2014; Stokes & Arnell, 2012). Changing state seems to be a more important factor than the rhythm in which interfering tones have been presented (Hall & Gathercole, 2011; Parmentier & Beaman, 2015). Yet, to our knowledge, no experiments have examined effects of irrelevant tones synchronized with the items to be remembered, and it is unclear whether disruption should still occur under these circumstances.

In sum, given that there are a number of conflicting expectations from previous work, we have no clear predictions as to whether tone series synchronized with the items to be presented and/or imagined singing of the digits in those tone pitches will help with digit list recall, or whether it will interfere. Given the possibility that the correct answer is a null effect, we use Bayesian statistical methods, in which the relative probability of the null and non-null hypotheses can be assessed, providing positive support to either the null or non-null hypothesis (for an introduction see Etz & Vandekerckhove, 2018).

Experiment 1 investigated compared effects of changing tones, a repeated single tone, and silence on the recall of printed digits. Experiment 2 extended the examination of changing tones to a situation in which the mapping between digits and tones was consistent and regular. Given that these conditions still did not show a benefit of the tones, Experiment 3 involved separate trial blocks for silent and changing-tone conditions to improve learning of the digit-to-tone mapping. All three experiments also explored the potential effects of having participants imagine singing the digits in the pitches of the accompanying tones.

Experiment 1

In Experiment 1, participants were given three different acoustic conditions (changing tones, steady tones, and silent) with lists of 8 digits. In the changing tones condition, a randomly selected tone was played simultaneously as a digit appeared on the screen. The steady tone condition was the same except that, instead of a new note for each digit, the same, constant note was repeated for each digit in the list. The silent condition involved the digit list memory test without any tones. Half of the participants were given no specific directions for this task on how to use the tones, while the other half were provided with pre-training instruction and an example of how to use the tones to help them imagine singing the tones presented. An illustration of the method for all experiments is shown in Figure 1.

The singing instructions were meant to be exploratory. Given that the tones were presented separately from the printed digits to be recalled, we found it unlikely that the participants would spontaneously sing the printed digits when there was no instruction to do so. Whereas rehearsal of tones to be retained would be expected to elicit some subvocalization (Reisberg et al., 1989), it was not expected for our tones that were not to be recalled except when participants were instructed to imagine singing, integrating the tone pitches with the digits. We do, however, expect that most individuals will tend to subvocalize the digits themselves, given that subvocalization seems necessary to account for the phonological similarity effect for printed items (Conrad, 1964). The singing instructions were included primarily to make it likely that integration of the tones into the recall response occurred for at least one group.

Method

Participants. All of the experiments were approved by the University of Missouri Institutional Review Board. The participants were recruited from an online data collection agency (https://www.prolific.co/) and compensated six dollars for their participation. Inclusion criteria were as follows: (a) the participants were 18 to 30 years of age; (b) they were native English speakers; (c) their nationality was British, American (U.S.), or Canadian; (d) they had no language-related disorder, cognitive impairment, dementia, or difficulty hearing; (e) they had normal or corrected-to-normal vision; and (f) they had an approval rating of at least 90% in previous submissions at Prolific. The criteria (a) through (e) were self-reported, and criteria (f) is computed by Prolific.

Based on the criteria, sixty-seven participants were gathered for the study initially; however, seven participants were excluded based on insufficient attention to the tones, another pre-established criterion for inclusion. Specifically, after each trial, participants had to identify if no tone was presented, one tone multiple times, or multiple different tones; participants with accuracy inferior to 80% were excluded. In the final sample, data from sixty participants were included for analyses. Thirty participants were randomly assigned to each of two groups, one without singing instruction (uninstructed) and one with imagine-singing instruction. The 30 participants in the imagine-singing group had a mean age of 23.80 (*SD* = 3.86, range = 18-30); 22 self-identified as a female, 6 as male, and 2 identified as other. The 30 participants in the uninstructed group had a mean age of 25.07 (*SD* = 2.85, range = 19-29); 23 self-identified as a female and 7 as male.

Materials. This experiment and all subsequent experiments were programmed with PsyToolKit (Stoet, 2010, 2017). The auditory stimuli were created with a free open-source audio software called Audacity (https://www.audacityteam.org). The tones were the following: B3 (246.94 Hz), C4 (261.63 Hz), D4 (293.66 Hz), E4 (329.63 Hz), F4 (349.23 Hz), G4 (392.00), A4 (440.00 Hz), B4 (493.88 Hz). Each tone was created as a 500-ms sine waveform with an amplitude of 0.5 on a scale from zero to one. The first 25 ms of the 500 ms wave form was used to create a linear fade-in from 0 (silent) to 0.5 (maximum amplitude). Likewise, the last 25 ms of the 500 ms waveform was used to create a linear fade-out from 0.5 (maximum amplitude) to 0 (silent). For each tone after the 500 ms waveform, a 250 ms silence was included to ensure a smooth transition within the program. The visual stimuli were integers from one to nine, presented at the center of the screen in white 30-point Times New Roman font on a black background.

Design. The general design of the experiment is a 2 × 3 × 8 design, with the absence or presence of imagine-singing instructions between participants and with two repeated measures: acoustic condition (silent vs. steady tone vs. changing tones) and serial position (1 to 8). There were three practice trials, one for each acoustic condition, and 48 experimental trials, 16 for each acoustic condition. The order of the acoustic conditions was randomized for the three practice trials for each participant, followed by the experimental trials in which the repetitions of the three trial types were randomized within each participant. On each trial, the eight digits were randomly selected from integers one to nine without replacement.

In the steady tone condition, one randomly-selected tone was played for all eight digits on a given trial. Each of the eight possible tones was used in two trials for that condition. In the changing tones condition, the eight different tones were presented one at a time, coincident with the visual presentation of the eight digits on a given trial. The order of the tones was pre-determined using a balanced Latin Square as shown in Table 1. For all participants, the tones were presented in the same order within a given list, but the order of the list and the digits associated to each tone were random for each participant. The results sounded musical to the investigators, although the resulting melodies would not consistently fit a stereotype of a simple melody like a children's tune within European or American culture.

The participants were presented with the stimuli and were to recall the digits in the presented order, and to indicate what acoustic condition had been presented. The instructed group was given further instruction to encourage them to imagine singing the digits in the melody formed by the accompanying tones, to help them remember the digits. These 30 participants were given two examples of what that might sound like. Each participant in the imagine-singing group received one singing example in a female voice and one in a male voice. For the female voice, the digits were *7-4-1-5-8-3-6-2* sung in the notes *G4 A4 B4 D4 B3 C4 F4*, and *E4*. For the male voice, the digits were *5-3-6-2-4-1-7-8*, sung in the notes *B3 E4 B3 E4 B4 A4 B4*, and *F4*.

One example was in a male voice and the other was in a female voice, in both cases singing each digit in the concurrently presented tone. Participants were instructed to only imagine singing in this way, not to sing aloud. They were given an example of what the trial would look like, and then the vocal example was played.

Procedure. All participants were tested in one experimental session lasting approximately 30 to 40 minutes. Before the beginning of each experimental session, electronic informed consent was given. The participants were required to have headphones on or speakers on before the beginning of the experimental session. Participants were allowed to self-adjust the volume to their comfort level after hearing the 8 tones. As shown in Figure 1, the participants initiated each trial by pressing the space bar of their keyboard. First, a fixation cross was presented at the center of the computer screen for 500 ms follow by a 1000 ms blank interval. After that, the eight digits were presented sequentially at a rate of one digit per second at the center of the computer screen. In the steady tone or changing tone condition, the tones were presented simultaneously with the onset of the visual stimuli for a duration of 500 ms. After the digits were presented, the participants were to type the digits that they could recall in the presented order, using the keyboard or number pad. The responses were displayed immediately on the center of the screen and the participants were able to backtrack to change a response with the backspace key until they pressed the Enter key. Once the participant pressed the Enter key, their answer was recorded and could not be changed.

The trials of different tone conditions were presented in random order. After the participants recalled the digits as best they could, they were given a multiple-choice question: How many different tones were presented on this trial? Participants could select one of three options: A) No tone was presented; B) Only one tone was presented multiple times; or C) Multiple different tones were presented. This question was added to ensure that the participants were carefully listening to each tone, an inclusion criterion as previously described in the section on participants.

Data analysis. For all experiments, responses were scored via a strict recall criterion. Based on this criterion, a digit must be recalled at its presentation position to be scored as correct. In all experiments, the proportion of correct responses was analyzed with the statistical software R (R Core Team, 2021). We report both Frequentist and Bayes factor statistical analyses, the former for descriptive purpose only and the latter for statistical inference. Our Bayes factor analyses were conducted with the “Bayes Factor” R package and the default priors (version 0.9.12-4.2; see R. Morey & Rouder, 2018; Rouder et al., 2009, 2012). To ensure reliable estimates, 100,000 iterations followed by 10,000 additional iterations were conducted until the proportional error of the computation was below 5%. In all analyses, participants were included as a random factor, and main effects and interactions were tested by omitting each effect from the full model. For Bayes Factor analyses, we reported either BF_{10}, which corresponds to evidence for the alternative hypothesis, or BF_{01} ($BF_{01} = 1/BF_{10}$), which corresponds to evidence for the null hypothesis. For all Frequentist statistical analyses, we used the “ez” (version 4.4-0; Lawrence, 2016) R package (ANOVAs) and the “lsr” (version 0.5; Navarro, 2015) R package (t-tests).

Results

For included participants, the check on hearing the tones was successful, with 0.94 proportion correct responses on what the acoustic condition was (SD=0.24) for the uninstructed group and 0.95 proportion correct responses in the group instructed to imagine singing (SD=0.21).

Figure 2 shows the proportions of correct responses as function of acoustic conditions (silent, steady tone, changing tones) for each instruction group (no singing instruction, imagine-singing instruction). Overall, participants with no singing instruction (*M* = .81, *SD* = .15) recalled digits better than those with the imagine-singing instruction (*M* = .68, *SD* = .20). Participants were also better in the silent condition (*M* = .77, *SD* = .19) relative to the steady tone (*M* = .75, *SD* = .20) and changing tones (*M* = .72, *SD* = .20), a detrimental effect of tone conditions across both groups. An ANOVA with acoustic conditions and instruction group supported these descriptive results. The results from the analysis revealed the presence of a main effect of singing instruction favoring no instruction, $F(1,58) = 7.59$, $\eta_p^2 = .12$, $BF_{10} = 3.45$, and a main effect of acoustic conditions favoring silence, $F(2,116) = 8.03$, $\eta_p^2 = .12$, $BF_{10} > 1000$. However, there was no interaction between singing instruction and acoustic conditions, $F(2,116) = 1.26$, $\eta_p^2 = .02$, $BF_{01} = 7.09$.

We further investigate the main effect of acoustic conditions by conducting separate Bayesian paired-sample t-tests. Results revealed comparable performance between the silent and the steady tone condition, $BF_{01} = 1.97$, Cohen's d = 0.21, although only anecdotal evidence was found in favor of the null hypothesis. The analyses also revealed that participants were better in the silent relative to the changing tones condition, $BF_{10} = 33.07$, Cohen's d = 0.46, and better in the steady tone condition relative to the changing tones conditions, $BF_{10} = 4.81$, Cohen's d = 0.36.

Discussion

Overall, the results of Experiment 1 were clear. With or without singing instructions, participants performed worse, not better, in the presence of changing tones, with no effect of a steady tone. Thus, even though the tones were coordinated with the digits, the results are in keeping with the irrelevant sound effect obtained without coordination of the memoranda and the tones (e.g., Tremblay et al., 2000).

We wondered whether a potential benefit of tones, at least in the group in which singing was imagined, might occur if the tones provided another modality of information that could identify the digit. In the first experiment, this was not the case, as the association of tones to digits was arbitrary and changed from trial to trial. In the second experiment, we kept the association fixed across trials, with low tones corresponding to low digits and increasing tone frequency (and perceived pitch) for increasing digits. This consistent mapping between digit magnitudes and tone frequencies seems analogous to the effects of consistency between digits and places on the number line from left to right (e.g., Dehaene et al., 1993). Keller et al. (1995)

showed that non-musicians can associate both series of digits and series of tones to higher-versus-lower places on a vertical scale at an accuracy much better than chance.

Experiment 2

Experiment 2 was identical to Experiment 1 except for the following changes. In Experiment 2, only the silent and changing-tones conditions were used (again in random order). The silent condition was identical to Experiment 1 while the changing tones condition was altered to have a tone consistent with each digit (1 to 8, as digit 9 was not used For Experiment 2). Lower tones correlated with lower digits and higher tones with higher digits.

Method

Participants. Sixty different participants were recruited from Prolific. The inclusion criteria were identical to Experiment 1 with the addition that participants had not taken part in Experiment 1. In this experiment, no participants were excluded for failing to attend to the tones. Half of the participants were randomly assigned to the no-instruction group and the other half to the imagine-singing instruction group. The mean age of the participants in the imagine-singing group was 24.37 (*SD* = 3.61, range = 19-30); 19 self-identified as a female, 10 as male, and 1 identified as other. The mean age of the participants in the uninstructed group was 22.93 (*SD* = 3.10, range = 18-30); 16 self-identified as a female, 10 as male, and 4 identified as other.

Materials, **Design, and Procedure.** The materials, the design, and the procedure in Experiment 2 were like Experiment 1 except that, rather than three conditions, there were only silent and changing tones conditions, and no constant tone condition. Most importantly, though, in the changing tones condition, unlike Experiment 1, each specific tone frequency was consistently associated with a specific digit (1 to 8, as digit 9 was not used for Experiment 2), with lower tone frequency associated with small digits and higher tone frequency associated with larger digits (see Figure 1). Participants were not informed of the tone-digit consistency in either group.

In this experiment, participants completed 2 practice trials (1 trial silent, 1 trial changing tones) and 60 experimental trials (30 trial silent, 30 trial changing tones). The acoustic conditions were randomized as in Experiment 1.

Results

As in Experiment 1, The check on hearing the tones was again successful, with 0.97 proportion correct responses on what the acoustic condition was (SD=0.18) for the uninstructed group and 0.98 proportion correct responses in the imagine-singing group (SD=0.15).

Figure 3 shows the proportions of correct responses as function of acoustic conditions (silent and changing tones) for each group (uninstructed and imagine-singing). Overall, participants with the imagine-singing instruction (*M* = .82, *SD* = .12) had slightly higher means than participants without singing instruction (*M* = .76, *SD* = .15). Participants were better in the silent condition (*M* = .81, *SD* = .13) relative to the changing tones (*M* = .77, *SD* = .15).

An ANOVA with acoustic conditions and singing instruction group as factors clarified these descriptive results. The results from the analysis revealed the presence of a main effect of acoustic condition favoring silence, $F(1,58) = 12.76$, $\eta_p^2 = .18$, $BF_{10} > 1000$. However, the results from the analysis were inconclusive regarding a main effect of singing instruction, $F(1,58) = 2.82$, $\eta_p^2 = .05$, $BF_{01} = 2.76$, and there was no interaction between singing instruction and acoustic condition, $F < 1$, $\eta_p^2 = .00$, $BF_{01} = 35.29$.

At first glance, the data may seem inconsistent with the statistical result, given that the condition SDs are slightly smaller for singing instruction group than for acoustic condition, and the difference between means is slightly larger for singing instruction group than for acoustic condition, yet only the acoustic condition produced a reliable effect. Further inspection showed that the basis of this pattern of results was that there was a high correlation between performance levels in the two acoustic conditions (silent and changing tones), so that the SD for the difference between acoustic conditions was only .09. Given this situation, the within-participant manipulation could be examined more powerfully than the between-participant manipulation.

Discussion

Overall, the results of Experiment 2 were clear. With or without singing instructions, participants performed worse in the presence of auditory information. These results are in line with the results of Experiment 1 and were observed despite the methodological changes, facilitating the association between the digits and the tones in Experiment 2.

Experiment 3

In Experiment 3, the two acoustic conditions (silent and changing tones) were presented in counterbalanced, separate blocks rather than being mixed and randomized as in the previous experiments. This change was implemented to further facilitate the formation of associations between tones and the digits. Specifically, because the trials were blocked, in the changing tone condition of this experiment, there is a tone associated with each digit consistently for each trial within the block, which could facilitate learning. This blocked presentation also should minimize the extent to which participants mentally apply the tones to the digits presented in the silent condition, revealing any potential benefit of the tones, or of imagining singing in those tones, that might have been obscured by the method used in the previous experiments.

Method

Participants. Based on this criterion that is identical to that of Experiment 1, sixty-one participants who did not participate previously were gathered for the study, one of which was excluded because of the predetermined criterion of insufficient attention to the tones. Half were randomly assigned to the no-instruction group and the other half to the imagine-singing instruction group. The mean age of the participants in the imagine-singing group was 24.96 (*SD* = 3.70, range = 18-30); 23 self-identified as a female, and 7 as male. The mean age of the participants in the uninstructed group was 25.63 (*SD* = 3.11, range = 19-30); 16 self-identified as a female, 13 as male, and 1 identified as other.

Design. Experiment 3 is identical to Experiment 2 except that the acoustic conditions were presented in two separate, counterbalanced blocks of trials instead of one mixed block. Half of the participants in each group (imagine-singing, uninstructed) started with the 30 silent digit span trials and then received the 30 changing tone trials, while the other half of the participants received the trial blocks in the reverse order.

Because we used separate blocks of silent and changing-tone trials in this experiment, the previous post-check on whether headphones were on seemed inadequate. It would be possible to figure out that the answer to the post-check was “tones present” for an entire trial block. Therefore, we added a pre-check. To start a trial, participants had to listen to the message, press the letter H or press the letter C and press the corresponding letter key to start the trial.

Results

The check on hearing the tones done after the trial was again successful, with 0.99 proportion correct responses on what the acoustic condition was (SD=0.12) for the uninstructed group and 0.99 proportion correct responses in the imagine-singing group (SD=0.12). The additional check on hearing done before the trial revealed similar success, with 0.99 (SD = 0.10) proportion correct responses for the uninstructed group and 0.99 proportion correct responses in the imagine-singing group (SD = 0.11 for both groups).

Figure 4 shows the proportions of correct responses as function of acoustic conditions (silent, changing tones) for each group (uninstructed, imagine-singing). Overall, participants without instruction (*M* = .82, *SD* = .13) were better than participants with the imagine-singing instruction (*M* = .72, *SD* = .18). Participants were also better in the silent condition (*M* = .79, *SD* = .16) relative to the changing tones (*M* = .75, *SD* = .18).

The results from the ANOVA with tone conditions and singing instruction confirmed these descriptive trends. The results from the analysis revealed inconclusive evidence of a main effect of singing instruction, $F(1,58) = 5.71$, η_p^2=.09, $BF_{10} = 1.39$, but revealed the presence of a main effect of acoustic condition, $F(1,58) = 13.97$, η_p^2 =.19, $BF_{10} > 10{,}000$, favoring silence over changing tones. There was no interaction between singing instruction and acoustic condition, $F< 1$, η_p^2 = .00, $BF_{01} = 53.03$.

Last, we wondered if performance in the changing-tone condition might increase as a function of trials across the trial block, given that the block was interrupted by silent trials. However, we did not find this to be the case. Inspection of the performance function for silent and changing-tone conditions, for the first and second block of trials separately, and for the two instruction groups separately (see Figure 5) show no practice effect across trials within a trial block. In an ANOVA to examine the effect of the first and second block, the block could not be directly entered as a factor because different participants received different conditions in the two blocks. The difference between blocks would emerge as a main effect of condition order (silent-tone, tone-silent), but there was anecdotal evidence against this effect, $F(1,56) = 2.48$, η_p^2=0.04, $BF_{01} = 2.59$. There was the expected effect of condition $F(1,56)=13.52$, η_p^2= 0.19, $BF_{10} > 10{,}000$, but there was evidence against the interaction between order and condition, $F< 1$, η_p^2= 0.00, $BF_{01} = 50.94$.

Discussion

With or without singing instructions, participants performed worse in the presence of auditory information. The persistence of an irrelevant sound effect even with instructions to imagine singing is important because it demonstrates that participants did appear to have followed the directions to imagine singing the tones. The fact that the tones actually provided information regarding the digits was apparently not usable. Perhaps that cueing potential was outweighed by the distraction from phonological information to melodic information.

General Discussion

In this study, we examined whether a tone series concurrent with printed digits facilitates immediate memory performance. One would expect that it does based on common lore, common educational practice in many countries, and published educational theory (Almutairi, & Shukri, 2016; Fisher & McDonald, 2001; Noll, 2003; Orlova, 2003; Yeni & Amelia, 2020). In the literature more closely tied to our use of a random order of tones synchronized with printed digits to be recalled, the expectations are split. We found no evidence supporting the usefulness of the tones, even when specific tones were consistently mapped to digits in a monotonic manner across trials, whether or not participants were encouraged to imagine singing the stimuli in the pitches of the corresponding tones.

Overall, the results were clear and can be summarized further as follows. The results of all three experiments showed that a tone series had a detrimental effect on recall. In Experiment 1, we found that performance was negatively affected by the changing tones when participants tried to recall digit lists, whether or not they were instructed to imagine singing the tones.

Instructions to imagine singing only worsened performance further. In Experiment 2, we established a fixed and systematic pairing between digits and the accompanying tones, with lower tones accompanying lower digits, but the detrimental effect of the changing tones remained. There was no longer a detrimental effect of imagining singing, but there was no statistical evidence that it helped. Last, in Experiment 3, we used the same procedure as in Experiment 2 but with silent and changing-tone trials presented in separate trial blocks. However, the results were the same as in Experiment 2. The effect of instruction groups was indeterminate, and a detrimental effect of changing tones remained.

The three experiments are compatible with the literature indicating that there are irrelevant sound effects that distract from verbal working memory (e.g., Jones & Macken, 1993; Tremblay et al., 2000). Unlike what might be expected from an object-oriented episodic record framework for the effect, though, we found it even when tones are synchronized with visually-presented verbal items and therefore do not enter irrelevant time points on the episodic record.

In Experiments 2 and 3, in which the digit-to-tone pairing was systematic and consistent across trials, the absence of a benefit of the tones can be viewed as surprising. Rather than being truly irrelevant, they are correlated with the correct digits. There is a vivid sensory memory for tones lasting several seconds (e.g., Cowan, 1984; Treisman & Rostron, 1972), which might have served as a recall cue. For example, if the memory indicated that the last three tones were descending, in Experiments 2 and 3 this would indicate that the last three digits were descending in a similar pattern. Apparently, any such cue was weak compared to the distracting effects of the tones.

The imagine-singing instruction might have been expected to integrate the tone pattern with the digit pattern, which would enrich the feature set of the items to be remembered. In

Experiments 2 and 3, because of the consistent mapping of tones to digits, memory of the imagined singing theoretically could assist in retrieving or guessing at digits. For example, memory of going high for two consecutive notes in the middle of the list would be a clue that two digits in the middle of the list were large, a property that might otherwise not have been encoded from the digits and could help in guessing them. However, the potential benefit was absent, suggesting that enriching the feature set of stimuli may not always be worth the effort.

We are unable to assess completely the roles of the various competing theories related to immediate memory based on the outcome of the study. One point that seems very clear, however, is that the irrelevant sound effect of Tremblay and Jones (2000) and others may still apply when the tones are synchronized with the information to be remembered, which might not be expected according to the view that the tones disrupt performance by adding to an object-oriented episodic record of the list (e.g., Jones & Tremblay, 2000). An alternative basis of irrelevant tone interference with memory is through distraction of attention away from the digits, a mechanism suggested by Cowan (1995, 1999) and verified in relation to irrelevant sound effects by a diminished irrelevant effect as a function of time across the experiment (Röer et al., 2014).

The absence of a helpful effect of tone cues for printed digits might be understood through a view in which interference effects occur through similarities in process rather than content (Marsh et al., 2009). The process of perceiving tones does not seem similar to the process of perceiving digits, but perhaps both are similar in terms of the salience of the temporal order dimension. Perhaps the presentation of tones and digits separately allows them to compete for temporal order perception and short-term memory encoding, in contrast to what might happen if they were presented in an integrated fashion. In future work, it would be of interest to include conditions in which both the tones and the digits have to be remembered, as these two types of memory should not compete for temporal order storage if memory for tones is part of a modular system for musical memory (Peretz & Coltheart, 2003).

In order for these results to be applied in practical settings, one still would have to learn whether the detrimental effect of tone series or melodies that are covertly sung applies also when they are overtly sung. The acoustic and motor feedback from overt singing could reinforce memory for the verbal items in a way that covertly imagining singing the items cannot do.

If there is an advantage of singing aloud over the present conditions, the interpretation of that finding would depend on a comparison of singing to simply speaking the items aloud. Recall that in a previous study of that nature (Quinlan et al., 2019), a benefit of singing over speaking occurred only in comparison to unsung items presented to the same participants. Therefore, its educational use would not be in enhancing memory overall, as is usually assumed, but rather in providing emphasis for some material in comparison to other material that is not sung.

One limitation of the present research is that we have not examined long-term retention of the items. Doing so would require using different words for each trial rather than digits repeated from trial to trial. It is possible that the distinctiveness afforded by tones would make lists in such an experiment memorable in the long term without facilitating immediate recall. It has been found that semantic structure can have a differential effect on long-term memory without assisting immediate memory (e.g., Glanzer & Schwartz, 1971), but it is unclear if the same is true of melodic structure. We also did not examine the possibility that repeated presentation of the same material will reveal a beneficial effect of tones and imagined singing. Unlike studies that can span weeks or even months, the current experiments last only about 30 minutes, and there was no dedicated learning session for the tone-to-digit associations. Moreover, we could not verify that participants complied with the imagine-singing instructions.

In sum, the extra features provided by tones in frequencies that covary with printed digits was of no benefit in remembering the digits and under some circumstances was detrimental. There is much left to learn about whether the results would change with different circumstances, such as overt motor involvement or spoken digits presented in different sung frequencies. One would expect more helpful results of music based on common lore, common educational practice in many countries, and published educational theory. Until extensions of the present research are carried out, for example with much longer periods of exposure to melodies associated with material to be learned, we suggest caution in using music to reinforce verbal information. The present study suggests that there are probably some undesirable aspects of dividing attention between verbal materials to be learned and accompanying, incidental tonal information, at least early in the learning process.

Open Practices Statement

The materials and the data can be found on the Open Science Framework page associated with this project: https://osf.io/wdxaq/?view_only=9768534a6be54cffba7c9a5a12d23ec9. The study was not preregistered.

References

Almutairi, M., & Shukri, N. (2016). Using songs in teaching oral skills to young learners: Teachers' views and attitudes. *International Journal of Linguistics* 8, 133-153. DOI: 10.5296/ijl.v8i6.10464

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R.E. Mayer, *The Cambridge handbook of multimedia learning*, second edition. Cambridge University Press. (p. 206-226). DOI: 10.1017/9781108894333

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review, 118*, 175-192. doi: 10.1037/a0022324.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology, 55*, 75-84.

Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1235-1249.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96*, 341–370. https://doi.org/10.1037/0033-2909.96.2.341

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford, UK: Oxford University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review, 24*, 1158–1170. DOI: 10.3758/s13423-016-1191-6

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396. https://doi.org/10.1037/0096-3445.122.3.371

Elliott, E.M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal, color-word interference task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27,* 654-667. DOI: 10.1037/0278-7393.27.3.654

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin & Review, 25*, 5–34. DOI 10.3758/s13423-017-1262-3

Fanuel, L., Portrat, S., Tillmann, B., & Plancher, G. (2018). Temporal regularities allow saving time for maintenance in working memory. *Annals of the New York Academy of Sciences, 1424*, 202–211.

Fanuel, L., Portrat, S., Dalla Bella, S., Tillmann, B., & Plancher, G. (2020). Temporal regularities during maintenance benefit short-term memory in the elderly. *Experimental Aging Research, 46,* 396-415.

Ferreri, L., & Verga, L. (2016). Benefits of music on verbal learning and memory: How and when does it work? *Music Perception, 34*, 167–182. https://doi.org/10.1525/mp.2016.34.2.167

Fisher, D., & McDonald, N. (2001). The intersection between music and early literacy instruction: Listen to literacy! *Reading Improvement, 38* (Fall), 106-115.

Guérard, K., Hughes, R. W., & Tremblay, S. (2008). An isolation effect in serial memory for spatial information. *The Quarterly Journal of Experimental Psychology, 61*, 752–762. Doi: 10.1080/17470210701402331

Guérard, K., Neath, I., Surprenant, A.M. et al. (2010). Distinctiveness in serial memory for spatial information. *Memory & Cognition*, *38*, 83–91. Doi: 10.3758/MC.38.1.83

Glanzer, M., & Schwartz, A. (1971). Mnemonic structure in free recall: Differential effects on STS and LTS. *Journal of Verbal Learning and Verbal Behavior*, *10*, 194-198. DOI: 10.1016/S0022-5371(71)80013-0

Hall, D., & Gathercole, S.E. (2011). Serial recall of rhythms and verbal sequences: Impacts of concurrent tasks and irrelevant sound. *Quarterly Journal of Experimental Psychology*, *64*, 1580–1592.

Hughes, R.W., Chamberland, C., Tremblay, S., & Jones, D.M. (2016). Perceptual-motor determinants of auditory-verbal serial short-term memory. *Journal of Memory and Language*, *90*, 126-146. Doi: 10.1016/j.jml.2016.04.006

Hughes, R.W., Marsh, J.E., & Jones, D.M. (2009). Perceptual-gestural (mis)mapping in serial short-term memory: The impact of talker variability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1411-1425 Doi: 10.1037/a0017008

Hughes, R.W., Marsh, J.E., & Jones, D.M. (2011). Role of serial order in the impact of talker variability in short-term memory: Testing a perceptual organization-based account. *Memory & Cognition*, *39*, 1435-1447. Doi: 10.3758/s13421-011-0116-x

Jacoby, L.L., & Craik, F.I.M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of Processing and Human Memory* (pp. 1-21). Hillsdale, N.J.: LEA.

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, 369–381. DOI:10.1037//0278-7393.19.2.369

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review, 7*, 544–558.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459–491. Doi: 10.1037/0033-295X.96.3.459

Keller, T.A., Cowan, N., & Saults, J.S. (1995). Can auditory memory for tone pitch be rehearsed? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 635-645. doi: 10.1037//0278-7393.21.3.635

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119–159. Doi: 10.1037/0033-295X.106.1.119

Lawrence, M. A. (2016). ez: Easy Analysis and Visualization of Factorial Experiments. R package version 4.4-0. https://CRAN.R-project.org/package=ez

Marsh, J.E., Hughes, R.W., & Jones, D.M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*. *110*, 23-38. doi: 10.1016/j.cognition.2008.08.003.

Marsh, J.E., Campbell, T.A., Vachon, F., Taylor, P.J., & Hughes, R.W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, & Psychophysics, 82*, 350–362. https://doi.org/10.3758/s13414-019-01800-w

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. Psychonomic Bulletin & Review, *1*, 296–301. DOI: 10.3758/BF03196573

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*, 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Nairne, J.S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251-269. doi: 10.3758/bf03213879

Navarro, D. J. (2015) Learning statistics with R: A tutorial for psychology students and other beginners. (Version 0.5) University of Adelaide. Adelaide, Australia Online at https://open.umn.edu/opentextbooks/textbooks/559

Noll, C. E. (2003). A comparison of bible verse memorization using traditional techniques versus using song. Doctoral dissertation, Johnson Bible College. Retrieved from https://files.eric.ed.gov/fulltext/ED480896.pdf

Orlova, N. (2003). Helping prospective EFL teachers learn how to use songs in teaching conversation classes. The Internet TESL Journal, 9(3). Retrieved from http://iteslj.org/Techniques/Orlova-Songs.html

Parmentier, F.B.R., & Beaman, C.P. (2015). Contrasting Effects of Changing Rhythm and Content on Auditory Distraction in Immediate Memory. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, *69*(1), 28–38.

Penney, C.G. (1989). Modality effects and the structure of short term verbal memory. *Memory & Cognition*, *17*, 398-422.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688-691. doi: 10.1038/nn1083.

Peynircioğlu, Z. F. (1995). Covert rehearsal of tones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 185–192. https://doi.org/10.1037/0278-7393.21.1.185

Plancher, G., Lévêque, Y., Fanuel, L., Piquandet, G., Tillmann, B. (2018). Boosting Maintenance in Working Memory with Temporal Regularities. *Journal of Experimental Psychology: Learning, Memory and Cognition, 44*, 812-818. doi: 10.1037/xlm0000481.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*, 411–440.

Purnell-Webb, P., & Speelman, C. P. (2008). Effects of music on memory for text. *Perceptual and Motor Skills, 106*, 927–957.

Quinlan, C. K., & Taylor, T. L. (2019). Mechanisms underlying the production effect for singing. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 73(4), 254–264. https://doi.org/10.1037/cep0000179

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reisberg, D., Smith, J. D., Baxter, D. A., & Sonenshine, M. (1989). "Enacted" auditory images are ambiguous; "Pure" auditory images are not. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 41*(3-A), 619–641. Doi: 10.1080/14640748908402385

Röer, J.P., Bell, R., & Buchner, A. (2014). Evidence for habituation of the irrelevant-sound effect on serial recall. *Memory & Cognition, 42*, 609–621. https://doi.org/10.3758/s13421-013-0381-y

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237. https://doi.org/10.3758/PBR.16.2.225

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. Journal of Mathematical Psychology, 56, 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Salamé, P., and Baddeley, A. (1982). Disruption of short term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*, 150 164.

Schendel, Z. A., & Palmer, C. (2007). Suppression effects on musical and verbal memory. *Memory & Cognition*, *35*, 640–650. https://doi.org/10.3758/BF03193302

Silverman, M. J. (2012). Effects of melodic complexity and rhythm on working memory as measured by digit recall performance. *Music and Medicine*, *4*, 22–27.

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolKit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Stokes, K.A., & Arnell, K.M. (2012). New considerations for the cognitive locus of impairment in the irrelevant-sound effect. *Memory & Cognition*, *40*, 918–931. DOI 10.3758/s13421-012-0194-4

Surprenant, A. M., & Neath, I. (2009). *Principles of memory*. New York: Psychology Press.

ISBN 9781138882928

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation*: *Cognition in education* (pp. 37–76). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory & Cognition, 35*, 628–639.

Tremblay, S., Nicholls, A. P., Alford, D., & Jones, D. M. (2000). The irrelevant sound effect: Does speech play a special role? Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(6), 1750–1754. DOI: 10.1037/0278-7393.26.6.1750

Treisman, M., & Rostron, A.B. (1972). Brief auditory storage: A modification of Sperling's paradigm. *Acta Psychologica, 36*, 161 170.

von Restorff, H. (1933). "Über die Wirkung von Bereichsbildungen im Spurenfeld" [The effects of field formation in the trace field]. *Psychologische Forschung* [*Psychological Research*] (in German). 18 (1): 299–342. doi:10.1007/BF02409636

Yeni, M., & Amelia, R. (2020). Teaching alphabet for young learners through song. *Journal of English Language and Education, 5*, 12-22. https://jele.or.id/index.php/jele/index

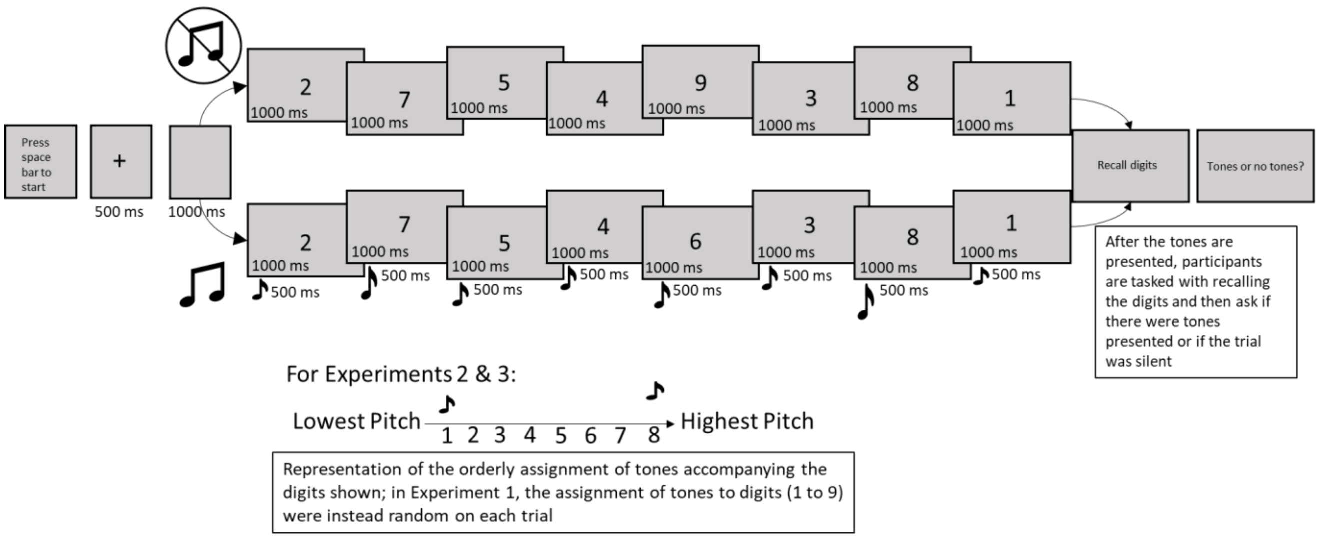
Table 1

Tone series that were presented in the changing-tones condition

	Serial Position in Series							.
Tone Series	1	2	3	4	5	6	7	8
1	A4	B3	B4	E4	C4	D4	G4	F4
2	B3	E4	A4	D4	B4	F4	C4	G4
3	E4	D4	B3	F4	A4	G4	B4	C4
4	D4	F4	E4	G4	B3	C4	A4	B4
5	F4	G4	D4	C4	E4	B4	B3	A4
6	G4	C4	F4	B4	D4	A4	E4	B3
7	C4	B4	G4	A4	F4	B3	D4	E4
8	B4	A4	C4	B3	G4	E4	F4	D4

Figure 1

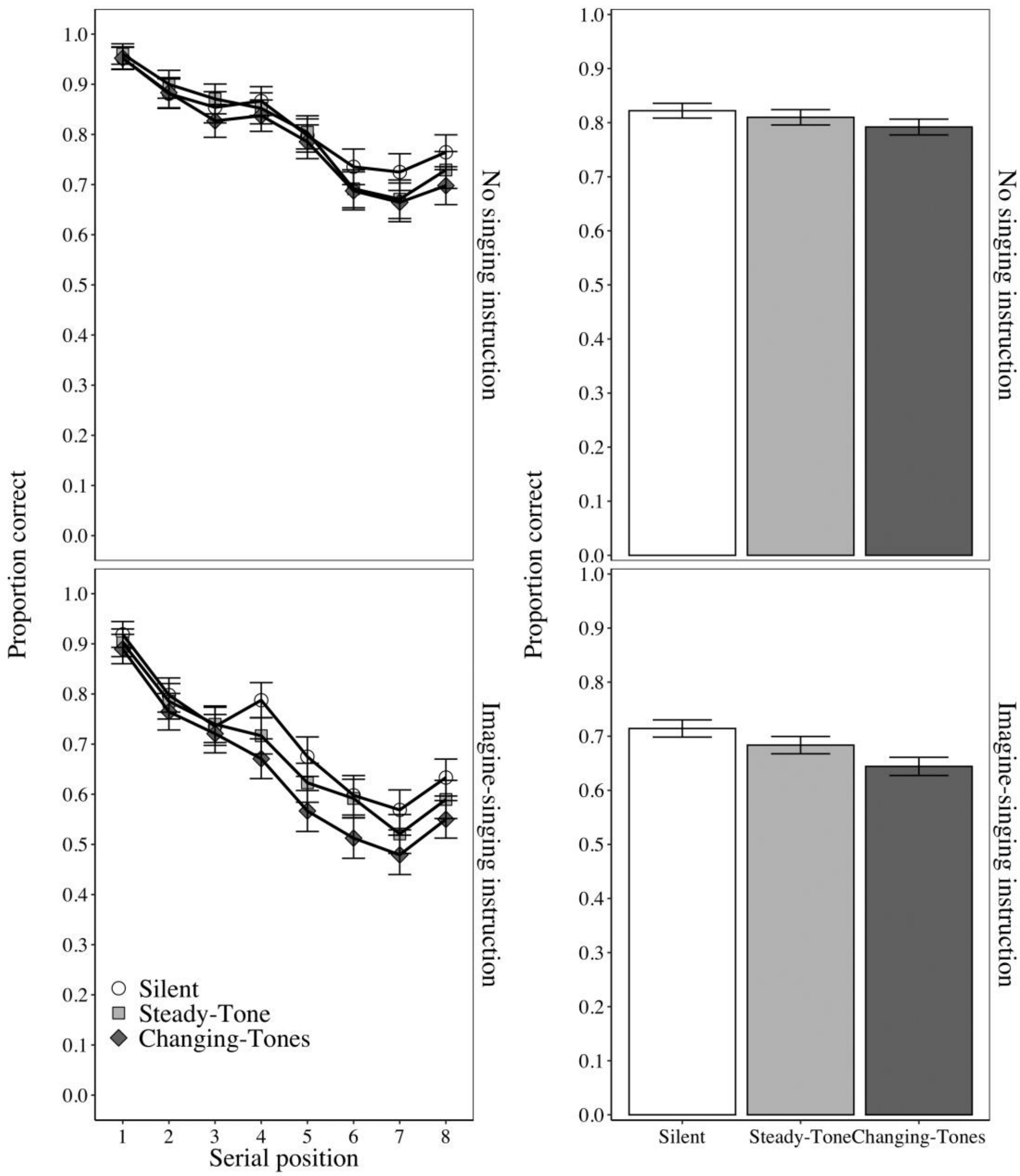
Illustration of the procedure for the silent and changing-tone conditions in three experiments



Note. There was also a steady-tone condition with the same tone for every note in a trial, in Experiment 1 only (not shown). Half of the participants in each experiment were instructed to imagine singing the digits in the presented tones.

Figure 2

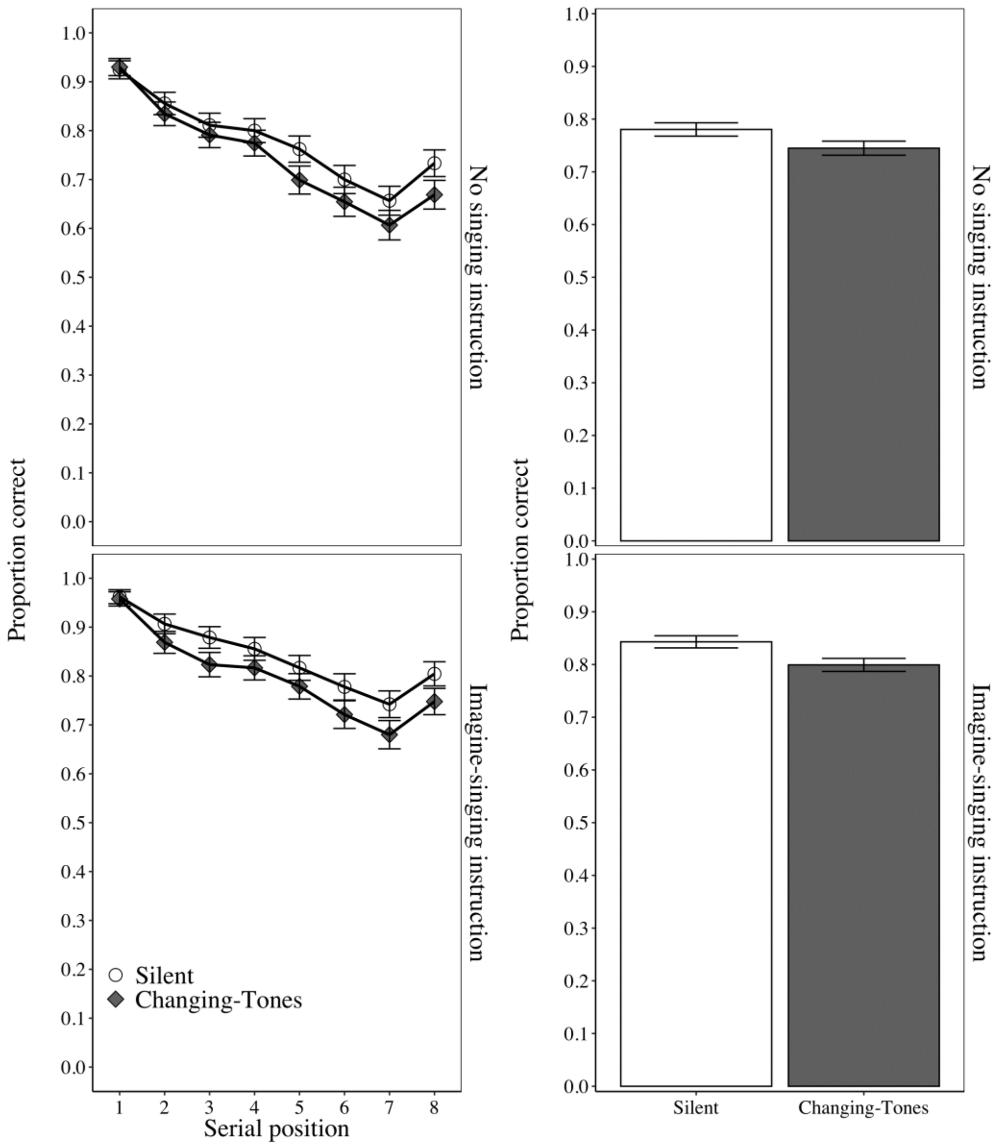
*Proportion of correct response as a function of tone condition (*silent, steady tone, changing tones*) and whether or not there were imagine-singing instruction (rows) in Experiment 1.*



Note. **Left column**, serial positions (1-8); **right column**, results averaged across serial positions. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.

Figure 3

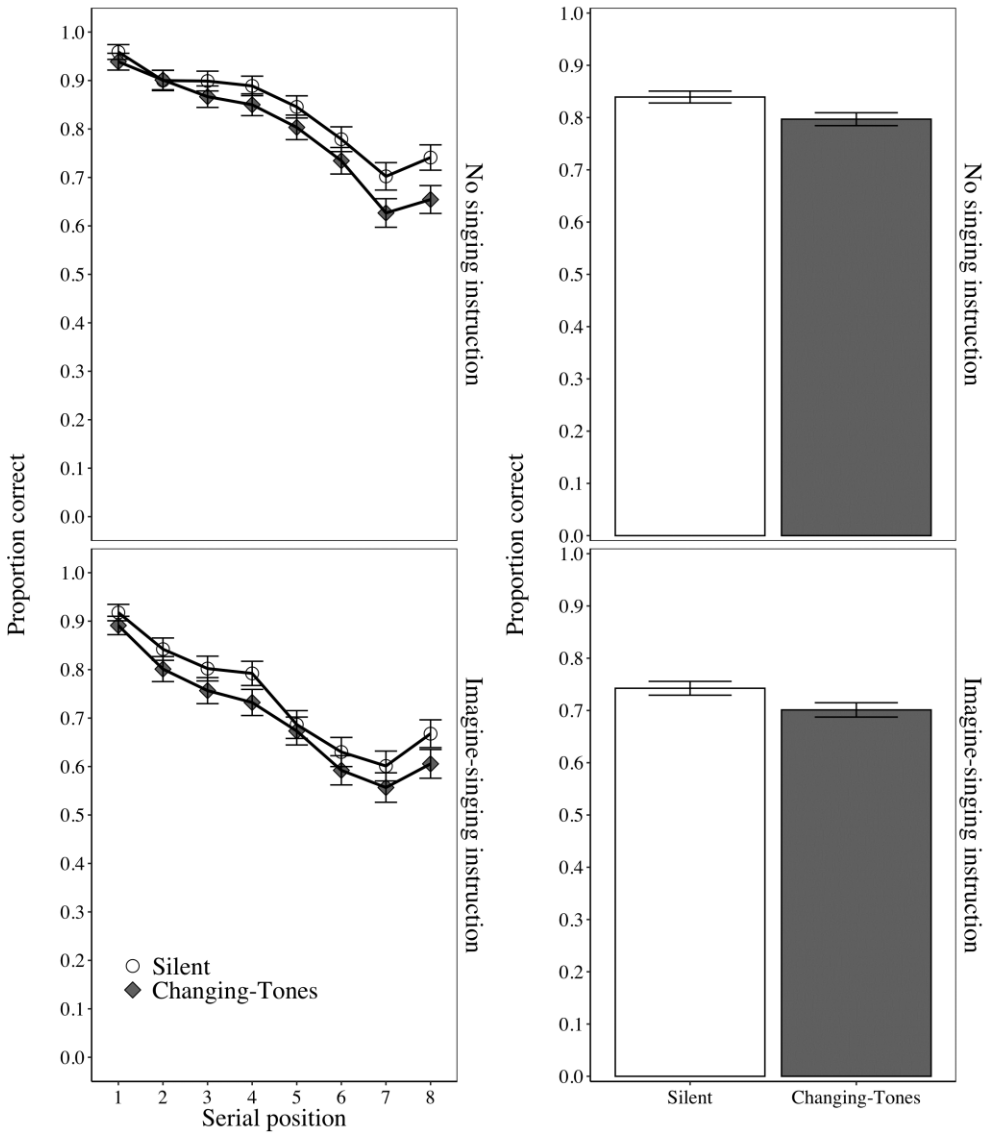
*Proportion of correct response as a function of tone condition (*silent, changing tones*) and whether or not there were imagine-singing instruction (rows) in Experiment 2.*



Note. **Left column**, serial positions (1-8); **right column**, results averaged across serial positions. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.

Figure 4

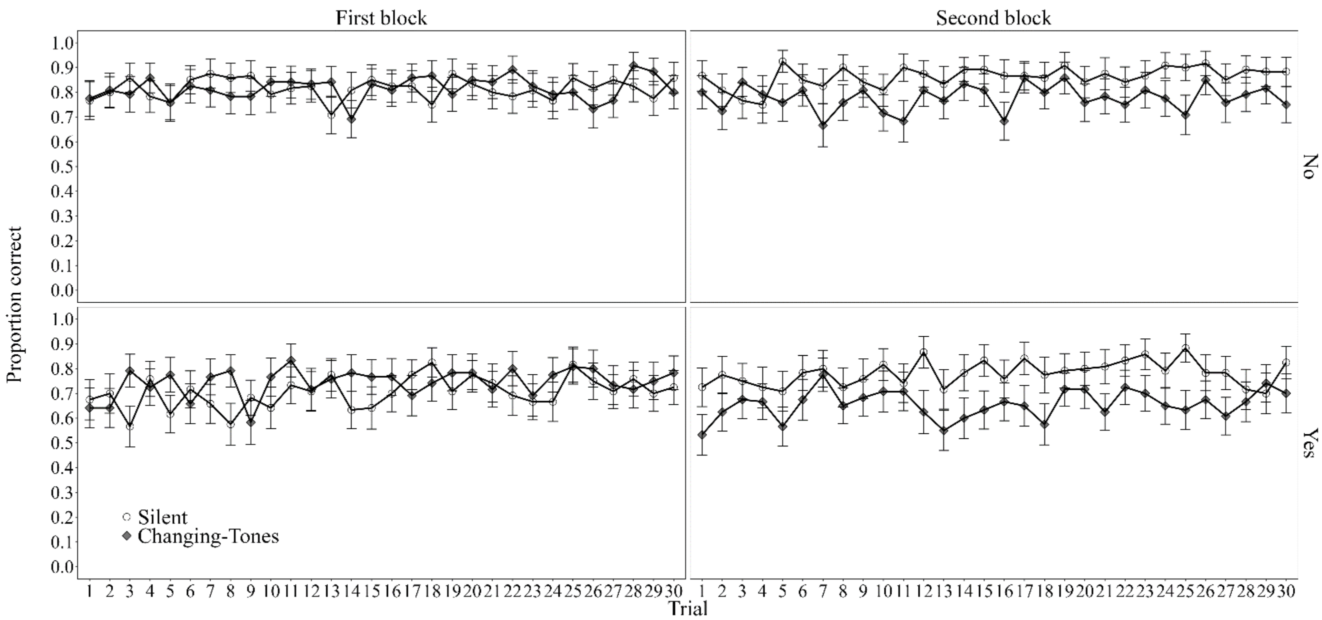
*Proportion of correct response as a function of tone condition (*silent, changing tones*) and whether or not there were imagine-singing instruction (rows) in Experiment 3.*



Note. **Left columns**, serial positions (1-8); **right columns**, results averaged across serial positions. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.

Figure 5

Proportion of correct response as a function of melody condition (silent, changing tones), block (first block, second block), trial (1 to 30) and whether or not there were imagine-singing instruction (no=no singing instruction, yes=imagine-singing instruction) in Experiment 3



Note. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.