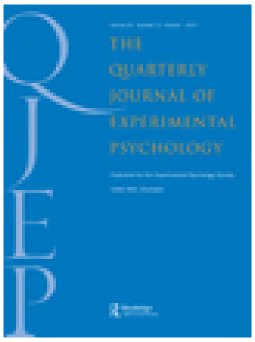

Routledge
Taylor & Francis Group

The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Resilient memory for melodies: The number of intervening melodies does not influence novel melody recognition

Steffen A. Herff, Kirk N. Olsen & Roger T. Dean

To cite this article: Steffen A. Herff, Kirk N. Olsen & Roger T. Dean (2017): Resilient memory for melodies: The number of intervening melodies does not influence novel melody recognition, The Quarterly Journal of Experimental Psychology, DOI: 10.1080/17470218.2017.1318932

To link to this article: http://dx.doi.org/10.1080/17470218.2017.1318932

View supplementary material

Accepted author version posted online: 12 Apr 2017.

Submit your article to this journal

Article views: 1

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=pqje20

Download by: [Macquarie University Library] **Date:** 15 April 2017, At: 04:42

Publisher: Taylor & Francis & The Experimental Psychology Society
Journal: *The Quarterly Journal of Experimental Psychology*
DOI: 10.1080/17470218.2017.1318932

Resilient memory for melodies:

The number of intervening melodies does not influence novel melody recognition

Steffen A. Herff[1], Kirk N. Olsen[1,2], Roger T. Dean[1]

[1]*The MARCS Institute for Brain, Behaviour and Development,*

Western Sydney University, Australia

[2]*Department of Psychology, Macquarie University, Australia*

Running Head: Memory for Melody

Corresponding Author:

Steffen A. Herff

The MARCS Institute for Brain, Behaviour and Development,

Western Sydney University, Australia

Post: Locked Bag 1797

Penrith, NSW 2751

Phone: +61 2 9772 6801

ACCEPTED MANUSCRIPT

Fax: +61 2 9772 6040

E-mail: s.herff@westernsydney.edu.au

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

In many memory domains, a decrease in recognition performance between the first and second presentation of an object is observed as the number of intervening items increases. However, this effect is not universal. Within the auditory domain, this form of interference has been demonstrated in word and single-note recognition, but has yet to be substantiated using relatively complex musical material such as a melody. Indeed, it is becoming clear that music shows intriguing properties when it comes to memory. This study investigated how the number of intervening items influences memory for melodies. In Experiments 1, 2, and 3, one melody was presented per trial in a continuous recognition paradigm. After each melody, participants indicated whether they had heard the melody in the experiment before by responding ‘old’ or ‘new’. In Experiment 4, participants rated perceived familiarity for every melody without being told that melodies reoccur. In four experiments using two corpora of music, two different memory tasks, transposed and untransposed melodies, and up to 195 intervening melodies, no sign of a disruptive effect from the number of intervening melodies beyond the first was observed. We propose a new ‘regenerative multiple representations’ conjecture to explain why intervening items increase interference in recognition memory for most domains but not music. This conjecture makes several testable predictions and has the potential to strengthen our understanding of domain specificity in human memory, while moving one step closer to explaining the ‘paradox’ that is memory for melody.

Keywords: Memory, Interference, Music Perception, Recognition, Familiarity.

Fast and accurate recognition of stimuli such as faces, names, smells, phone numbers, street signs, book titles, animals, food, and music relies on an important interplay between perception and memory. The demand of this difficult day-to-day task is exacerbated when we are faced with numerous items of similar nature. In human memory, cumulative interference from an increase in the number of intervening items impairs recognition performance in a variety of domains (Sadeh, Ozubko, Winocur, & Moscovitch, 2014). For example, in written word recognition, a systematic decrease in recognition performance is observed as the number of intervening items increases (Poon & Fozard, 1980). Digit tasks (Donaldson & Murdock, 1968), letter trigrams (Olson, 1969), word lists and pairs (Bui, Maddox, Zou, & Hale, 2014; Hockley, 1992), faces (Rakover & Cahlon, 2001), and a wide variety of everyday visual objects (Konkle, Brady, Alvarez, & Oliva, 2010; Nickerson, 1965) have been thoroughly investigated with results that report similar behavioural phenomena.

However, evidence suggests that the detrimental effect of intervening items on recognition performance is not universal. In the visual domain, there is a large decrease in word recognition accuracy as the number of intervening words between the first and second presentation of a target word increases from 2 to 32 (Friedman, 1990b). Interestingly, this effect is not observed with drawings, where recognition performance remains constant between 2, 8, and 32 intervening items (Berman, Friedman, & Cramer, 1991; Friedman, 1990a, 1990b). In the auditory domain, a disruptive effect of intervening items has been demonstrated in recognition tasks using spoken words and single-notes (Campeanu, Craik, Backer, & Alain, 2014; D. Deutsch, 1970, 1975). In single-note recognition, interference arises from intervening pitches but not intervening spoken numbers (D. Deutsch, 1970). This suggests strong stimulus-specific interference effects in single-note recognition. However, cumulative disruptive effects have yet

ACCEPTED MANUSCRIPT

to be substantiated in other real-world auditory contexts, such as multi-note melody recognition in the musical domain. Indeed, as detailed below, it has been suggested that memory for melodies may be resilient to interference from intervening items (Dowling, 1991; Dowling, Kwak, & Andrews, 1995). Thus a systematic investigation of the possible interference effects in music and melodies is important to further elucidate memory's domain specificity (Fougnie, Zughni, Godwin, & Marois, 2015) and the question of whether memory for music is 'special' (Jackendoff & Lerdahl, 2006; Schulkind, 2009; Stevens, 2015). This study conducted such an investigation.

Memory for Melody

Melodies are one of the most ubiquitous aspects of music and is crucial for musical enculturation (Corrigall & Trainor, 2014). We are exposed to many melodies on a day-to-day basis (Krause, North, & Hewitt, 2014) and often many new melodies are heard before we encounter an old one again. Listeners who have experienced involuntary music imagery (e.g., earworms) know that simple melodies have great potential for memorability (Bailes, 2007, 2015; Halpern & Bartlett, 2011; Williams, 2015). However, memory for melody has been described as a 'paradox' (Halpern & Bartlett, 2010; Schulkind, 2009; Stevens, 2015). It can be long lasting (Bailes, 2007; Baird & Samson, 2014; Cuddy & Duffin, 2005; Halpern & Bartlett, 2011; Jacobsen et al., 2015; Vanstone, Cuddy, Duffin, & Alexander, 2009), yet "memory for novel melodies is surprisingly poor" (Lange & Czernochowski, 2013, p. 137) and "(…) even the simplest kind of recognition test for melodies shows how poor musical memory can be, in comparison to other kinds of memory" (Halpern & Bartlett, 2010, p. 234). These findings raise important questions about the nature of fundamental memory phenomena in memory for melody.

ACCEPTED MANUSCRIPT

Here, we systematically investigated two fundamental memory phenomena in memory for melody: recency-in-memory and cumulative disruptive effects from the number of intervening melodies (both discussed in detail below).

Interference and Temporal Delay in Novel Melody Recognition

A potential source of confusion in the literature on memory for music is that different forms of memory – and the conditions under which memory is measured – are often not well articulated. This means that apparent contradictory evidence may occur when in fact different results arise because researchers evaluate different forms of memory under different experimental conditions. In the context of memory for melody in particular, apparent contradictory evidence such as the claims reviewed above, that memory for melody is sometimes ‘surprisingly poor‘ yet sometimes ‘surprisingly good’, provides one such example. To avoid this potential source of confusion in the present study, it is important to note that we specifically investigated continuous recognition of novel monophonic melodies that resemble melodies that are encountered in normal day-to-day experiences.

Recently, Schellenberg and Habashi (2015) showed that recognition performance of novel ~30 s melodies is not disrupted by the mere passing of time over periods up to a week. However, their study focused on the passing of time and not on whether the number of intervening items influenced melody recognition. This is important because memory decay over time is only one of two main mechanisms of forgetting (Eysenck & Keane, 2015). Interference is the other mechanism, which describes a continuous decrease in memory performance, not because of the passing of time, but from the additional information that is learned after or even before the encoding of a stimulus into memory. Establishing the extent of interference as well as

ACCEPTED MANUSCRIPT

effects of decay in memory for melody provides an important step towards a full account of memory for melody in particular, and memory for complex auditory information in general. This will, in turn, assist in a comprehensive understanding of memory phenomena observed in the context of music; for example, the observation that memory for music is often ‘spared’ in people who suffer from debilitating deficits in memory, such as those associated with dementia and severe brain injury (Baird & Samson, 2014; Cuddy & Duffin, 2005; Jacobsen et al., 2015; Schulkind, 2009).

In the context of interference and decay, a series of melody recognition experiments conducted by Dowling and colleagues (Dowling, 1991; Dowling et al., 1995) used very short (~ 3 s) melodies that recurred after periods of silence or periods filled with other melodies. Results indicated that melody recognition predominantly relied on pitch-interval information when delays are filled with other intervening melodies. Pitch-interval information refers to the relative pitch distance between notes, rather than their absolute pitches (i.e., pitch intervals like “a major third” rather than absolute pitches like ‘C4’ and ‘E4’). After observing relatively stable memory performance over the first two minutes, Dowling, et al. (1995) proposed that there might be underlying processes involved in pitch-interval based recognition of novel melodies that may be resilient to the presentation of intervening items (in this case, melodies). Such resilience to the effects of interference in memory for melodies was investigated here by systematically manipulating the number of intervening melodies in four melody recognition experiments.

Pitch Information and Melody Transposition

As mentioned above, several kinds of pitch information are important for accurate melody recognition: surface information such as absolute pitch and abstract information of pitch

ACCEPTED MANUSCRIPT

relations expressed, for example, as pitch intervals or pitch contour (Bartlett & Dowling, 1980; Dowling & Fujitani, 1971; Krumhansl, 2000; Levitin, 1994; Schellenberg, Stalinski, & Marks, 2014). Transposition of a melody into another key only retains relative pitch information. While participants are capable of recognizing untransposed melodies, recognition of transposed melodies tends to be worse (Dowling & Fujitani, 1971; Plantinga & Trainor, 2005; Schellenberg et al., 2014). Some literature suggests that for delays beyond one minute, memory for melody predominantly uses relative pitch information (Bachem, 1954) (see also, Krumhansl, 2000, for a review). However, this does not necessarily imply that memory is resilient to transposition. Indeed, more recent research shows that transposing melodies also disrupts memory for melody after longer delays of up to a day (Schellenberg & Habashi, 2015).

Taken together, these findings suggest that interference in melody recognition might depend on whether or not both absolute or relative frequency information is retained. This hypothesis is investigated in the present study. Specifically, Experiments 1 and 2 investigated melody recognition with the second occurrence of each melody untransposed, whereas in Experiments 3 and 4 the second occurrence of each melody was transposed.

Task Awareness in Melody Recognition

Participants' awareness of the type of musical memory task nature of a musical memory task can potentially influence performance (Halpern & Bartlett, 2010). For example, age-related decline in memory appears to be smaller with indirect (Fleischman, Wilson, Gabrieli, Bienias, & Bennett, 2004) rather than explicit memory tasks (Gaudreau & Peretz, 1999; Halpern & O'Connor, 2000). Therefore, the four experiments reported here used two different continuous memory tasks with varying degrees of task-specific awareness (Shepard & Teghtsoonian, 1961).

ACCEPTED MANUSCRIPT

Continuous memory tasks are commonly used to investigate recognition performance and the disruptive effects of intervening items (Berman et al., 1991; Campeanu et al., 2014; Ferris, Crook, Clark, Mccarthy, & Rae, 1980; Friedman, 1990b; Hockley, 1992; Sadeh et al., 2014). In such a task, stimuli are presented successively and participants are asked to respond after each item; most commonly, whether they perceive the present stimulus to be ‘old’ or ‘new’ (Shepard & Teghtsoonian, 1961). In this case, ‘old’ refers to items that have been presented before, and vice versa. Importantly, the memory-related aspects of a continuous memory task can be expressed with varying degrees of memory-task awareness. For example, the ‘old’ or ‘new’ request is an explicit memory task, while rating the familiarity of each melody is a less explicit memory task. Therefore, in the present study, Experiments 1, 2, and 3 implemented explicit memory task instructions, whereas Experiment 4 implemented less explicit memory task instructions.

Recency-In-Memory

In a continuous memory task, the case of zero intervening melodies is the only condition that reflects immediate repetition. Due to recency-in-memory effects, we expected that immediate repetition of a melody should lead to higher memory performance when compared to any case that has at least one intervening item (Berz, 1995; Dowling, 1973; Greene & Samuel, 1986; Jahnke, 1963; Roberts, 1986). Indeed, recognition performance should be lower for all other numbers of intervening items, as the first intervening melody will disrupt any recency-in-memory effect. No effect of the number of intervening melodies beyond the first would also support the findings of Schellenberg and Habashi (2015), who reported no significant disruptive effect of temporal delay on melody recognition. Here, Experiment 1 specifically investigated a

ACCEPTED MANUSCRIPT

small number of intervening items (up to 6) to compare the zero intervening melody condition to each of the other intervening melody conditions. Experiments 2, 3, and 4 presented a far larger number (up to 197) of intervening melodies to investigate longer-term unfolding cumulative disruptive effects on recognition. In the following section, we provide a summary of the aim, design, and hypothesis of the study as a whole.

Aim, Design, and Hypotheses

Four experiments were designed to further elucidate mechanisms that underpin memory for novel melodies by investigating the effect of distinctively different intervening melodic items (details in the stimulus section). The importance of absolute and relative frequency information was also investigated by implementing melody transposition, and the potential for differential outcomes between memory tasks was investigated by measuring both recognition and perceived familiarity.

Following the literature reviewed above, a significant performance decrease as the number of intervening items increases is evidence that memory for melody is similar to other memory domains. A lack of such an effect supports Dowling, et al.’s (1995) suggestion that memory for melody might be special because other intervening melodies do not disrupt it. Furthermore, such results would also provide evidence of the domain specificity of human memory. Finally, immediate repetition of a melody was hypothesized to lead to better performance than when a melody’s repetition occurs after any number of different intervening melodies.

Experiment 1: Melody Recognition with Zero to Six Intervening Melodies

ACCEPTED MANUSCRIPT

As a starting point, Experiment 1 was designed to investigate how the first six intervening melodies affect melody recognition performance. Specifically, the number of intervening items was varied between zero and six in a continuous recognition paradigm while recognition performance was measured. The choice of six intervening item conditions was deemed a sufficient starting point to demonstrate cumulative disruptive effects, as has been the case in other domains such as written and spoken word recognition (Campeanu et al., 2014; Friedman, 1990b).

Method

Participants. Twenty-eight undergraduate students were recruited from the Western Sydney University (M_{age} = 22.1 years, SD_{age} = 8.5, seven male). Recruitment criteria were less than two years of musical training and no hearing impairments. Participants volunteered for this experiment and were reimbursed with chocolate and the opportunity to learn about the research after the experiment.

Stimuli and Equipment. Testing took place in a sound attenuated booth at the MARCS Institute for Brain, Behaviour, and Development, Sydney, Australia. The experiment was programmed in E-Prime (Psychology Software Tools, 2012). Stimuli were presented through Sennheiser 25 HD headphones at a volume comfortable to the user. Experiment 1 focused on recognition of novel melodies, and 60 novel monophonic melodies were composed by the first author of this article. All melodies were 12 seconds in duration and unmistakably tonal. The melodies were composed in 12-tone equal temperament, the most common tuning system used in Western tonal music (Milne, Sethares, & Plamondon, 2007). The key for each melody was

ACCEPTED MANUSCRIPT

randomly chosen before composition; half of the melodies were composed in major and the other half in minor. All notes were sounded with the same grand piano timbre at the same velocity using Pianoteq 4 STAGE (Version 4.2). The meter was balanced across the melodies between 4/4 and 3/4, the two dominant meters in Western tonal music (London, 2012). The tempi were pseudo randomized between 80-165 beats per minute (bpm) with a mean bpm set to the most common 120 bpm (Franek, van Noorden, & Rezny, 2014; Moelants, 2002). Not all tempi between 80 and 165 bpm were possible to realize, as the meter was fixed at 4/4 and 3/4 and the duration fixed at 12 seconds. The rhythmic structure was kept simple with not more than two levels of metrical division (Winold, 1975). Musical scores relating to representative examples of melodies presented in Experiment 1 can be found in Appendix A. The online supplement S1-Stimuli.zip contains the stimuli of all Experiments as well as a musical feature analysis of all melodies (more detail is provided in the Stimuli and Material section of Experiment 3).

Even though all melodies were novel, there was a slight possibility that some melodies might have resembled familiar tunes. Therefore, a pilot study was conducted and every melody that was perceived to sound similar to another known melody by at least one participant was removed from the corpus. Twelve researchers from the MARCS institute who were not involved in this project volunteered to participate in the pilot study. All participants were oblivious to the origin of the stimuli. A total of five melodies were removed from the initial pool of 60. Representative examples of the remaining stimuli can be seen in Appendix A. An uninvolved expert listener with an extensive and sophisticated background in music (*Ollen Musical Sophistication Index* of 845, see Ollen, 2006, where on a scale of 0-1000, >500 is deemed to be musically sophisticated) described the melodies as follows:

ACCEPTED MANUSCRIPT

> *“(…) I guessed they were theme tunes from TV programs, film music, or adverts. They sounded like the sort of melodies one would typically come across in everyday life.”*

Procedure. Up to three participants could participate in the experiment simultaneously, but each participant could neither hear the stimuli presented to others, nor see the other participants as they completed the task. Participants provided informed consent and sat comfortably in front of a computer. All instructions were presented on the computer screen. Participants were instructed that “*in this experiment, you will hear many different melodies, one after another. However, sometimes a melody will be repeated. Each melody may repeat more than once OR may not repeat at all*.” They were asked to “*listen to each melody and respond to whether you have heard this melody before in this experiment*.” Responses were made using two different keys of the keyboard (the ‘A’ key and the ‘-’ key), one associated with ‘*New*’ and the other with ‘*Old*’. The response keys were counterbalanced between participants. While each melody was played, the screen showed ‘Listen!’ in black letters on a white background. As soon as a melody finished, the ‘Listen!’ text disappeared and the participant made their response. Therefore, one melody and one response constituted one trial. The next trial was initiated as soon as a participant gave a response. Participants had the opportunity to practice the task in six practice trials and were allowed to adjust the volume to their personal preference during the practice trials. However, the volume was then fixed to their chosen level for the main experiment trials. After completion of the experiment, participants were asked to fill out a short demographic questionnaire. A schematic example of a series of trials is shown in Figure 1.

ACCEPTED MANUSCRIPT

Figure 1 about here

The number of intervening melodies between the first and second presentation of an identical melody was manipulated between zero and six. To avoid list order effects, the order of the melodies was randomized for every participant. In a continuous recognition task this means that the two melody presentations of a condition with few intervening items can be ‘embedded’ in between two melody presentations of a condition with more intervening items. Furthermore, it is possible that the first and second presentation of one melody ‘overlap’ with the first and second presentation of another melody. When melody presentations are embedded, then a larger intervening-items condition will comprise first and second melody presentations of a smaller intervening-item condition. When melody presentations overlap, much like the links in a chain, the intervening items between the two presentations of a target melody will comprise the first, but not the second presentation of another target melody. For example, if A1 and B1 are the first presentations of melodies A and B, and A2 and B2 are the second presentations of melodies A and B, then both scenarios are possible: A1B1A2B2 (overlap), A1B1B2A2 (embedded). Since the number of intervening melodies was manipulated from zero to six with the same number of trials in each gap size, the possible permutations of list order were highly constrained. As a result, 49 of the 55 melodies were presented twice over the course of the experiment with controlled numbers of intervening items. The remaining six melodies were used as ‘dummy’ melodies. Dummy melodies randomly filled the remaining item list slots and were not included in the analyses. Overall, every participant listened to 130 trials that included seven melodies in each of the different numbers of intervening melody conditions (0, 1, 2, 3, 4, 5, 6), and 32

ACCEPTED MANUSCRIPT

dummy trials that were not included in the analysis but enabled a fully randomized list order for each participant.

Accounting for Participant Response Tendencies. Recognition paradigms are prone to effects of response biases (Snodgrass & Corwin, 1988). For example, there is evidence that response tendencies can change over the course of an experiment (Berch, 1976; Donaldson & Murdock, 1968). In order to take this into account, we built participant-wise *Dynamic Response Tendency* models that predict the baseline tendency for each participant to press ‘old’, and how this tendency changes throughout the experiment. These generalized linear mixed effects models were trained on ‘old’ responses on first melody presentations based on trial number. The fitted models were then used to predict the probability of pressing ‘old’ on melody repetition trials based solely on trial number. As a result, the Dynamic Response Tendency models were used in statistical analyses in each Results section to account for the individual response tendencies of each participant, and the way such tendencies may change over the course of the experiment.[1]

Statistical Approach. The statistical approach is similar in all experiments and will be detailed here. We used generalized linear mixed effects models to investigate the effect of the number of intervening melodies on binary melody recognition data (Experiments 1, 2, and 3) and linear mixed effects models for continuous perceived familiarity data (Experiment 4) (Baayen, Davidson, & Bates, 2008). This approach takes into account crossed random effects of subjects

---

[1] Detailed assessment showed that there were no response tendency shifts in Experiments 1, 2, and 4 (all dynamic response tendency coefficient *p*-values > .20). Experiment 3 showed dynamic response tendency shifts over the 196 trials ($p < .001$). The model controlled for the shifts. The nature of the response tendency shifts in Experiment 3 is detailed in Appendix C.

ACCEPTED MANUSCRIPT

and melodies that possess different levels of memorability (Baayen, 2008; Baayen et al., 2008; Judd, Westfall, & Kenny, 2012; Kass & Raftery, 1995; Kruschke, 2010, 2013; Nathoo & Masson, 2016). The models were implemented in the R software platform (R-Core-Team, 2013) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2013).

A model comparison approach was used to compare evidence in favour of the null hypothesis relative to evidence in favour of a competing hypothesis (Kruschke, 2011). Bayesian information criteria (BIC) (Schwarz, 1978) are reported. BIC values, which penalize additional parameters strongly, were used as the basis of model selection. Differences in BIC values between models are reported as ΔBIC for significant model improvements. A ΔBIC of two or greater is assumed “positive” in favour of the model with lower BIC. A ΔBIC difference of six or greater is considered “strong” evidence (Kass & Raftery, 1995). ΔBIC can be used to estimate the Bayes factor, a measurement of how much evidence there is supporting one hypothesis or model compared to another (Kass & Raftery, 1995; Lewandowsky & Farrell, 2010, p. 186; Nathoo & Masson, 2016; Wagenmakers, 2007). A ΔBIC of six represents a Bayes factor of twenty, which can be interpreted as twenty times more evidence for the model with the lower BIC. A Bayes factor cannot only provide evidence against the Null-Hypothesis, but also evidence for it. Furthermore, direct model comparisons using goodness-of-fit were conducted with likelihood-ratio tests, with *p*-values reported throughout (Wilks, 1938). The models were provided with a random intercept for *Melody* in order to account for possible effects of individual melodies. Accounting for participant response tendencies was achieved in each model by implementing random intercepts for *Subject* as well as a fixed factor for the aforementioned *Dynamic Response Tendency* models.

ACCEPTED MANUSCRIPT

At the beginning of each results section, mixed effects models were used to assess whether overall performance is above chance. This was achieved by testing if melody repetition predicts significantly more 'old' responses (Experiments 1-3) or higher perceived familiarity (Experiment 4) while taking random participant response biases and melody variation into account. Coefficient *p*-values are reported in the beginning of each results section. All figures that depict recognition performances in each Results section below report response bias-corrected hit rates by subtracting the participant-wise false alarm rates from participant-wise hit rates. As a result of this transformation, performance around zero indicates the inability to recognize the melodies (Snodgrass & Corwin, 1988).

At the end of each Results section for each experiment, a final assessment of the effects of the number of intervening melodies is reported using models with maximal random effects (Barr, Levy, Scheepers, & Tily, 2013). These models include random slopes for participants over the number of intervening items; that is, all random effects that could possibly play a role given the experimental design. For generalized linear models, *p*-values as calculated by the lme4 package are reported. For linear models, conservative Kenward-Roger approximations were used to adjust the degrees of freedom (Kenward & Roger, 1997) using the R package pbkrtest (Halekoh & Højsgaard, 2014a, 2014b).

Results

Figure 2 shows mean false alarm rates and hit rates for every participant and melody, and bias-corrected hit rates are depicted in Figure 3. In summary, participants performed significantly above chance ($Z = 24.387$, $p < .0001$) and the number of intervening melodies did not influence recognition performance beyond the first intervening item.

ACCEPTED MANUSCRIPT

Figure 2 about here

Specifically, a generalized linear mixed effects model was constructed to investigate how the number of intervening items influences melody recognition. The model is built to predict ‘old’ responses on melody repetitions. The base model included the systematic factor for *Dynamic Response Tendency* and random intercepts for *Subject* and *Melody* (*BIC* = 1659.7; *LogLik* = -815.38), and improved significantly when provided with information about the *Number of Intervening Items* (*BIC* = 1655.8; *LogLik* = -791.75, p < .0001, ΔBIC = 3.9). Coefficient estimation for each number of intervening melodies confirmed a significantly lower recognition performance for all numbers of intervening items beyond immediate repetition (all *p*-values < .0001). A base model excluding immediate repetitions (*BIC* = 1442.7; *LogLik* = -707.17) did not improve when provided with the number of intervening items (*BIC* = 1474.7; *LogLik* = -705.48, *p* > .64). This shows that the number of intervening items beyond the first and up to six does not carry predictive value when it comes to melody recognition performance. This result is illustrated in Figure 3.

Figure 3 about here

ACCEPTED MANUSCRIPT

A final statistical assessment of the effect of number of intervening items was made using the maximal random effect structure. The number of intervening melodies did not improve models that excluded data relating to immediate repetition (*p* > .95). As hypothesized, significant differences between melodies were observed, as shown by a significant decrease in model performance when the random intercept for melody was removed (*p* < .00001, ΔBIC = 68.1). Figure B.1 in Appendix B displays uncorrected hit-rates for each of the different intervening item conditions.

Discussion

Experiment 1 showed with up to 6 intervening melodies that beyond the effect of the first intervening melody, additional intervening melodies did not significantly decrease recognition performance. This finding cannot be explained by a floor effect that would prevent further decreases in performance after the first intervening melody, as recognition performance was above chance in all six intervening melody conditions (see Figure 2). The results are consistent with Schellenberg and Habashi’s (2015) findings of a lack of disruptive effects of intervening time on melody recognition. Using relatively long (12 s) melodies, Experiment 1 also supports the hypothesis of underlying processes involved in melody recognition that bypass the interference by intervening items found in other memory domains (Dowling et al., 1995).

Recent research using five four-part chord progressions followed by three-note arpeggiated continuations has shown that it can take approximately 20 s to reappraise a prior melody as a whole; that is, 20 s to integrate a melody into long term memory (Bailes, Dean, & Pearce, 2013). Buchsbaum, Padmanabhan, and Berman (2011) investigated this issue in auditory-verbal stimuli by combining a continuous recognition paradigm and fMRI. They found

ACCEPTED MANUSCRIPT

systematically different activation patterns in response to conditions that presented ≥ 4 intervening items. So it could be that the disruptive effect of the number of intervening melodies only manifests after the critical period of 20-30 s. In this case, the few conditions in Experiment 1 (4, 5, and 6 intervening melodies) that extended up to and beyond this critical period may not be sufficient to properly rule out this possibility. Therefore, there may still be an interference effect from the number of intervening melodies on melody recognition that occurs beyond those used in Experiment 1. This possibility was investigated in Experiment 2.

Experiment 2: Melody Recognition with Four to Thirteen Intervening Melodies

Experiment 2 aimed to replicate and extend the results of Experiment 1 with a larger number of intervening items. Experiment 2 systematically investigated four to thirteen intervening melodies. This ensured that all number of intervening melodies were presented beyond the potentially critical period of 20 to 30 seconds after stimulus presentation. Following from the results of Experiment 1, no differences in recognition performance were hypothesized for any number of intervening melodies.

Method

Participants. Experiment 2 tested 20 participants from the Western Sydney University (M_{age} = 21.1 SD_{age} = 3.6, five male) who provided informed consent and were reimbursed with chocolate. The data of three participants were excluded because two reported a high level of musical expertise (active musicians), and one made the same response (identical key press) in each trial. All participants reported normal hearing and did not participate in Experiment 1.

ACCEPTED MANUSCRIPT

Stimuli and Equipment. Stimuli and equipment were identical to Experiment 1; the only difference was a reduction from 55 to 50 randomly chosen melodies from the set of 60 melodies described in Experiment 1 in order to decrease the experiment's duration. Musical scores relating to representative examples of melodies presented in Experiment 2 can be found in Appendix A. The online supplement S1-Stimuli.zip contains the stimuli of all Experiments as well as a musical feature analysis of all melodies.

Procedure. The procedure of Experiment 2 closely followed that of Experiment 1. However, instead of zero to six intervening melodies, Experiment 2 presented four to thirteen intervening melodies. For each participant, a list of 100 trials was randomly populated with melodies that reoccurred once after four to thirteen intervening melodies. Each number of intervening melodies occurred at least four times for each participant. This ensured at least 80 analyzable trials with 40 different melodies for each participant. The remaining 20 trials were filled without restraints on the number of intervening melodies in order to allow sufficient permutations of list order. Similar to the procedure in Experiment 1, dummy melodies randomly filled any remaining item list slots and were not included in the analyses. Instructions were identical to Experiment 1.

Results

Figure 4 shows mean false alarm rates and hit rates for every participant and melody, and bias-corrected hit rates are depicted in Figure 5. Again, participants performed significantly above chance ($Z = 15.97$, $p < .0001$), with results that replicated and extended those reported in

ACCEPTED MANUSCRIPT

Experiment 1. The number of intervening melodies did not influence recognition performance when up to thirteen intervening melodies are presented.

Figure 4 and Figure 5 about here

In statistical terms, a mixed effects model predicting 'old' responses on repetition trials (*BIC* = 929.83; *LogLik* = -451.67) with a random intercept of *Subject* and *Melody* and a systematic factor for *Dynamic Response Tendency* did not improve when provided with the information of the *Number of Intervening Melodies* (*BIC* = 936.13; *LogLik* = -451.51; *p* > .56). This shows that the number of intervening items cannot be used to predict recognition performance in memory for melody between 4 and 13 intervening melodies. This effect is also illustrated in Figure 5.

A final statistical assessment of the effect of number of intervening items was made using the maximal random effect structure. *Number of Intervening Melodies* did not yield a significant result (*p* > .50). As hypothesized, the model performed significantly worse without a random intercept for *Melody (p* < .00001, ΔBIC = 47.72*)*. This result is consistent with Experiment 1.

Discussion

Experiment 2 extends the results of Experiment 1, with no significant differences observed when the number of intervening melodies increased to thirteen. Thirteen intervening melodies equates to a temporal delay of about two and a half minutes. Thus, no differences in recognition performance between four and thirteen intervening melodies indicate that the passing

ACCEPTED MANUSCRIPT

of time again has no effect on melody recognition performance within the first few minutes of melody recognition. This is consistent with previous research in the musical domain (Schellenberg & Habashi, 2015). However, this result is somewhat surprising from non-musical memory research, because a variety of domains do show interference effects of time (Bui et al., 2014; Campeanu et al., 2014; Hockley, 1992; Konkle et al., 2010; Olson, 1969; Sadeh et al., 2014) that include other auditory stimuli such as words (Buchsbaum et al., 2011; Campeanu et al., 2014).

The second presentation of each melody in Experiment 1 and Experiment 2 was physically identical to the first presentation. Therefore, both absolute and relative pitch (or frequency) information were available in the task (Bartlett & Dowling, 1980; Dowling & Fujitani, 1971; Krumhansl, 2000; Levitin, 1994; Schellenberg et al., 2014). Though relative pitch information seems to be predominantly used in long-term memory for melodies (Krumhansl, 2000), recent studies show that absolute pitch information is also retained (Schellenberg et al., 2014). One possible explanation for our findings in Experiments 1 and 2 could be that absolute frequency information compensated for an interference effect of the number of intervening melodies. It may be that an effect of intervening melodies will be observed when only relative frequency information is available. To test this possibility, two additional experiments were designed with melodies that were transposed on their repetition.

The results of Experiment 1 and Experiment 2 could also be limited to the specific corpus of melodies that resembled advertisement jingles or movie themes. To investigate whether the results can be replicated in other melody corpora, Experiments 3 and 4 also used a new set of melodies taken from a large corpus of European folk songs. Furthermore, prior research in the visual domain has shown cumulative disruptive interference for complex and meaningful visual

ACCEPTED MANUSCRIPT

stimuli (photographs) between 40, 80, 120, 160, and up to 200 intervening items (Nickerson, 1965). Potentially, the slope of the cumulative disruptive interference in memory for melodies may be too shallow to be detected with only a span of 13 intervening items. Similar to complex and meaningful visual stimuli, cumulative disruptive interference may emerge when investigating larger scale differences in the number of intervening items. While Experiments 1 and 2 investigated a relatively early time course of memory for melody, Experiments 3 and 4 were designed to investigate long-term effects of intervening melodies with the use of up to 195 intervening melodies.

Experiment 3: Melody Recognition with up to 195 Intervening Melodies

Experiment 3 was designed to investigate the influence of large numbers of intervening items when absolute frequency information is removed through melody transposition. Furthermore, the experiment investigated potential long-term effects of the number of intervening melodies on melody recognition with a new stimulus set derived from a large corpus of European folk songs.

Method

Participants. Thirty-two undergraduate students from the Western Sydney University (M_{age} = 21.03 SD_{age} = 5.64, six male) participated in this experiment. Participants were required to have received less than two years of musical training (five participants had musical training *M* = .34 years, *SD* = .90). Participants reported normal hearing and did not participate in the previous experiments. Participation was reimbursed with course credit as part of university course requirements.

ACCEPTED MANUSCRIPT

Stimuli and Material. Experiment 3 used an exhaustive stimulus selection procedure to create the final set of stimuli. A European folk song corpus of 8,397 monophonic melodies was analyzed (CCARH; Sapp, 2005). All melodies were deconstructed into their underlying musical features using the MIDI Toolbox (Eerola & Toiviainen, 2004a; 2004b, p. 96), FANTASTIC (Müllensiefen, 2009, p. 37), as well as several self-implemented routines to measure tonality (Dean, Bailes, & Drummond, 2014), autocorrelation between pitch values (Dean, Bailes, & Dunsmuir, 2014; Dean & Dunsmuir, 2015), and pitch as well as rhythmic balance and evenness (Milne, Bulger, Herff, & Sethares, 2015). Deconstructing melodies into their underlying musical features was necessary to ensure in a later step that the final subsample of melodies adequately represented the underlying corpus in respect to various musical features. An in-depth description of the musical features can be found in the MIDI Toolbox and FANTASTIC manuals (Eerola & Toiviainen, 2004a; 2004b, p. 96; Müllensiefen, 2009, p. 37). Due to the vast number of musical features (116), a principal component analysis was used to reduce the dimensions. Twenty-one significant underlying components were identified using a permutation based Monte Carlo Simulation (Parallel analysis) with a 95% confidence level (Clarkson & Jennrich, 1988; O'Connor, 2000). The score on every principal component was calculated for every melody in order to cluster the melodies in the dimension-reduced space. A hierarchical cluster analysis using Euclidean distances resulted in one large cluster. A cluster analysis using reduction in log-likelihood was implemented as a distance measurement and a nine-cluster solution emerged from the single large cluster using Bayesian Information Criteria (average silhouette measure of cohesion and separation = .01 with a smallest to biggest cluster ratio of 2.587; (Rousseeuw, 1987). The nine-cluster solution was accepted.

ACCEPTED MANUSCRIPT

In total, 110 melodies were randomly drawn, and the specific number of melodies from each cluster was determined by the relative cluster size. These 110 melodies were then subjected to a perceptual pilot study in order to identify melodies that, despite being European folk songs, evoke high degrees of familiarity from Australian listeners. Twelve (M_{age} = 33, SD_{age} = 12.5) participants provided familiarity ratings on a 100-point visual analog scale. The familiarity response distributions for every individual melody were compared to the response distribution of all other melodies using non-parametric Kolmogorov-Smirnov tests. Outlier melodies that showed significantly different response distributions were removed because they could not be classified as ‘novel’ for the purpose of the experiment. This procedure resulted in 98 melodies that were mathematically derived from a large corpus of European folksong melodies and perceptually tested to be novel to Australian listeners. These 98 melodies had a mean duration of 10.86 s and a mean pitch range of 7.93 semitones. Musical scores comprising representative examples of melodies presented in Experiment 3 can be found in Appendix A.

Procedure. The procedure of Experiment 3 closely followed the procedure of the previous experiments and is visualized in Figure 6. Participants provided informed consent and instructions were presented in a standardized format on the computer. A demographic questionnaire was also digitally administered. Participants were instructed to *“carefully listen to the melodies and, after each melody, please respond whether you have heard it before in this experiment”*. After a melody was presented, two buttons appeared on the screen, one labeled “*New*” and the other one “*Old”*. Responses were made using the computer mouse[2]. The 98 melodies were presented in 196 trials. Every participant received a unique randomization list that

---

[2] Note that in addition to the above-described procedure, remember/ know judgments (Tulving, 1985; Yonelinas, 2002) and confidence ratings were also measured after each participants’ recognition response (data not shown).

ACCEPTED MANUSCRIPT

was automatically generated at the beginning of the experiment. Melodies occurred twice throughout the experiment and, similar to previous studies (Schellenberg et al., 2014), were transposed on their repetition by six semitones up (to the key most distant from the original)[3].

Up to three participants were tested simultaneously, but each participant could neither hear the stimuli presented to others, nor see the other participants as they completed the task. The experiment was programmed in Max MSP 6.0 and executed in Max Runtime 6.0 (Cycling74, 2014) and took approximately 45 minutes to complete.

Figure 6 about here

Results

Figure 7 shows mean false alarm rates and mean hit rates for every participant and melody. Figure 8 shows the probability of Bias-Corrected Hits with increasing number of intervening items. Overall, participants performed significantly above chance ($Z = 12.77$, $p < .0001$) with results that replicated and extended those reported in Experiments 1 and 2: the number of intervening melodies did not influence recognition performance when up to 195 intervening melodies were presented with a different melody corpus and with transposition on their second presentation.

---

[3] Data from a pilot study as well as Experiments 3 and 4 show that participants did not use the pitch height of melodies as the basis of their judgments. Several mixed effects models were built to evaluate this. The melodies varied in average pitch height. If pitch height had been used as a systematic cue to respond ‘old’ then melodies with higher average pitch height would show an increased number of ‘old’ responses. However, average pitch height variation (as calculated by FANTASTIC) between the melodies did not carry predictive value for familiarity or recognition judgments during the second presentations (all *p*-values > .18). This means that higher average pitch height does not systematically shift response tendencies toward old-responses.

ACCEPTED MANUSCRIPT

Figure 7 and Figure 8 about here

As before, mixed effects models addressed the predictive value of the number of intervening items on melody recognition. A base model with a systematic factor for *Dynamic Response Tendency* and random intercepts for *Participant* and *Melody* (*BIC* = 3975.0; *LogLik* = -1969.8) did not significantly improve when provided with information about the *Number of Intervening Melodies* (*BIC* = 3979.8; *LogLik* = -1969.8, *p* >.07). This result again shows that the number of intervening items cannot be used to predict recognition performance in memory for melody when up to 195 intervening melodies are presented. This is illustrated in Figure 8.

A final model assessment using maximal random effect structure confirms that the number of intervening items does not carry systematic predictive value in the context of memory for melody (*p* > .60). Similar to the previous experiments, without a random intercept for *Melody (p* < .00001, ΔBIC = 14.8) the model performed significantly worse, showing significant differences between melodies. Figure B.2 in Appendix B displays uncorrected hit-rates over the different intervening item conditions in Experiment 3.

Discussion

Descriptively, performance in Experiment 3 was worse than in Experiments 1 and 2, reflecting the impact of transposition. This observation is consistent with recent findings that surface features in music, such as absolute frequencies, play a significant role in memory for

ACCEPTED MANUSCRIPT

melodies even at relatively long delays (Schellenberg & Habashi, 2015). Interestingly, melodies were still resistant to interference by intervening items, and that resistance was demonstrated here with up to 195 intervening items. Thus, relative pitch information alone sufficed to support the resilience. This is consistent with previous research showing that relative pitch information is important for long-term memory for melodies (Dowling & Bartlett, 1981). The results so far might be specific to tasks that focus on the explicit *recognition* of melodies. Experiment 4 addressed the possibility that cumulative disruptive interference arises when participants are not aware that their memory is being tested. This hypothesis is tested here by measuring changes in perceived *familiarity* between first and second presentations of novel melodies.

Experiment 4: Perceived Familiarity

Experiment 4 was designed to investigate the effect of the number of intervening items in a different assessment of memory than melody recognition. Rather than measuring recognition performance, participants were instructed to rate perceived familiarity for every melody without being told that melodies reoccur. If the number of intervening melodies has a significant influence on perceived familiarity, then it may be that the results observed in Experiments 1, 2, and 3 are underpinned by a mechanism specific to explicit memory task instructions. In an explicit memory task such as the recognition paradigm implemented in Experiments 1, 2, and 3, participants are aware that their memory is being tested. Prior research has shown that the degree to which participants are aware of the nature of a memory tasks can influence fundamental memory phenomena (Fleischman et al., 2004; Gaudreau & Peretz, 1999; Halpern & Bartlett, 2010; Halpern & O'Connor, 2000). Specifically, some researchers posit that conscious recollection and a general feeling of familiarity are underpinned by different neurological

ACCEPTED MANUSCRIPT

processes (Yonelinas, 2002). If there is a pattern of results supporting no significant effect of the number of intervening melodies on perceived familiarity, then the findings reported in Experiment 1, 2, and 3 are not specific to an explicit melody recognition paradigm.

Although in Experiment 4 participants were not informed that melodies may be repeated throughout the experiment, it was hypothesized that perceived familiarity should increase from the first to the second presentation of each melody. Furthermore, it was hypothesized that the number of intervening melodies does not affect this change in perceived familiarity between the first and second presentation of a melody.

Method

Participants. Thirty undergraduate students from the Western Sydney University volunteered to participate (M_{age} = 23.6, SD_{age} = 6.234, eight male). Participants did not have formal musical training nor did they report any hearing disabilities or participate in the previous experiments. Participation was reimbursed with university course credit.

Stimuli and Equipment. Stimuli and equipment were identical to Experiment 3 and musical scores relating to representative examples of melodies presented in Experiment 4 are presented in Appendix A.

Procedure. The procedure was to identical to Experiment 3 with the following exception: Participants were asked to indicate "*how familiar you perceive each melody to be*" rather than being asked to make recognition judgments. A vertical 100-point visual analogue scale was used. The familiarity scale had a spatial extent of 10 cm on the computer display and

ACCEPTED MANUSCRIPT

was labeled *unfamiliar* at the bottom and *familiar* at the top. Note that participants were not informed about the reoccurring nature of the melodies.

Results

In summary, Experiment 4 replicated and extended the results of the previous three experiments in a task that measured perceived familiarity, rather than a binary recognition response. Figure 9 shows mean perceived familiarity towards first and second presentation for each participant and each melody. Figure 10 shows perceived familiarity with increasing number of intervening items. Recall that participants were not instructed that melodies reoccur. After testing, participants were invited to speculate about the purpose of this experiment. No participant suspected that their memory was being tested. As hypothesized, perceived *familiarity* increased from the first (*M* = 38.68, *SE* = 2.89) to the second occurrence *(M* = 47.12, *SE* = 3.214) of the melodies (*t* = 11.53, *p* < .0001).

Figure 9 and 10 about here

A model predicting within-participant standardized *z*-scores of *Familiarity* with a random intercept of *Subject* and *Melody* (*BIC* = 13660.51; *LogLik* = -6813.26) performed significantly worse (*p* < .001, ΔBIC = 129.24) than the same model provided with the additional information of melody *Occurrence*; that is, the first or second occurrence of a melody (*BIC* = 13531.27; *LogLik* = -6744.40). Furthermore, a model using the maximal random effect structure also

ACCEPTED MANUSCRIPT

confirmed the significant effect of *Occurrence* ($p < .001$). These data further support the hypothesis that overall familiarity changes significantly between the first and second occurrence of a melody, even when taking random subject and item intercepts into account. Similar to all previous experiments, the above model performed significantly worse without a random intercept for *Melody* (*BIC* = 13788.31; *LogLik* = -6877.16, $p < .001$, ΔBIC = 257.04). This result provides evidence that there were significant differences between melodies.

The second hypothesis predicted a pattern of results that support no cumulative disruptive interference of the number of intervening melodies on the change in perceived familiarity between the first and second occurrence of a melody. This hypothesis was also supported. A mixed effects model with a random intercept for *Melody* and *Subject* and systematic factor of *Dynamic Response Tendency* predicting the change of familiarity in participant-wise *z*-scores between the first and second occurrence of a melody (*BIC* = 7797.5; *LogLik* = -3879.3) did not significantly improve ($p > .66$) when provided with information regarding the *Number of Intervening Melodies* (*BIC* = 7805.1; *LogLik* = -3879.2). This is illustrated in Figure 10. The maximal random effect structure also showed that the *Number of Intervening Melodies* was not a useful predictor of perceived familiarity ($p > .73$).

Discussion

The results of Experiment 4 show that the lack of interference by intervening melodies (Experiments 1-3) is not limited to explicit memory task instructions, but also occurs in an indirect memory task that measures perceived familiarity. This was an important test, since previous work has suggested that conscious recollection and the feeling of familiarity are two distinct mechanisms in recognition (Yonelinas, 2002). However, the present findings suggest that the lack of a disruptive effect of the number of intervening melodies is not underpinned by a

ACCEPTED MANUSCRIPT

mechanism specific to explicit memory task instructions or conscious recollection, but is also observed in indirect assessments of memory such as perceived familiarity.

General Discussion

The present study undertook a multi-experiment investigation of how the number of intervening melodies affects memory for melody. Over four experiments involving transposed and untransposed melodies, melody recognition and perceived familiarity assessments, and up to 195 intervening melodies in two different corpora of music, there was no indication of a disruptive effect from the number of intervening melodies on memory for melody beyond immediate repetition.

The present study provides support for the suggestion of Dowling, et al. (1995) that there may be underlying automatic processes involved in the recognition of novel melodies that are not disrupted by the presentation of intervening melodies. In the present investigation, the number of intervening items and the passing of time were closely intertwined, as additional time was required to include additional intervening items. This means that the results here also support Schellenberg and Habashi's findings (Schellenberg & Habashi, 2015) that, in contrast to other domains (Karnekull, Jonsson, Willander, Sikstrom, & Larsson, 2015), the mere passing of time does not interfere with melody recognition. The results also converge with the majority of literature showing that successful recognition of melodies is possible, even when they have only been heard a few times; in our case, only once (Bartlett, Halpern, & Dowling, 1995; Dowling, 1991; Dowling et al., 1995; Halpern & Bartlett, 2010, 2011; Halpern & Müllensiefen, 2008; McAuley, Stevens, & Humphreys, 2004; Müllensiefen & Halpern, 2014; Peretz & Gaudreau, 1998; Schellenberg & Habashi, 2015; Schellenberg et al., 2014).

ACCEPTED MANUSCRIPT

However, our findings are relatively surprising when one considers that a disruptive effect from the number of intervening items is reported in the context of many domains other than music (Buchsbaum et al., 2011; Bui et al., 2014; Campeanu et al., 2014; Hockley, 1992; Konkle et al., 2010; Nickerson, 1965; Olson, 1969; Sadeh et al., 2014). Note that we do not claim that memory for melody is exceptionally good compared to other stimuli. The present data allow no direct cross-domain performance comparisons. However, the lack of a disruptive effect from the number of intervening items observed here is has been reported in only a few cases in other domains (Berman et al., 1991; Tillmann & Dowling, 2007).

In the following, our results will be discussed in terms of melodic transposition, melody distinctiveness, melody recognition and perceived familiarity, music's temporal organization and its relation to memory's domain specificity, as well as a novel regenerative multiple representations conjecture that offers a pathway for future research designed to investigate the psychological mechanisms that may explain our findings.

Melodic Transposition

The findings of Experiments 1 and 2 showed no effect of the number of intervening melodies when pitch information was available to participants in the form of each note's original pitch (absolute frequency information) and the relative pitch intervals between notes (relative frequency information). Experiments 3 and 4 extended these findings, showing there is also no effect of the number of intervening melodies when only relative frequency information is available. Absolute frequency information appears to serve as additional information that is used to aid melody recognition. As indicated by our data as well as previous research, this is demonstrated in better recognition performance towards untransposed compared to transposed

ACCEPTED MANUSCRIPT

melodies (Bartlett & Dowling, 1980; Dowling & Fujitani, 1971; Krumhansl, 2000; Levitin, 1994; Plantinga & Trainor, 2005; Schellenberg et al., 2014). This raises the question: do changes in other important physical properties of music lead to similar results? For example, a change of musical timbre between first and second presentations of a melody would be a useful manipulation to test this hypothesis.

Here, we tested two different melody corpora representing different musical ‘genres’ (modern advertisement jingles vs. European Folk melodies). However, the melodies within each corpus were distinctly different to each other. An important future direction that is beyond the scope of the present investigation is the question regarding how similarity between intervening melodies and the target melodies affect recognition performance. Within the two melody corpora tested here, we did not observe cumulative disruptive interference from the number of intervening melodies. However, the present findings do not necessarily generalize to cases where intervening melodies are significantly more, or less, similar to the target melodies. In single-note recognition, for example, the degree of similarity between the target pitch and intervening pitches greatly mediates interference, with greater dissimilarity leading to greater interference (D. Deutsch, 1972). Interestingly, substantially different intervening items such as spoken numbers do not cause cumulative disruptive interference (D. Deutsch, 1970). Future studies could further investigate cumulative distractors and how their similarity to a target melody influences recognition performance. A major difficulty that such studies will need to overcome is establishing the perceptual similarity of melodies such as used in this work, as well as their cumulative effects.

Melody Distinctiveness

ACCEPTED MANUSCRIPT

While the number of intervening melodies did not show any significant effect on melody recognition, the melodies within our two corpora did. Some melodies showed high recognition performance even after large numbers of intervening melodies, while others failed to be recognised after only one intervening melody. This suggests that some melodies were not successfully encoded in the first place. Our data suggest that once a melody is successfully encoded, the number of intervening melodies does not influence the retrieval process.

It is reasonable to expect that the specific combinations of underlying musical features in the melodies provide predictive value when it comes to melody recognition. Initial evidence for this hypothesis is found in a recent investigation using a range of musical features in melodies to predict melody recognition performance (Müllensiefen & Halpern, 2014). This study showed that less common motifs relative to a corpus could predict correct recognitions. This is analogous to the visual domain, where better long-term recognition for vivid pictures or oddities is reported (Konkle et al., 2010; Standing, 1973). The importance of musical features for prediction of successful melody recognition is a promising avenue for future investigation. The present study provides behavioural data that will aid in the development of mathematical models that use musical features as recognition predictors. The data further facilitate the endeavor of building predictive models of melody recognition by demonstrating that the number of intervening items, a predictor that carries substantial predictive power in other memory domains, does not seem to apply to melody recognition.

Melody Recognition and Perceived Familiarity

Unlike the situation in Experiments 1-3, participants in Experiment 4 were only instructed to report their feeling of perceived familiarity. Nevertheless, none of the experiments

ACCEPTED MANUSCRIPT

revealed an influence of the number of intervening items on melody recognition or familiarity. In the familiarity task, an increase in perceived familiarity was observed between the first and second presentation of each melody. This suggests that participants formed memory representations of the melodies during the first presentations that increased perceived familiarity when a melody was heard again. Interestingly, the number of intervening melodies did not influence the increase in familiarity. This shows that the lack of a disruptive effect reported here with up to 195 intervening melodies is not limited to melody recognition instructions, but can be extended to less explicit measurements of memory such as perceived familiarity. Further studies can investigate whether this finding is replicated with even less explicit measurements of memory, such as preference ratings in the form of mere exposure effects, or reaction time.

Temporal Organization and Domain Specificity

One inherent feature of music is that it continuously unfolds through time and comprises successively organized (rhythmic) events (Jackendoff & Lerdahl, 2006). A melody develops from its first note to its last, and the stimulus as a whole is complete only when the last note has been sounded. It has long been suspected that this temporal organization of music lies at the heart of some important psychological phenomena related to music. For example, in a study by Dowling and colleagues (Dowling, Tillmann, & Ayers, 2001), participants listened to short phrases from classical minuets. After 4-5, 15, or 30 seconds, participants were required to discriminate between the initial phrase and similar lures. Discrimination performance increased with greater temporal delays, but only if the delay was filled with a continuation of the music. No such improvement was observed when the delay was filled with silence or a purely rhythmic continuation. The findings were attributed to an ongoing process of feature binding that assist in

ACCEPTED MANUSCRIPT

forming coherent representation of the melodies. Dowling and Tillmann (2014) suggest that this process runs as long as ‘similarity’, ‘continuity’, and ‘coherence’ of the stimulus is not interrupted. The authors conclude that “the important thing is not that the delay be filled, but that it be filled with musically meaningful material that engages the listener” (Dowling et al., 2001, p. 270). This condition is somewhat analogous to the continuous memory paradigm used in the present study. Specifically, any detrimental effect to memory from the number of intervening items or the passing of time may have been compensated by a domain-specific increase of performance when the ‘delay’ (in our case, the period of time comprising melodies) is filled with meaningful melodies. A notable difference between the Dowling, et al. (2001) study that used classical minuets and the present study is that here, delays were filled with different melodies, rather than continuations of the target stimulus. Furthermore, the continuation of the listening experience was interrupted by participants’ responses after each melody. Nevertheless, we did not observe a disruptive effect from the number of intervening items. This may indicate that findings reported in Dowling, et al. (2001) in the context of intervening delays filled with related musical content extends to cases where the delay is filled with unrelated musical material at least that of a similar style. Indeed, such findings have been replicated using non-classical guitar music (Dowling, Magner, & Tillmann, 2016). Similar findings have also been reported in poetry but not in prose (Tillmann & Dowling, 2007): while memory performance declines over time for prose, the mere passing of time has no effect on memory performance for poetry.

Interestingly, an effect of the number of intervening items on recognition is again absent with simple drawings, but it is observed with photographs (Berman et al., 1991; Friedman, 1990a; Konkle et al., 2010). The reasons for this discrepancy are currently unclear but we provide one possible explanation below. The present results show again that findings related to

ACCEPTED MANUSCRIPT

memory in one domain may not necessarily generalize to others (Fougnie et al., 2015). However, when observing several different domains that act similarly, but differently to others, one wonders whether these domains share underlying processes. In this case, a fair overarching question is as follows: what are the similarities of music, poetry, and drawings that lead to a lack of an effect from the number of intervening items and the mere passing of time? In the following, we propose a novel *regenerative multiple representations* (RMR) explanation. This perspective provides clear predictions, is falsifiable, and offers future possibilities for mathematical implementation. However, the conjecture is still speculative and should only serve as pointer for potential future research.

A Novel Perspective: Regenerative Multiple Representations

The temporal structure of music is realized as a relational structure of underlying temporally spaced components. Examples of such underlying components are notes, pitch intervals, rhythms, or short note phrases within a melody. We define such relational organization as the strong interdependent connection between each element that defines and gives meaning to an object; the 'process of perceptual synthesis' that integrates the fragmented features of an auditory stimulus (D. Deutsch, 1986). In the following, we intend to emphasize the importance of the underlying relational organization, rather than the temporal organization that emerges out of the relational organization.

Relational organization is relevant for perception of most objects in our day-to-day experience. For example, an everyday object such as a chair consists of underlying geometrical shapes. While the relational organization of such an object is often clear, the relevance of different layers in this organization might depend not only on the object, but also on the

ACCEPTED MANUSCRIPT

observer. We have learned that a chair is important to be perceived as an integrated whole. The exact underlying components are often of little relevance. A Joiner experienced in constructing chairs, however, may have a different perception of the same chair that is likely to include additional underlying components. Thus, our prior experience can inform perceptual relevance. In formulating the RMR conjecture, we assume that perception directly influences memory; objects that are perceived as a whole will predominantly be remembered as an integrated whole, objects that are perceived as underlying components may only be remembered as those components.

Usually, we tend to perceive objects as either a whole or as their components, but not both at the same time (Goldstein, 2013, pp. 100-114). This observation has been exhaustively studied in the perceptual grouping and object recognition literature (Wolfe et al., 2012). In general, the perception of objects as a whole is favored because it is often more meaningful to the perceiver. In music, it has also long been theorized that melodies are integrated in memory as a whole (see (D. Deutsch, 1986; Krumhansl, 1991, p. 295). Following the assumption that memory is guided by perception (Malmberg & Annis, 2012), the perception of an object as a whole might then lead to a representation in memory of the object as a whole. Such a representation could then be subject to interference and decay, until the representation drops below some form of recognition threshold, rendering retrieval impossible. This process is often implemented mathematically in memory models in the form of a decay parameter, an interference parameter, or both (Norman, 2013; Oberauer & Lewandowsky, 2011; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012).

However, there could plausibly be a group of objects that are simultaneously perceived as both an integrated relation *and* as two or more sets of components that create that relation. The

ACCEPTED MANUSCRIPT

reason this group of objects may be perceived this way is that the perceiver has learned over time that it is important to pay attention to several aspects of the object, as well as their relations. As a result, the object is perceived simultaneously on both levels and thus, multiple representations are formed?. We speculate that such multiple representations may be what music, poetry, and drawings have in common. Outside of the memory domain, the hypothesis that music might be represented as a complex whole as well as its underlying components has been proposed as early as 1873 (as discussed in (Schneider, 1997, p. 119). In the context of the present study, an integrated melody is an example of the object as a 'whole', whereas particular features such as notes, intervals, or short note clusters within the melody are an example of its underlying components. Stimuli that elicit multiple representations may have an advantage: if one representation fades below the threshold of retrieval, it may still be reconstructed or regenerated by cross-referencing with other representations. We suggest that *regeneration* is triggered when a single representation is accessed that is below the recognition threshold.

For example, a person may find it difficult to immediately answer whether the first and last tone of *Mary Had A Little Lamb* are the same pitch (a question that tests memory for underlying components). However, the same person may still be capable of humming the complete melody of *Mary Had a Little Lamb* (an example of retrieving the integrated whole), which results in the access to information necessary to correctly answer the initial question. In this example, if the person is prompted to answer quickly, performance may be low. However, if the person is provided with additional time to make full use of their integrated representation of the melody as a whole, then they should easily *regenerate* the specific information regarding the underlying components and perform the task with relative ease. This conjecture predicts that once encoded, such regenerative multiple representations are robust against the interference of

ACCEPTED MANUSCRIPT

time (decay) and intervening items, since most of the multiple representations would be theoretically required to drop below the hypothetical retrieval threshold before retrieval is impossible. The representations might be long lasting, though not necessarily permanent. Such an effect would also be independent of overall memory performance from one group of stimuli compared to another, since it only describes what might happen to objects once they are encoded. Indeed, some stimuli are likely to be harder to encode than others.

Taken together in the context of melody recognition, this conjecture speculates that: (1) melodies consist of underlying perceivable components such as notes, intervals, or short note clusters; (2) perceivers have learned that it is important to perceive the underlying components of a melody and subsequently form a memory representation of them; (3) perceivers have learned how underlying components of a melody are related; (4) perceivers are capable of integrating the relational dependency into a whole (e.g., a melody). Consequently, multiple representations specific to the underlying components of the melody and the melody as a relation, an integrated whole, are developed; (5) representations in general are subject to forgetting. However, if one representation fades, some of the others which remain intact can regenerate it. ‘Regenerate’ here refers to information that can be retrieved by using information in other memory representations, even though the original memory representation is lost. The greater number of intact representations, the more resilient a memory is likely to be. If and how many multiple representations are formed depends on both stimulus properties and observer specifics (e.g., prior knowledge and experience). In music, most listeners have likely learned how to integrate melodies over years of exposure, but also learned that the individual underlying entities of music such as notes and pitch relations are crucial for the formation and understanding of such melodies. These melodic expectancies are a candidate-mechanism for the aforementioned

ACCEPTED MANUSCRIPT

regeneration process, as they help predict what comes next in a melody (Margulis, 2005; Pearce, 2014; Schellenberg, 1996) and could potentially be used to interpolate forgotten parts of a melody. As a result, the RMR conjecture predicts that the resilience of memory for melody should continuously build up as a listener becomes more familiar with the tuning system in which it is heard. Training listeners on a new tuning system and continuously testing for cumulative disruptive interference could investigate this prediction. Other implications of the RMR conjecture are discussed in the following.

Implications of Regenerative Multiple Representations

The regenerative multiple representations conjecture may also explain why an effect of the number of intervening items is observed in *photographs* of everyday objects (Konkle et al., 2010), but not in *drawings* of everyday objects (Berman et al., 1991). Drawings may guide attention to the underlying components of a represented object, which leads to multiple representations of the object as a whole, as well as its underlying components. However, photos of everyday objects would predominantly be perceived as the object itself. An additional component unique to the drawing, potentially contributing to regenerative events might be the brushstroke style. This leads to the testable hypothesis that memory for words written in longhand shows no effect of the number of intervening items, while memory for plain printed words does. In addition to the representation of a word as its integrated meaning, longhand might draw attention to individual strokes (underlying components) and their spatial relation to each other, forming another representation of the underlying components, whereas plain printed words would for most observers only result in a representation of the integrated meaning. Consistent with this, participants can perceive and use information related to the mode of production of

ACCEPTED MANUSCRIPT

words in longhand (Babcock & Freyd, 1988; Knoblich, Seigerschmidt, Flach, & Prinz, 2002; Tse & Cavanagh, 2000).

At this stage, the RMR conjecture cannot predict *exactly* how many memory representations plain printed words have. However, the RMR conjecture can predict that words written in longhand have the same representations that plain printed words have, plus additional representations of the underlying strokes and should therefore be more resilient to intervening items effects. Interestingly, the RMR conjecture also predicts resilience towards cumulative disruptive interference in cases that have previously shown such interference, given that the observer forms more memory representations. An expert in photography composition, for example, is likely to perceive more components of photographs than average naïve observers. These additional percepts form additional memory representations, which should provide the expert with additional resilience towards cumulative disruptive interference.

The RMR conjecture also predicts the finding of an effect of the number of intervening items in words and prose, but not in poetry (Bui et al., 2014; Tillmann & Dowling, 2007). Usually, a word is predominantly represented as a whole and seldom as its actual underlying components. For example, correct letters in a word that are placed at an incorrect position are often perceptually processed as if they were in the correct position (Rayner, Pollatsek, Ashby, & Clifton Jr, 2012a). The same applies to whole sentences. Usually the meaning is represented while individual words are sometimes ‘skipped’ during reading (Rayner, Pollatsek, Ashby, & Clifton Jr, 2012b). In poetry, on the other hand, attention may be shifted towards the underlying components (e.g., individual words), as they carry greater importance and may be sequenced irregularly. As a result, more representations of underlying components, such as underlying words and how they interact with each other, are likely to manifest in poetry while still

ACCEPTED MANUSCRIPT

maintaining an overall representation for every line, stanza, and the entire poem as an integrated whole. In summary, informed by prior experience with a stimulus, the degree to which different properties of a stimulus are salient changes.

Providing a task that forces participants to focus on underlying elements, the whole, or both simultaneously would be a suitable context to test the multiple regenerative representations conjecture. Furthermore, one would expect to find similar results in domains that also favour multiple representations. One example of such a domain is dance. Dance consists of dynamic movement as well as underlying postures, and both have been shown to contribute to recognition of contemporary dance postures (Vicary, Robbins, Calvo-Merino, & Stevens, 2014).

Regenerative Multiple Representations in Relation to other Memory Models

Regenerative multiple representations is similar to previous memory theories in so far as it draws from multiple-trace theories (Hintzman, 1984, 1988). It also does not challenge global matching models that describe recognition as the *'match'* response of memory that reflects a familiarity distribution to a cue (Clark & Gronlund, 1996). However, even in-depth memory models such as REM, SLiM, and BCDMEM predict interference effects from lags or delays that were not observed in the present study's findings (Dennis & Humphreys, 2001; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997).

The regenerative multiple representations conjecture also bears resemblance to Paivio's dual-coding theory (Paivio, 1969). It assumes two representations (words and images) that assist in retrieval of each other, therefore increasing the chance of remembering a stimulus. Multiple regenerative representations postulate any number of representations, and not necessarily words and images. This is important for a theory applied to music or other stimuli where multiple

ACCEPTED MANUSCRIPT

representations beyond two are likely. Indeed, Paivio’s dual-coding theory could be described as a special case of the present multiple regenerative representations conjecture, where prior experience informs our perception to focus on words and their associated images.

The notion of perceptual relevance that is informed by prior knowledge is somewhat similar to the notion of *pertinence* described in D. Deutsch (1986) (see also J. A. Deutsch and Deutsch (1963). Pertinence describes a weighting of a perception based on a current situation, as well as long-term factors such as prior knowledge. However, while pertinence primarily influences awareness, we suggest that perceptual relevance directly influences perception and subsequent formation of future memories: a chair that is perceived as an integrated whole will be remembered as an integrated whole rather than the underlying components.

The regenerative multiple representations perspective is speculative, with many open questions. For example, how can the relative strength of representations be measured? Can perceptual relevance be manipulated to form multiple representations and, as a result, facilitate learning and memory encoding? The conjecture does allow for many specific and informative predictions and mathematical implementations that can be tested in future research endeavors. Ongoing work in our lab has tested and found support for some of the predictions made by the RMR conjecture. For example, one hypothesis concerned whether an effect of the number of intervening items will manifest if note based melodies are played in tuning systems that are completely unfamiliar to participants (e.g., microtonalities) (Herff, Olsen, Dean, & Prince, submitted). In this case, listeners should still be capable of perceiving some of the underlying component, but may fail to integrate the stimuli into perceptually coherent melodies. Based on the RMR perspective, the prediction was that the number of intervening items will have a significant negative impact on memory performance, because listeners do not have the multiple

ACCEPTED MANUSCRIPT

representations necessary to utilize the regeneration process. This prediction was supported in Herff, et al. (submitted). Other predictions of the RMR conjecture concerning melodies that only vary in pitch or rhythm, rather than melodies with variations in melodic and rhythmic information combined, have also been tested and supported in Herff, Olsen, Prince, and Dean (2017).

Conclusion

Human memory is fallible and prone to interference. In many domains, memory performance decreases as the number of intervening items between the first and second presentation of a stimulus increases. However, this phenomenon is not universal, and music in particular has proven to possess intriguing properties when it comes to memory: these were further investigated here. Regardless of whether one or 195 intervening melodies were presented, performance in all our experiments was above chance and not affected by the number of intervening melodies. This finding was observed using two different musical corpora. Furthermore, transposition of each melody's repetition left only relative frequency information, yet the number of intervening melodies still did not affect memory for melody. In addition to explicit recognition, one experiment measured memory for melody in the form of perceived familiarity, with results that were consistent with the other findings of the study.

To explain the findings, we offer a novel yet speculative *regenerative multiple representations* conjecture that bears resemblance to Paivio's dual-coding theory (Paivio, 1969). The conjecture assumes that: previous experience influences our perception; perception determines which memory representations are formed; memory representations are subject to decay and interference; there are stimuli where previous experience informs us to simultaneously

ACCEPTED MANUSCRIPT

perceive these stimuli in multiple ways; multiple perception leads to multiple representations; multiple representations can regenerate each other, making them resilient to decay and interference; melodies in familiar tuning systems belong to the category of objects that we simultaneously perceive in multiple ways, such as underlying components (e.g., notes, intervals or phrases) as well as an integrated, coherent whole (i.e., the melody).

Future studies investigating these assumptions in the context of music (e.g., familiar tonal versus unfamiliar atonal stimuli) as well temporally dynamic domains such as dance (e.g., dynamic movement versus underlying postures) and words (e.g., longhand versus printed) will provide empirical evidence or counterevidence regarding the *regenerative multiple representations* conjecture, thus facilitating its development into an empirically informed theory of human memory that can begin to answer the question: why does the sheer number of intervening items have no influence on memory for melody, when it does for almost all other memory domains?

ACCEPTED MANUSCRIPT

Author Note

We thank Andrew Milne, Daniel Müllensiefen, and the MARCS Institute Music Cognition and Action group for valuable comments on a previous version. Correspondence concerning this article should be addressed to Steffen A. Herff, MARCS Institute, Locked Bag 1797, Penrith, NSW 2751; s.herff@westernsydney.edu.au, + 61 2 9772 6801, https://www.westernsydney.edu.au/marcs

ACCEPTED MANUSCRIPT

References

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R.* New York: Cambridge University Press.

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. doi:10.1016/J.Jml.2007.12.005

Babcock, M. K., & Freyd, J. J. (1988). Perception of dynamic information in static handwritten forms. *The American Journal of Psychology*, 111-130.

Bachem, A. (1954). Time factors in relative and absolute pitch determination. *J Acoust Soc Am, 26*(5), 751-753.

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music*, *35*(4), 555-570.

Bailes, F. (2015). Music in mind? An experience sampling study of what and when, towards an understanding of why. *Psychomusicology: Music, Mind, and Brain, 25*(1), 58.

Bailes, F., Dean, R. T., & Pearce, M. T. (2013). Music Cognition as Mental Time Travel. *Scientific Reports*, *3*. doi:10.1038/srep02690

Baird, A., & Samson, S. (2014). Music evoked autobiographical memory after severe acquired brain injury: Preliminary findings from a case series. *Neuropsychological Rehabilitation, 24*(1), 125-143. doi:10.1080/09602011.2013.858642

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3). doi:10.1016/j.jml.2012.11.001

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of Transposed Melodies - a Key-Distance Effect in Developmental Perspective. *Journal of Experimental Psychology-Human Perception and Performance, 6*(3), 501-515. doi:10.1037/0096-1523.6.3.501

Bartlett, J. C., Halpern, A. R., & Dowling, W. J. (1995). Recognition of Familiar and Unfamiliar Melodies in Normal Aging and Alzheimers-Disease. *Memory and Cognition, 23*(5), 531-546. doi:10.3758/Bf03197255

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1*(4).

Berch, D. B. (1976). Criterion change in continuous recognition memory: A sequential effect. *Bulletin of the Psychonomic Society, 7*(3), 309-312. doi:10.3758/BF03337199

Berman, S., Friedman, D., & Cramer, M. (1991). ERPS during Continuous Recognition Memory for Words and Pictures. *Bulletin of the Psychonomic Society, 29*(2), 113-116.

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception*, 353-364.

Buchsbaum, B. R., Padmanabhan, A., & Berman, K. F. (2011). The Neural Substrates of Recognition Memory for Verbal Information: Spanning the Divide between Short- and Long-term Memory. *J Cogn Neurosci*, *23*(4), 978-991. doi:10.1162/Jocn.2010.21496

Bui, D. C., Maddox, G. B., Zou, F., & Hale, S. S. (2014). Examining the lag effect under incidental encoding: Contributions of semantic priming and reminding. *Quarterly Journal of Experimental Psychology*, *67*(11), 2134-2148. doi:10.1080/17470218.2014.909506

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Campeanu, S., Craik, F. I. M., Backer, K. C., & Alain, C. (2014). Voice reinstatement modulates neural indices of continuous word recognition. *Neuropsychologia, 62,* 233-244. doi:10.1016/J.Neuropsychologia.2014.07.022

CCARH. European Folk Songs. Retrieved from http://kern.ccarh.org/cgi-bin/browse?l=essen/europa

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3*(1), 37-60. doi:10.3758/Bf03210740

Clarkson, D. B., & Jennrich, R. I. (1988). Quartic Rotation Criteria and Algorithms. *Psychometrika, 53*(2), 251-259. doi:10.1007/Bf02294136

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children: evidence from behavioral and electrophysiological methods. *Developmental Science, 17*(1), 142-158. doi:10.1111/desc.12100

Cuddy, L. L., & Duffin, J. (2005). Music, memory, and Alzheimer's disease: is music recognition spared in dementia, and how can it be assessed? *Medical Hypotheses, 64*(2), 229-235. doi:10.1016/J.Mehy.2004.09.005

Cycling74. (2014). Max-MSP. Retrieved from http://cycling74.com/products/max/

Dean, R. T., Bailes, F., & Drummond, J. (2014). Generative Structures in Improvisation: Computational Segmentation of Keyboard Performances. *Journal of New Music Research*, 1-13. doi:10.1080/09298215.2013.859710

Dean, R. T., Bailes, F., & Dunsmuir, W. T. M. (2014). Time series analysis of real-time music perception: approaches to the assessment of individual and expertise differences in perception of expressed affect. *Journal of Mathematics and Music, 8*(3), 183-205. doi:10.1080/17459737.2014.928752

Dean, R. T., & Dunsmuir, W. T. M. (2015). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, 1-20. doi:10.3758/s13428-015-0611-2

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review, 108*(2), 452-478.

Deutsch, D. (1970). Tones and numbers: specificity of interference in immediate memory. *Science, 168*(3939), 1604-1605. doi:10.1126/science.168.3939.1604

Deutsch, D. (1972). Mapping of interactions in the pitch memory store. *Science, 175*(4025), 1020-1022.

Deutsch, D. (1975). The organization of short-term memory for a single acoustic attribute. In D. Deutsch & J. A. Deutsch (Eds.), *Short-term memory* (pp. 107-151). New York: Academic Press.

Deutsch, D. (1986). Auditory Pattern Recognition. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (Vol. II, Cognitive Processes and Performance, pp. 32-31 - 32-49). New York: Wiley.

Deutsch, J. A., & Deutsch, D. (1963). Attention - Some Theoretical Considerations. *Psychological Review, 70*(1), 80-90. doi:10.1037/h0039515

Donaldson, W., & Murdock, B. B. (1968). Criterion Change in Continuous Recognition Memory. *Journal of Experimental Psychology, 76*(3P1), 325-330. doi:10.1037/H0025510

ACCEPTED MANUSCRIPT

Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics, 14*(1), 37-40.

Dowling, W. J. (1991). Tonal Strength and Melody Recognition after Long and Short Delays. *Perception and Psychophysics, 50*(4), 305-313. doi:10.3758/Bf03212222

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology: A Journal of Research in Music Cognition, 1*(1), 30.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America, 49*(2), Suppl 2:524+.

Dowling, W. J., Kwak, S., & Andrews, M. W. (1995). The time course of recognition of novel melodies. *Perception and Psychophysics, 57*(2), 136-149. doi:10.3758/BF03206500

Dowling, W. J., Magner, H., & Tillmann, B. (2016). Memory improvement with wide-awake listeners and with nonclassical guitar music. *Psychomusicology: Music, Mind, and Brain, 26*(1), 26.

Dowling, W. J., & Tillmann, B. (2014). Memory Improvement While Hearing Music: Effects of Structural Continuity on Feature Binding. *Music Perception, 32*(1), 11-32. doi:10.1525/Mp.2014.32.1.11

Dowling, W. J., Tillmann, B., & Ayers, D. F. (2001). Memory and the experience of hearing music. *Music Perception, 19*(2), 249-276. doi:10.1525/Mp.2001.19.2.249

Eerola, T., & Toiviainen, P. (2004a). Midi Toolbox: MATLAB Tools for Music Research. Retrieved from http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/

ACCEPTED MANUSCRIPT

Eerola, T., & Toiviainen, P. (2004b, October 10-14, 2004). *MIR In Matlab: The MIDI Toolbox.* Paper presented at the 5th International Conference on Music Information Retrieval, Universitat Pompeu Fabra Barcelona, Spain.

Eysenck, M. W., & Keane, M. T. (2015). *Cognitive psychology: A student's handbook* (Seventh Edition ed.). London and NEw York: Taylor & Francis.

Ferris, S. H., Crook, T., Clark, E., Mccarthy, M., & Rae, D. (1980). Facial Recognition Memory Deficits in Normal Aging and Senile Dementia. *Journals of Gerontology, 35*(5), 707-714.

Fleischman, D. A., Wilson, R. S., Gabrieli, J. D. E., Bienias, J. L., & Bennett, D. A. (2004). A longitudinal study of implicit and explicit memory in old persons. *Psychology and Aging, 19*(4), 617-625. doi:10.1037/0882-7974.19.4.617

Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology-General, 144*(1), 30. doi:10.1037/a0038211

Franek, M., van Noorden, L., & Rezny, L. (2014). Tempo and walking speed with music in the urban context. *Frontiers in Psychology, 5*, 1361. doi:10.3389/Fpsyg.2014.01361

Friedman, D. (1990a). Cognitive event-related potential components during continuous recognition memory for pictures. *Psychophysiology, 27*(2), 136-148.

Friedman, D. (1990b). ERPS during Continuous Recognition Memory for Words. *Biological Psychology, 30*(1), 61-87. doi:10.1016/0301-0511(90)90091-A

Gaudreau, D., & Peretz, I. (1999). Implicit and explicit memory for music in old and young adults. *Brain and Cognition, 40*(1), 126-129.

Goldstein, E. (2013). *Sensation and Perception* (9th Edition ed.): Cengage Learning.

ACCEPTED MANUSCRIPT

Greene, R. L., & Samuel, A. G. (1986). Recency and Suffix Effects in Serial-Recall of Musical Stimuli. *Journal of Experimental Psychology-Learning Memory and Cognition, 12*(4), 517-524.

Halekoh, U., & Højsgaard, S. (2014a). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models–the R package pbkrtest. *Journal of Statistical Software, 59*(9), 1-32.

Halekoh, U., & Højsgaard, S. (2014b). pbkrtest: Parametric Bootstrap and Kenward Roger Based Methods for Mixed Model Comparison.

(Version R package version 0.4-0,). Retrieved from URL http: //CRAN.R-project.org/package=pbkrtest.

Halpern, A. R., & Bartlett, J. C. (2010). Memory for Melodies. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (Vol. 36, pp. 233-258). New York: Springer.

Halpern, A. R., & Bartlett, J. C. (2011). The Persistence of Musical Memories: A Descriptive Study of Earworms. *Music Perception, 28*(4), 425-431. doi:10.1525/Mp.2011.28.1.425

Halpern, A. R., & Müllensiefen, D. (2008). Effects of timbre and tempo change on memory for music. *Quarterly Journal of Experimental Psychology, 61*(9), 1371-1384. doi:10.1080/17470210701508038

Halpern, A. R., & O'Connor, M. G. (2000). Implicit memory for music in Alzheimer's disease. *Neuropsychology, 14*(3), 391-397. doi:10.1037//0894-4105.14.3.391

Herff, S. A., Olsen, K. N., Dean, R. T., & Prince, J. (submitted). Memory for melodies in unfamiliar tuning systems: Investigating effects of recency and number of intervening items.

ACCEPTED MANUSCRIPT

Herff, S. A., Olsen, K. N., Prince, J., & Dean, R. T. (2017). Interference in memory for pitch-only and rhythm-only sequences. *Musicae Scientiae,* Advance online publication, 1029864917695654. doi:doi:10.1177/1029864917695654

Hintzman, D. L. (1984). Minerva-2 - a Simulation-Model of Human-Memory. *Behavior Research Methods Instruments & Computers, 16*(2), 96-101.

Hintzman, D. L. (1988). Judgments of Frequency and Recognition Memory in a Multiple-Trace Memory Model. *Psychological Review, 95*(4), 528-551. doi:10.1037/0033-295x.95.4.528

Hockley, W. E. (1992). Item Versus Associative Information - Further Comparisons of Forgetting Rates. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*(6), 1321-1330. doi:10.1037//0278-7393.18.6.1321

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition, 100*(1), 33-72. doi:10.1016/J.Cognition.2005.11.005

Jacobsen, J. H., Stelzer, J., Fritz, T. H., Chetelat, G., La Joie, R., & Turner, R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain*. doi:10.1093/brain/awv135

Jahnke, J. C. (1963). Serial position effects in immediate serial recall. *Journal of Verbal Learning and Verbal Behavior, 2*, 284–287. doi:10.1016/S0022-5371(63)80095-X

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem. *Journal of Personality and Social Psychology, 103*(1), 54-69. doi:10.1037/A0028347

ACCEPTED MANUSCRIPT

Karnekull, S. C., Jonsson, F. U., Willander, J., Sikstrom, S., & Larsson, M. (2015). Long-Term Memory for Odors: Influences of Familiarity and Identification Across 64 Days. *Chemical Senses, 40*(4), 259-267. doi:10.1093/Chemse/Bjv003

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association, 90*(430), 773-795. doi:10.1080/01621459.1995.10476572

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*(3), 983-997. doi:10.2307/2533558

Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *The Quarterly Journal of Experimental Psychology: Section A, 55*(3), 1027-1046.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139*(3), 558-578. doi:10.1037/a0019165

Krause, A., North, A., & Hewitt, L. (2014). Music Selection Behaviors in Everyday Listening. *Journal of Broadcasting & Electronic Media, 58*(2), 306-323. doi:10.1080/08838151.2014.906437

Krumhansl, C. L. (1991). Music Psychology - Tonal Structures in Perception and Memory. *Annual Review of Psychology, 42*, 277-303. doi:10.1146/Annurev.Ps.42.020191.001425

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*(1), 159-179. doi:10.1037//0033-2909.126.1.159

Kruschke, J. K. (2010). Null Hypothesis Significance Testing *Doing Bayesian data analysis: A tutorial introduction with R* (pp. 265-287): Academic Press.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Kruschke, J. K. (2011). Bayesian Assessment of Null Values Via Parameter Estimation and Model Comparison. *Perspectives on Psychological Science, 6*(3), 299-312. doi:10.1177/1745691611406925

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General, 142*(2), 573-603. doi:10.1037/a0029146

Lange, K., & Czernochowski, D. (2013). Does this sound familiar? Effects of timbre change on episodic retrieval of novel melodies. *Acta Psychol (Amst), 143*(1), 136-145. doi:10.1016/j.actpsy.2013.03.003

Levitin, D. J. (1994). Absolute Memory for Musical Pitch - Evidence from the Production of Learned Melodies. *Perception and Psychophysics, 56*(4), 414-423. doi:10.3758/Bf03206733

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. New York: Oxford University Press.

Malmberg, K. J., & Annis, J. (2012). On the relationship between memory and perception: Sequential dependencies in recognition memory testing. *Journal of Experimental Psychology: General, 141*(2), 233. doi:10.1037/a0025277

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception: An Interdisciplinary Journal, 22*(4), 663-714.

McAuley, J. D., Stevens, C., & Humphreys, M. S. (2004). Play it again: did this melody occur more frequently or was it heard more recently? The role of stimulus familiarity in episodic recognition of music. *Acta Psychol (Amst), 116*(1), 93-108. doi:10.1016/J.Actpsy.2004.02.001

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review, 105*(4), 724-760. doi:10.1037/0033-295x.105.4.734-760

Milne, A. J., Bulger, D., Herff, S. A., & Sethares, W. A. (2015). Perfect Balance: A Novel Principle for the Construction of Musical Scales and Meters *Mathematics and Computation in Music* (pp. 97-108). London: Springer.

Milne, A. J., Sethares, W., & Plamondon, J. (2007). Isomorphic controllers and dynamic tuning: Invariant fingering over a tuning continuum. *Computer Music Journal, 31*(4), 15-32. doi:10.1162/comj.2007.31.4.15

Moelants, D. (2002). *Preferred tempo reconsidered.* Paper presented at the Proceedings of the 7th international conference on music perception and cognition.

Müllensiefen, D. (2009). *FANTASTIC: Feature ANalysis Technology Accessing STatistics (In a Corpus): Technical Report*. Retrieved from

Müllensiefen, D., & Halpern, A. R. (2014). The Role of Features and Context in Recognition of Novel Melodies. *Music Perception,* 418-435. doi:10.1525/MP.2014.31.5.418.

Nathoo, F. S., & Masson, M. E. J. (2016). Bayesian alternatives to null-hypothesis significance testing for repeated-measures designs. *Journal of Mathematical Psychology,* 144-157.

Nickerson, R. S. (1965). Short-Term-Memory for Complex Meaningful Visual Configurations - a Demonstration of Capacity. *Canadian Journal of Psychology, 19*(2), 155-160. doi:DOI 10.1037/h0082899

Norman, D. A. (2013). *Models of human memory*. New York: Elsevier.

ACCEPTED MANUSCRIPT

O'Connor, B. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test.

. *Behavior Research Methods, Instrumentation, and Computers,, 32*, 396-402. doi:10.3758/Bf03200807

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: a computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic Bulletin & Review, 18*(1), 10-45. doi:10.3758/S13423-010-0020-6

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19*(5), 779-819. doi:10.3758/S13423-012-0272-4

Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings.* The Ohio State University.

Olson, G. M. (1969). Learning and retention in a continuous recognition task. *Journal of Experimental Psychology, 81*(2), 381-384. doi:10.1037/h0027756

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*(3), 241-263. doi:10.1037/h0027272

Pearce, M. T. (2014). IdyOM Project.   Retrieved from https://code.soundsoftware.ac.uk/projects/idyom-project

Peretz, I., & Gaudreau, D. (1998). Exposure effects on music preference and recognition. *Memory and Cognition, 26*(5), 884-902. doi:10.3758/BF03201171

Plantinga, J., & Trainor, L. J. (2005). Memory for melody: infants use a relative pitch code. *Cognition, 98*(1), 1-11. doi:10.1016/J.Cognition.2004.09.008

ACCEPTED MANUSCRIPT

Poon, L. W., & Fozard, J. L. (1980). Age and Word-Frequency Effects in Continuous Recognition Memory. *Journals of Gerontology, 35*(1), 77-86. doi:10.1093/geronj/35.1.77

Psychology Software Tools, I. (2012). [E-Prime 2.0]. Retrieved from Retrieved from http://www.pstnet.com

R-Core-Team. (2013). R: A language and enviornment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rakover, S. S., & Cahlon, B. (2001). *Face recognition: Cognitive and computational processes* (Vol. 31). Amsterdam: John Benjamins Publishing.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012a). Word Perception I: Some Basic Issues and Methods *Psychology of reading* (pp. 49-88): Psychology Press.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012b). Word Perception II: Word Identification in Text *Psychology of reading* (pp. 135-162): Psychology Press.

Roberts, L. A. (1986). Modality and Suffix Effects in Memory for Melodic and Harmonic Musical Materials. *Cognitive Psychology, 18*(2), 123-157. doi:Doi 10.1016/0010-0285(86)90010-1

Rousseeuw, P. J. (1987). Silhouettes - a Graphical Aid to the Interpretation and Validation of Cluster-Analysis. *Journal of computational and applied mathematics, 20*, 53-65. doi:10.1016/0377-0427(87)90125-7

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2014). How we forget may depend on how we remember. *Trends Cogn Sci, 18*(1), 26-36. doi:10.1016/J.Tics.2013.10.008

Sapp, C. S. (2005). *Online Database of Scores in the Humdrum File Format.* Paper presented at the ISMIR.

ACCEPTED MANUSCRIPT

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58*(1), 75-125.

Schellenberg, E. G., & Habashi, P. (2015). Remembering the melody and timbre, forgetting the key and tempo. *Memory and Cognition*. doi:10.3758/s13421-015-0519-1

Schellenberg, E. G., Stalinski, S. M., & Marks, B. M. (2014). Memory for surface features of unfamiliar melodies: independent effects of changes in pitch and tempo. *Psychological Research, 78*(1), 84-95. doi:10.1007/s00426-013-0483-y

Schneider, A. (1997). “Verschmelzung”, tonal fusion, and consonance: Carl Stumpf revisited. In M. Leman (Ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology*. New York: Springer.

Schulkind, M. D. (2009). Is Memory for Music Special? *Neurosciences and Music III: Disorders and Plasticity, 1169*, 216-224. doi:10.1111/J.1749-6632.2009.04546.X

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics, 6*(2), 461-464.

Shepard, R. N., & Teghtsoonian, M. (1961). Retention of Information under Conditions Approaching a Steady-State. *Journal of Experimental Psychology, 62*(3), 302-309. doi:10.1037/H0048606

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*(2), 145-166. doi:10.3758/BF03209391

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology-General, 117*(1), 34-50. doi:10.1037//0096-3445.117.1.34

ACCEPTED MANUSCRIPT

Standing, L. (1973). Learning 10,000 Pictures. *Quarterly Journal of Experimental Psychology, 25*(May), 207-222. doi:10.1080/14640747308400340

Stevens, C. J. (2015). Is memory for music special? *Memory Studies, 8*(3), 263-266. doi:10.1177/1750698015584873

Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory and Cognition, 35*(4), 628-639.

Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition, 74*(3), B27-B32.

Tulving, E. (1985). Memory and Consciousness. *Canadian Psychology-Psychologie Canadienne, 26*(1), 1-12. doi:10.1037/H0080017

Vanstone, A. D., Cuddy, L. L., Duffin, J. M., & Alexander, E. (2009). Exceptional preservation of memory for tunes and lyrics: case studies of amusia, profound deafness, and Alzheimer's disease. *Annals of the New York Academy of Sciences, 1169*, 291-294. doi:10.1111/j.1749-6632.2009.04763.x

Vicary, S. A., Robbins, R. A., Calvo-Merino, B., & Stevens, C. J. (2014). Recognition of dance-like actions: Memory for static posture or dynamic movement? *Memory and Cognition, 42*(5), 755-767. doi:10.3758/S13421-014-0395-0

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14*(5), 779-804. doi:10.3758/Bf03194105

Wilks, S. S. (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics, 9*(1), 60-62. doi:10.2307/2957648

ACCEPTED MANUSCRIPT

Williams, T. I. (2015). The classification of involuntary musical imagery: The case for earworms. *Psychomusicology: Music, Mind, and Brain, 25*(1), 5.

Winold, A. (1975). Rhythm in twentieth-century music. In G. Wittlich (Ed.), *Aspects of Twentieth-Century Music* (pp. 208-269). Englewood Cliffs, New Jersey: Prentice-Hal.

Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R. S., Klatzky, R., . . . Merfeld, D. M. (2012). *Chapter 4: Perceiving and Recognizing Objects* (Fourth Edition ed.). Sunderland: Sinauer Associates, Inc.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441-517. doi:10.1006/Jmla.2002.2864

ACCEPTED MANUSCRIPT

Figure Captions

Figure 1: Schematic example of three trials in Experiment 1, with ‘Melody’ representing a stimulus and ‘Old? New?’ representing a participant’s response. One melody and one response constituted one trial. The two grey fields in the figure represent the same melody. Therefore, this shows an example of one intervening melody between the first presentation of the ‘grey’ melody and its repetition. Here, we systematically manipulate the number of intervening melodies until a target melody repeats (see text for details).

Figure 2: Hit rates and false alarm rates in Experiment 1. The left panel shows mean data participant-wise, and the right panel shows mean data melody-wise. The reference line represents chance level. Mixed effects models were built to take inter-melody and inter-participant variation into account.

Figure 3. Mean bias corrected hit rate for all seven conditions of intervening melodies in Experiment 1. Note that zero represents chance recognition level. The zero intervening melodies condition produced significantly higher recognition performance than all other conditions. Performance was statistical identical in one through to six intervening melodies. Error bars show 95% confidence intervals.

Figure 4: Hit rates and false alarm rates in Experiment 2. The left panel shows mean data participant-wise, and the right panel shows mean data melody-wise. The reference line represents chance level.

Figure 5. Mean bias corrected hit rate for all 10 conditions of intervening melodies in Experiment 2. Note that zero represents chance level. Recognition performance was statistically identical in all interveing melody conditions. Error bars show 95% confidence intervals.

Figure 6: Schematic example of three trials in Experiment 3. One melody and one response constituted one trial. The two grey fields represent the same melody; however, the second presentation of each melody (the ‘grey’ melody in this figure) in Experiment 3 was transposed. Therefore, this shows an example of one intervening melody between the first presentation of the ‘grey’ melody and its transposed repetition.

ACCEPTED MANUSCRIPT

Figure 7: Hit rates and false alarm rates in Experiment 3. The left panel shows the data participant-wise, and the right panel melody-wise. The reference line represents chance level.

Figure 8: Prediction line of a generalized mixed effects model that predicts the probability of bias corrected recognition (y-axis). Performance was significantly above chance. However, the number of intervening items did not carry predictive value. The grey area around the prediction line represents a 95% confidence interval.

Figure 9: Perceived familiarity raw data for Experiment 4. The left panel shows participant-wise differences in perceived familiarity between the first and second occurrence of a melody. The right panel shows melody-wise differences. Second presentations of melodies elicit significantly higher familiarity ratings than first presentations.

Figure 10: Raw data and prediction line of a linear mixed effects model predicting perceived familiarity on the second occurrence of a melody in Experiment 4, based on the number of intervening items. Familiarity increased significantly between first and second presentation of a melody. However, the number of intervening items did not carry predictive value. The grey area around the prediction line represents a 95% confidence interval.

ACCEPTED MANUSCRIPT

Appendix A

Representative Examples of Melodies used in all Experiments

Experiments 1 and 2

Figure A1 about here

Experiments 3 and 4

Figure A2 about here

ACCEPTED MANUSCRIPT

Appendix B

Uncorrected Hit Rates

Experiment 1

Figure B1 about here

Figure B.1: Uncorrected hit rates in Experiment 1. In our analysis, we corrected the hit rates to capture and account for participants' response biases, as well as potential shifts of these response biases over the course of the experiment.

ACCEPTED MANUSCRIPT

Experiment 3

Figure B2 about here

Figure B.2: Uncorrected hit rates in Experiment 3. In our analysis, we corrected the hit rates to capture and account for participants’ response biases, as well as potential shifts of these response biases over the course of the experiment. When comparing Figure B.2 with Figure 8, performance appears better in the uncorrected hit rates and performance seems to significantly increase over the course of the experiment. However, this is due to systematic response-tendency shifts (see further detail in Appendix C). When the analysis takes these response-tendency shifts into account, then performance is stable across all intervening item conditions (cf. Figure 8).

ACCEPTED MANUSCRIPT

Appendix C

Results from the Dynamic Response Tendencies Analysis

Participants' response tendencies can change over the course of an experiment. In the main statistical analyses presented throughout the paper, any shifts in response tendencies were taken into account by building *Dynamic Response Tendency* models that predict the baseline tendency for each participant to press 'old', and how this tendency changes throughout each experiment. To do this, a generalized mixed effects model was trained to predict 'old' responses on first melody presentations (False Alarms) based on trial number. The *Dynamic Response Tendency* models were then used to predict responses on second presentations of each melody. The predictions of the models were then included as a fixed factor in the main statistical analyses. Results of *Dynamic Response Tendency* models are summarized here: Experiments 1, 2, and 4 did not exhibit significant dynamic response tendency shifts (all coefficients *p*-values > .20). However, Experiment 3 did show significant dynamic response tendency shifts over the course of the experiment ($p < .001$, coefficient = 2.59). As can be seen in Figure C1, the average probability of producing false alarms increased in Experiment 3 as trial numbers increased. This is likely due to the design of Experiment 3's continuous recognition task, which presented nearly twice as many trials as Experiments 1 and 2.

Figure C1 about here

ACCEPTED MANUSCRIPT

Figure C1: Average dynamic response tendency shift in Experiment 3. The graph shows how the probability of producing a false alarm increased with trial number. The dynamic response tendency shift was controlled in the statistical analyses and implemented as a predictor in the mixed effects models reported in the main Results sections. The grey area around the prediction line represents a 95% confidence interval.

ACCEPTED MANUSCRIPT