ELSEVIER

Acta Psychologica 116 (2004) 93–108

acta psychologica

www.elsevier.com/locate/actpsy

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 ☆

J. Devin McAuley [a,*], Catherine Stevens [b], Michael S. Humphreys [c]

[a] *Department of Psychology, Bowling Green State University, Bowling Green, OH 43403, USA*
[b] *MARCS Auditory Laboratories, The University of Western Sydney, New South Wales, Australia*
[c] *School of Psychology, The University of Queensland, Queensland, Australia*

Received 25 August 2003; received in revised form 4 February 2004; accepted 6 February 2004

Abstract

Episodic recognition of novel and familiar melodies was examined by asking participants to make judgments about the recency and frequency of presentation of melodies over the course of two days of testing. For novel melodies, recency judgments were poor and participants often confused the number of presentations of a melody with its day of presentation; melodies heard frequently were judged as have been heard more recently than they actually were. For familiar melodies, recency judgments were much more accurate and the number of presentations of a melody helped rather than hindered performance. Frequency judgments were generally more accurate than recency judgments and did not demonstrate the same interaction with musical familiarity. Overall, these findings suggest that (1) episodic recognition of novel melodies is based more on a generalized "feeling of familiarity" than on a specific episodic memory, (2) frequency information contributes more strongly to this generalized memory than recency information, and (3) the formation of an episodic memory for a melody depends either on the overall familiarity of the stimulus or the availability of a verbal label.

☆ A portion of this research was conducted while the second author was an Australian Research Council Postdoctoral Fellow at the University of Queensland.
* Corresponding author. Tel.: +1-419-372-0285; fax: +1-419-372-6013.
E-mail address: mcauley@bgnet.bgsu.edu (J.D. McAuley).

0001-6918/$ - see front matter
doi:10.1016/j.actpsy.2004.02.001

PsycINFO classification: 2343
Keywords: Episodic memory; Music; Melody recognition

1. Introduction

Memory, for music in particular, can be quite powerful, evoking a strong sense of nostalgia. Although there have been many investigations of *what* we remember about the music we experience (Dowling, 1978; Dowling & Bartlett, 1981; Halpern, 1984; Hebert & Peretz, 1997; Jones & Ralston, 1991), there is relatively little known about *how* we remember it. This article is concerned with the latter question. Specifically, we are interested in factors that affect the formation of episodic memories for melodies.

Suppose that yesterday you heard the opening bars of "Twinkle, Twinkle, Little Star" played on the piano. Today, it is played again and you are asked to decide whether you have ever heard it before. For a familiar tune, such as this, it is likely that you have heard it played in different keys in a variety of styles. To simply recognize a melody, it is not necessary to recall specific memories of it. Instead, all that is required is a judgment about whether or not you know the melody. Consider a second, but related question: do you remember hearing "Twinkle, Twinkle, Little Star" yesterday? This question is different in at least one important way. The instructions are to make a memory decision that is restricted to yesterday's events. The unresolved problem for memory theorists is how this feat is accomplished. One issue that arises is that it is not always clear when such a decision requires an episodic memory (one that is specific to yesterday's events) or when it may be based instead on a generalized memory or "feeling of familiarity" (Humphreys, Bain, & Pike, 1989).

When a novel stimulus is first encountered, any memory for that stimulus will be by definition unique to that episode (Tulving, 1972). However, we will refer to the memory established by the first occurrence of a stimulus as *non-episodic* if memories for subsequent encounters are not differentiated from the first memory. This definition is in keeping with the spirit of Tulving's (1972) definition of episodic memory and is neutral with respect to the proposal that separate episodic and semantic memory systems are involved.

Many researchers have assumed that contextual cues play an important role in the retrieval of an episodic memory (Anderson & Bower, 1972; Bain & Humphreys, 1988; Dennis & Humphreys, 2001; Gillund & Shiffrin, 1984; Tulving, 1983). However, just what constitutes context and how it provides relevant information is not well understood, even though context effects have been the focus of extensive investigation in memory research (Bower, 1981; Fernandez & Glenburg, 1985; Godden & Baddeley, 1980; Humphreys, 1976). Dalton (1993) (see also Glenburg, 1979) distinguishes between two types of contextual manipulations (global and local) in recognition studies. According to this distinction, global context would include physical environment, mood, and stimulus characteristics that do not change (or change infrequently) during the study episode and therefore might potentially aid source discrimination. Local context cues, in contrast, would include stimuli that co-vary (or are associated) with target items (e.g., the word "strawberry" paired with "jam" in a list of to-be-remembered word pairs), and therefore might potentially help in the retrieval of particular items from memory.

In the music domain, researchers have primarily considered the role of local context cues in recognition, including the particular lyrics associated with a given melody (Serafine, Crowder, & Repp, 1984) and the rhythm through which a melody is expressed (Jones & Ralston, 1991). Much of the research in this area concerns the degree of independence of memory representations of the various components, such as melody and lyrics, of a musical event (Crowder, Serafine, & Repp, 1990; Hebert & Peretz, 1997; Jones & Ralston, 1991; Samson & Zatorre, 1991). Overall, these studies have shown that melody recognition is generally much poorer than lyric/word recognition and that the interference between the two is asymmetric; lyrics impact on melody recognition, but not the other way around (Samson & Zatorre, 1991; Serafine et al., 1984). One difference between the melodies and lyrics used in these studies is that the words comprising the lyrics are typically familiar to participants prior to the experiment, whereas the melodies are typically not (i.e., the melodies are novel).

One possible reason for poor recognition of novel melodies may be that people have difficulty encoding novel melodies, or encode them differently between study and test. [1] A second possible reason for poor recognition of novel melodies may be that people's memories are non-episodic. A common assumption in recognition studies using the study-test paradigm is that participants restrict their recognition judgments to the study episode, but this does not have to be the case. This assumption has been incorporated in many models of memory as a context cue that restricts retrieval to only those items in the study list (Clark & Gronlund, 1996; Dennis & Humphreys, 2001; Gillund & Shiffrin, 1984; Humphreys et al., 1989). However, in the absence of a context cue, or similar source information to isolate the study episode, people may be forced to rely on whatever information is available to make a recognition decision about a test item. This might be the verbal context (such as the lyrics presented with a melody), for which more accurate source information is available, or it might be a more generalized "feeling of familiarity" that increases in strength with stimulus repetition and collapses over study episode (Chalmers & Humphreys, 1998; Humphreys et al., 1989). We will hereafter refer to recognition based on an overall feeling of familiarity as a generalized memory.

Some direct evidence for the use of a generalized memory in the recognition of novel stimuli comes from a series of studies conducted by Chalmers and Humphreys

[1] If this was occurring it seems likely that musically trained participants should encode melodies more consistently than musically untrained participants. There is some evidence to support this claim. Halpern, Bartlett, and Dowling (1995) report musical-training differences in the ability to recognize transpositions of novel melodies. More recently, Andrews, Dowling, Bartlett, and Halpern (1998) report musical-training differences in the identification of melodies presented at different tempos (rates), although these melodies were familiar, rather than novel.

(1998) examining judgments about the recency-of-presentation and number-of-presentations of unfamiliar words (the verbal analogue of a novel stimulus). Prior to the study-test sequence subjects had been given familiarization trials on some of the words. During familiarization training, words were either presented with their definitions or without definitions. For words presented without their definition, recency judgments were poor and interacted with how many times that word had been experienced; more frequent words were judged more recent, suggesting that judgments about recency were based on a general measure of memory strength, rather than a specific episodic memory for the unfamiliar word.

2. Overview

The current study applied the Chalmers and Humphreys' design (adapted from Huppert & Piercy, 1978, Experiment 2) to musical stimuli in order to compare memory for novel and familiar melodies. Participants listened to novel and familiar melodies on one of two days either once or three times. Following the exposure phase on the second day, which followed the first day by 24 h, people were given two tests. For the recency test, participants discriminated between melodies they last heard on that day (same-day melodies) and those they last heard on the previous day (previous-day melodies), by responding "Today" or "Yesterday" to each melody. For the frequency test, participants discriminated between melodies they heard three times and those they heard only once, by responding "Three times" or "once" to each melody.

2.1. Generalized-memory hypothesis

If people base decisions about frequency and recency on a generalized memory that increases in strength with item repetition (frequency) and collapses over study episode (recency), then crossing frequency and recency is potentially detrimental to performance. That is, from the perspective of a generalized memory, melodies presented more frequently have the potential to be judged as having occurred more recently than they actually were, and melodies presented more recently have the potentially to be judged as having occurred more frequently than they actually were. For the present study, this translated to expectations that the proportions of "Today" and "Three times" responses would be largest for same-day melodies presented three times, smallest for previous-day melodies presented only once, and intermediate for same-day melodies presented once, and previous-day melodies presented three times.

Two interesting asymmetries are also possible with a generalized memory. If frequency contributes more to overall memory strength than does recency (e.g., there is little or no forgetting between the two study episodes), then, for the recency task, there should be more false alarms to previous-day melodies presented three times than correct "Today" responses to same-day melodies presented once. This would leave performance on the frequency task relatively intact. If, on the other hand, recency contributes more strongly to memory strength than does frequency (e.g., a sharply negatively accelerated learning function so that there is little difference between one and three presentations), then for the frequency task, there should be more false alarms to same-day melodies presented once, than correct "Three times" responses to previous-day melodies presented three times. This, conversely, would leave performance on the recency task relatively intact.

2.2. Episodic-memory hypothesis

If people base decisions about frequency and recency on an episodic memory (i.e., one that does not collapse over study episode) then crossing frequency and recency should not be detrimental to performance. That is, from the perspective of an episodic memory, the more presentations of a melody during a study episode, the better people should be at judging "when" they heard that melody; similarly, the more recently a melody has been presented, the better people should be at judging "how many times" that melody was presented. For the present study, this translates into expectations that the pattern of recency and frequency responses should be very different from that based on a generalized memory. In particular, for the recency task, increasing the number of presentations of a melody should (1) increase (or not change) the proportion of correct "Today" responses and (2) decrease (or not change) false alarms (responding "Today" to a melody studied "Yesterday"). An analogous pattern of "Three times" responses should occur for the recency manipulation in the frequency task.

3. Method

3.1. Participants

Participants were recruited from first-year psychology courses at the University of Queensland and the University of Western Sydney (age range: 17–40 years, mean: 20.5 years). Musically trained participants ($n = 42$) had at least six years formal music experience, either instrumental or voice (mean: 8.7 years). Untrained participants ($n = 52$) had less than two years formal training (mean: 0.9 years).

3.2. Design

The present study implemented a $2 \times 2 \times 2 \times 2$ mixed factorial design. The within-subject variables were melody type (novel versus familiar), day-of-presentation (same day versus previous day), and number-of-presentations (once versus three times). The between-subjects variable was musical training (trained versus not trained). The presentation order of the separate tests and the melodies used in each test were counterbalanced between participants, so that collectively, participants made recency and frequency judgments about each melody in each of the eight conditions.

3.3. Materials

The stimuli consisted of 24 familiar and 24 novel melodies. The familiar melodies (see Table 1) were a subset of those previously used by Halpern (1984) and DeWitt and Samuel (1990). They contained between 8 and 33 notes (mean = 15.6), were in a range of different major and minor keys, and used a variety of different rhythms. The novel melodies were composed for the purposes of the experiment, with the aim of making them at least as distinctive as the familiar melodies; they were single-note melodies, composed in a range of different major and minor keys using a variety of rhythms, speeds, melodic contours, and articulation styles. Each contained between 7 and 17 notes (mean = 12.3).

3.4. Equipment

The novel and familiar melodies were played on a Casio CT-670 stereo MIDI keyboard using a synthesized piano voice, and sequenced randomly onto audiocassette with a five-second inter-stimulus interval between melodies. There were eight different tapes (one for each of the different counterbalancing orders) and no melody ap-

Table 1
List of familiar melody titles with mean familiarity ratings (pre- and post-test) and probability of listeners correctly recalling the associated title or lyric

Familiar melody	Familiarity rating (max. = 2.00)		Probability recall of title/lyric
	Pre	Post	
Baby-face	1.64	1.66	0.53
Daisy, Daisy	1.27	1.26	0.30
Deck the Halls	1.98	1.92	0.86
Doe, a Deer	1.94	1.93	0.88
For He's a Jolly Good Fellow	1.97	1.99	0.89
Happy Birthday	1.98	2.00	0.95
White Christmas	1.44	1.71	0.72
Jingle Bells	1.98	1.98	0.95
Oh Come All Ye Faithful	1.97	1.97	0.91
Old MacDonald	2.0	1.97	0.91
Pop Goes the Weasel	1.94	1.92	0.89
Raindrops Keep Falling on My Head	1.89	1.93	0.91
Row, Row, Row, your Boat	1.92	1.92	0.89
Rudolf, the Red-Nosed Reindeer	1.99	1.99	0.95
Seventy-six Trombones	1.16	1.05	NA
She'll be Coming 'Round the Mountain	1.83	1.89	0.91
Silent Night	1.93	1.96	0.89
Singing in the Rain	1.63	1.68	0.82
Somewhere Over the Rainbow	1.78	1.94	0.93
Twinkle, Twinkle, Little, Star	1.93	1.93	0.84
We Wish You a Merry Christmas	1.99	1.94	0.88
When the Saints Go Marching In	1.71	1.89	0.89
Yesterday	1.69	1.77	0.63
As Time Goes By	0.27	0.57	0.09

peared twice in succession on any of the tapes. The experimental sessions were run in small groups and participants listened to melodies through loudspeakers at a comfortable listening level.

3.5. Procedure

The experiment consisted of two parts: an exposure phase and a test phase. During the exposure phase, participants rated the familiarity of all 48 melodies on a three-point scale: 0 for very unfamiliar—never encountered, 1 for familiar—sometimes encountered, and 2 for very familiar—encountered frequently. The exposure phase was split into two experimental sessions; the second session followed the first session by 24 h. In each session, participants heard 24 melodies, of which 12 were novel and 12 were familiar, with half of the novel and familiar melodies presented three times and the other half presented once, with the melodies randomly ordered within each counterbalancing condition.

After the presentation of the second set of melodies in the second session (Day 2), participants were given two memory tests, following a 10-min filled retention interval. One test asked participants to make recency judgments and the other test asked participants to make frequency judgments. Each test involved presenting 24 melodies. Participants were instructed that all of the melodies had been heard earlier, either earlier in that session (same day), or in the first session (previous day). They were also told that half the melodies had been presented once and half had been presented three times. For the recency test, participants were asked to respond "Today" to melodies they believed they last heard on that day and "Yesterday" to melodies they believed they last heard on the previous day, and to guess when in doubt. For the frequency test, participants were asked to respond "Once" or "Three times" to melodies they believed they heard once or three times, and to guess when in doubt. After participants completed both tests, all 48 melodies were presented a final time and were rated on the same three-point familiarity scale; participants also wrote down any title, lyric, movie or event they associated with each melody.

The first session lasted approximately 15 min and the second session lasted approximately 50 min. In the remainder of the paper, the two sets of melodies (comprising the first and second sessions) will be referred to as previous-day and same-day melodies, respectively, because of their temporal relation to the two memory tasks.

4. Results

Mean familiarity ratings for the study (exposure) phase for the familiar and novel melodies were 1.74 and 0.34, respectively, confirming their classification as familiar and novel stimuli. One clear exception was the classification of "As Time Goes By", which received a mean pre-test rating of only 0.27. Three other familiar melodies received relatively low mean ratings: "Daisy, Daisy" (1.27), "Seventy-Six Trombones" (1.16), and "White Christmas" (1.44). All other mean ratings for the familiar melodies were above 1.6 and for three of the four familiar melodies that received relatively low ratings these ratings were still much higher than the mean ratings for the novel melodies. A preliminary analysis of the proportion of "Three times" and "Today" responses that compared inclusion and exclusion of the familiar melodies that received low ratings revealed no obvious differences; thus, all familiar melodies were retained for the analyses reported below.

Recency and frequency task performance was examined by comparing the proportions of "Today" and "Three times" responses for the within-subject variables of day-of-presentation (same-day versus previous-day), number of presentations (three times versus once), and melody type (novel versus familiar) and the between-subject musical-training variable. For the recency task, a hit meant responding "Today" to a same-day melody, and a false alarm meant responding "Today" to a previous-day melody. Similarly, for the frequency task, a hit meant responding "Three times" to melodies presented three times and a false alarm meant responding "Three times" to melodies presented once. [2]

Table 2 summarizes the results of the recency and frequency tasks. Mean proportions of "Today" and "Three times" responses for the novel and familiar melodies are shown, with standard errors reported in parentheses. Hits rates are indicated in bold. An ANOVA revealed no main effect or interactions for the musical-training variable for either the recency task [$F(1,92) = 1.50$, $\text{MSE} = 0.05$, $p > 0.2$] or for the frequency task [$F(1,92) = 0.65$, $\text{MSE} = 0.05$, $p > 0.4$] and so we chose to collapse over this variable in reporting the analyses. [3] Below, we describe the results separately for the two tasks in terms of hits and false alarms.

4.1. Recency performance

An ANOVA on the proportion of "Today" responses showed a significant main effect of day-of-presentation [$F(1,93) = 142.2$, $\text{MSE} = 0.095$, $p < 0.01$]; collapsed over the number of presentations (one versus three) and melody type (familiar versus novel), the mean hit rate of 0.61 and the mean false-alarm rate of 0.35 indicated that participants could generally perform the recency task. However, recency judgments about the familiar melodies were much better than for the novel melodies. Comparing the overall hit rates for familiar and novel melodies shows that the mean hit rate

[2] A signal detection analysis was not performed on these data because there were too few items per condition per subject to provide accurate estimates of d'. As an alternative to d' analysis for measuring sensitivity, we simply compared differences between the hit and false-alarm rates, rather than first transforming these to z-scores, as described by Snodgrass and Corwin (1988).

[3] Given previous demonstrated effects of musical training on recognition performance (Halpern et al., 1995), our lack of an effect of musical training was a little bit surprising. However, the criterion used to separate musically trained and untrained participants was not that stringent. We classified a participant as musically experienced if they had at least 6 years of formal musical training. This may not have been sufficient enough to demonstrate differences in memory performance due to musical training. In addition, some participants may have had at least 6 years of training, but that training may have occurred many years prior to their experiment participation. A better index of musical training might have been the number of hours each week that a participant *currently* practices or performs.

Table 2
Mean proportions of "Today" and "Three times" responses for novel and familiar melodies for the four combinations of recency and frequency: hit rates are in bold

Response	Melody	Study condition			
		Previous day, Once	Previous day, Three times	Same day, Once	Same day, Three times
"Today"	Novel	0.38 (0.03)	0.59 (0.03)	**0.46 (0.03)**	**0.61 (0.03)**
	Familiar	0.21 (0.02)	0.21 (0.03)	**0.64 (0.03)**	**0.76 (0.03)**
"Three times"	Novel	0.25 (0.03)	**0.66 (0.03)**	0.20 (0.02)	**0.53 (0.03)**
	Familiar	0.30 (0.03)	**0.77 (0.03)**	0.32 (0.03)	**0.80 (0.02)**

Data are collapsed across musical training, with standard error reported in parentheses ($n = 94$).

for the familiar melodies was 0.69, but was only 0.53 for the novel melodies; closer examination of the novel melody data shows that only the proportion of "Today" responses to same-day novel melodies presented "Three times" was significantly greater than 0.5 [$t(93) = 3.61$, $p < 0.01$]. Comparing false-alarm rates reveals the opposite pattern; the false-alarm rate for familiar melodies was 0.21, which increased to 0.48 for the novel melodies; in this case, closer inspection of the novel melody data shows that only the proportion of "Today" response to previous-day novel melodies presented once was significantly less than 0.5 [$t(93) = -3.78$, $p < 0.01$].

The low level of recency performance for the novel melodies can be attributed to interference with the number of presentations of a melody. As shown in Table 2, increasing the number of presentations of novel same-day and previous-day melodies increased *both* hit rates and false-alarm rates. Post hoc analysis of novel melody performance revealed significant differences between comparisons of (1) previous-day melodies presented once and previous day melodies presented three times [$t(93) = -4.55$, $p < 0.01$], (2) previous-day melodies presented once and same-day melodies presented three times [$t(93) = -5.36$, $p < 0.01$], (3) previous-day melodies presented three times and same-day melodies presented once [$t(93) = 2.73$, $p < 0.01$], and (4) same-day melodies presented once and same-day melodies presented three times [$t(93) = -3.30$, $p < 0.01$].

Overall, recency judgments about the novel melodies were most consistent with the generalized-memory hypothesis; the smallest proportion of "Today" responses (0.38) occurred for the novel previous-day melodies presented once, the largest proportion of "Today" responses (0.61) occurred for novel same-day melodies presented three times, with intermediate values obtained for novel same-day melodies presented once (0.46) and novel previous-day melodies presented three times (0.59). It is worth noting that the interference with the number of presentations of a melody was so strong that participants made more incorrect "Today" responses to previous-day melodies presented three times than they made correct "Today" responses to same-day melodies presented once [$t(93) = 2.73$, $p < 0.01$].

The results for the familiar melodies contrasted with those for the novel melodies; for familiar melodies, increasing the number of presentations of same-day melodies from one to three increased the hit rate from 0.64 to 0.76, respectively [$t(93) = -2.80$, $p < 0.01$], but for previous-day melodies held the false-alarm rate constant ($M = 0.21$). Post hoc tests revealed reliable differences between all pairwise comparisons (all p's < 0.01) except once and thrice presented previous-day melodies [$t(93) = -0.32$, $p = 0.75$]. Overall, recency judgments about familiar melodies were most consistent with the episodic-memory hypothesis.

One question that arose during the analysis was whether there was an overall bias to respond "Today" to the familiar melodies. This was not the case. In fact, it turned out that the opposite was true. The proportion of "Today" responses was somewhat smaller for familiar melodies than for novel melodies (0.45 versus 0.50), $F(1,93) = 4.76$, MSE = 0.122, $p < 0.05$. Inspection of Table 2 suggests that the "Today" bias for novel melodies is partly due to the very high false-alarm rate to novel melodies presented three times on the previous day. We will return to this point in the discussion.

4.2. Frequency performance

An ANOVA on the proportion of "Three times" responses showed a significant main effect of number of presentations [$F(1,93) = 363.82$, MSE = 0.09, $p < 0.01$]; collapsed over day of presentation (previous-day versus same-day) and melody type (familiar versus novel), the mean hit rate of 0.69 and the false-alarm rate of 0.27 indicated that participants could generally perform the frequency task. However, participants' frequency judgments were more accurate than their recency judgments and the effect of melody type was different.

For the frequency task, familiar melodies produced more hits *and* false alarms than did novel melodies [$F(1,93) = 25.5$, MSE = 3.52, $p < 0.01$] and the effect was more pronounced for familiar melodies presented on the same day than on the previous day [$F(1,93) = 8.00$, MSE = 0.07, $p < 0.01$] and more pronounced for familiar melodies presented three times than for those presented once [$F(1,93) = 7.02$, MSE = 0.07, $p < 0.01$]. For the familiar melodies, post hoc t-tests revealed significant differences for all pairwise comparisons (all p's < 0.01) except the once presented previous-day and same-day conditions [$t(93) = -0.426$, $p = 0.67$] and thrice presented previous-day and same-day conditions [$t(93) = -0.675$, $p = 0.50$]. For the novel melodies, post hoc t-tests revealed significant differences for all pairwise comparisons (all p's < 0.01) except the once presented previous-day and same-day conditions [$t(93) = 1.52$, $p = 0.13$].

Finally, unlike the ANOVA on recency judgments, the ANOVA on frequency judgments did not indicate a significant main effect of day of presentation; proportions of "Three times" responses for same-day and previous-day melodies were not reliably different: 0.49 versus 0.46, respectively [$F(1,93) = 3.1$, MSE = 0.07, $p = 0.08$]. There were no other significant main effects or interactions.

5. Discussion

The primary question addressed by this research was the role of stimulus familiarity in episodic recognition of music. Overall, very different patterns of recency and frequency responses were observed with familiar and novel melodies. Familiar melodies afforded accurate judgments about *both* recency and frequency, whereas novel melodies afforded accurate judgments about frequency, but not recency. The results of this study support the use of an episodic memory in the recognition of familiar melodies, but the use of a generalized memory in the recognition of novel melodies.

Support for the use of a generalized memory in the recognition of the novel melodies derives from the pattern of interference that occurred between number of presentations and day of presentation. As predicted by the generalized-memory hypothesis, the smallest proportion of "Today" responses occurred for previous-day melodies presented once, the largest proportion of "Today" responses occurred for same-day melodies presented three times, with intermediate values obtained for same-day melodies presented once and previous-day melodies presented three times. The obtained interference was so strong that novel melodies heard three times yesterday tended to elicit more "Today" responses than novel melodies that had been presented on that day, but only once. The analogous pattern of interference was not found between number of presentations and day of presentation for the frequency judgment task. This later finding is consistent with one of the two possible asymmetries we outlined in the description of the generalized-memory hypothesis; based on the present study, it appears that frequency information contributes more strongly to overall memory strength than does recency information.

The main support for the use of an episodic memory in the recognition of the familiar melodies derives from the overall improvement observed in recency performance (compared with the novel melodies) and the lack of interference between number of presentations and day of presentation for *both* the recency and frequency tasks. In contrast to the novel melodies, the number of presentations of a familiar melody improved judgments about *when* a melody was last heard, rather than interfering with it; increasing the number of presentations from one to three increased the hit rate, but did not change the false-alarm rate. The most striking difference was the false-alarm rate to melodies presented "Three times" yesterday. For the novel melodies, the false-alarm rate was 0.59, whereas for the familiar melodies, it was only 0.21 (which was no different from the false-alarm rate for familiar melodies presented once). A strong prediction of the episodic-memory hypothesis was that the false-alarm rate should decrease with increased repetition. However, the constant false-alarm rate found in the present study is at least consistent with an episodic memory if it is assumed (as it is in Dennis & Humphreys, 2001) that participants only successfully reinstate the same-day context and there is some similarity between same-day and previous-day contexts.

There is a clear relation between the present results and a previous study by Peretz, Gaudrea, and Bonnel (1998) involving novel and familiar melodies. Peretz and colleagues examined effects of prior exposure and retention interval on memory performance for novel and familiar melodies using a preference task and a recognition task. In this study, subjects were found to have exceptionally good recognition of novel melodies even at a 1-month retention interval, which at first glance seems to contrast with our findings. However, the recognition task used in the Peretz et al. study is more akin to our frequency task than our recency task. That is, subjects in the

Peretz et al. study were not asked to discriminate the recency of the novel melodies, only to make a recognition judgment, which could have been based on a generalized memory rather than an episodic memory. Moreover, if we compare our frequency performance for novel melodies with their recognition performance then the results of the two studies are quite consistent. In our study, we observed essentially equivalent frequency performance at 10-min and 24-h retention intervals, and so we might anticipate some memory for frequency even after a month. It is conceivable that a generalized memory is much more long lasting than an episodic memory.

The differential pattern of findings reported in this study is remarkably similar to the qualitative differences observed in previous studies comparing memory performance of individuals with and without amnesia. Using a design identical to the present study, Huppert and Piercy (1978, Experiment 2) reported that patients with Korsakoff's syndrome (an organic amnesia) were unable to make accurate recency decisions about complex picture stimuli, independent of stimulus frequency, whereas a control group could. In fact, the Korsakoff's individuals produced a pattern of "Today" responses that looks nearly identical to our novel melody recency data. Like our conclusion for the novel melodies, the conclusion from the Huppert and Piercy study was that the Korskakoff's individuals failed to form episodic memories for the pictures.

So why do novel melodies make normal individuals appear amnesic? In the present study, there are at least three possible reasons for poor recency performance for the novel melodies other than the failure to form an episodic memory. One reason may be that the novel melodies have a higher level of inter-item similarity than the familiar melodies. Previous studies of recognition memory involving novel musical stimuli have shown that listeners have difficulty discriminating between melodies with similar contours (Bartlett & Dowling, 1980; Dowling & Bartlett, 1981; Dowling & Fujitani, 1971) and with similar rhythms (Jones & Ralston, 1991; White, 1960). In the present study, we attempted to control for musical similarity by composing novel melodies with as least as many salient differences (non-overlapping contours and rhythms) as the familiar melodies, but there still may have been some differences in inter-item similarity.

A second possible reason for poor recency performance for novel melodies, but not familiar ones, is that the novel melodies may not have been as distinctive or "catchy" as the familiar melodies, and thus they were not as well attended to. Although possible, neither of these two explanations seems likely given that people were able to make accurate frequency judgments about novel melodies. If something about the composition of the novel melodies made them less distinctive, catchy, or overall more similar to each other than the familiar melodies, leading to poor judgments about recency, we would have also expected poor judgments about frequency. This was not the case.

An additional possibility is that with novel melodies participants might be impaired at episodic learning of recency but not frequency. It is certainly true that in an episodic-memory system (e.g., Dennis & Humphreys, 2001) recency judgments would be impaired if context did not discriminate the two lists or occasions. However, the ability of the participants to determine the recency of the familiar melodies shows that an inability to discriminate between contexts is not responsible. It is still possible to speculate that different kinds of context or different kinds of episodic memory are involved in recency and frequency judgments. With this assumption differential learning is possible but not very parsimonious.

Conversely, in order to explore some of the reasons why recency performance might be *enhanced* for the familiar melodies, rather than *impaired* for the novel melodies, we considered that all of the familiar melodies had potentially well-known titles, and so could have been encoded both perceptually (as melodies) *and* verbally, which may have indirectly assisted episodic recognition. By contrast, the novel melodies did not have an associated verbal label. Some support for this hypothesis was revealed by a significant, but relatively weak, positive correlation ($r = 0.53, p < 0.05$) between the probability of recalling the correct title/lyric of a melody in the post-test and participants' ability to discriminate recency (measured as hit rate minus false-alarm rate). Thus, the failure to form an episodic memory for a novel melody may be linked indirectly to not having an associated verbal label. This raises an interesting question: can people form episodic memories for familiar musical excerpts that do not necessarily have an associated verbal label? More generally, the fact that this question can be raised with respect to melodies means that it can be raised with respect to any non-verbal stimulus such as a face or an odor.

Finally, we would like to note that our argument for the involvement of a non-episodic memory in the recognition of previously unfamiliar melodies and of an episodic memory in the recognition of already familiar melodies is compatible with the involvement of both episodic and non-episodic memories in the recognition of the latter. In the dual-processing framework we might assume that performance with familiar melodies reflected the contribution of both familiarity (non-episodic) and recollection (episodic) whereas only familiarity was available for novel melodies (Jacoby, 1991). Similarly, the present results could also be interpreted in the remember/know framework as differences in remember and know responses to novel and familiar musical stimuli (see Gardiner & Radomski, 1999).

From this perspective, the present pattern of results could conceivably be interpreted in terms of differences in perceptual fluency (Jacoby, 1991; Wagner & Gabrieli, 1998). A possibility suggested by one of the reviewers is that if one interprets the question about frequency of presentation as a question about fluency strength and if fluency strength is similar for the previous-day and same-day melodies, then frequency judgments should not be affected much by day of presentation, as was found for both the novel and familiar melodies. Conversely, because the recency task is presumably not a direct measure of fluency, it is thus potentially subject to misattribution, much like the way fluency affects liking judgments (Szpunar, Schellenberg, & Pliner, in press). This misattribution may be specific to novel melodies if episodic recognition of the novel melodies is familiarity-based (consistent with the generalized-memory view) without a recollective component. An issue raised by this interpretation is the precise nature of the relationship between recognition familiarity and perceptual fluency, which is not known; see Wagner and Gabrieli (1998) for a review.

In conclusion, we feel that one of the highest priorities in future studies is to clarify whether the recognition of previously encountered stimuli from amongst stimuli that have never been encountered differs from the recognition of stimuli encountered in a particular situation from amongst familiar stimuli. In this regard, the present findings contribute to the body of evidence supporting a distinction between the recognition of novel and familiar stimuli (Chalmers & Humphreys, 1998; Dalton, 1993; Russo, Ward, Geurts, & Sheres, 1999). This distinction is important because the properties we need to explain the recognition of novel stimuli may not be the same as the properties needed to explain the recognition of already familiar stimuli. For example, our results as well as the Peretz et al. results suggest that the non-episodic (familiarity-based) memory that seems to underlie the recognition of novel melodies may be very long lasting. A long lasting familiarity would not be very useful in the recognition of already familiar stimuli because it would not differentiate stimuli that had been encountered today from stimuli that had been encountered yesterday or even last month.

A second priority should be to determine whether the intuitions about the importance of physically or mentally reinstating some of the details about the learning situation are correct (Bain & Humphreys, 1988; Dennis & Humphreys, 2001). Once we have answers to these questions it may be possible to overcome the very difficult methodological and conceptual problems associated with determining whether two processes underlie the recognition of already familiar words (Dennis & Humphreys, 2001; Humphreys, Dennis, Chalmers, & Finnigan, 2000).

Acknowledgements

The authors wish to thank Sam Glucksburg, Doug Nelson, Zehra Peynircioglu, Richard Shiffrin, John Wixted, and an anonymous reviewer for their helpful comments and suggestions for improvements to an earlier version of this manuscript.

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123.

Andrews, M. W., Dowling, W. J., Bartlett, J. C., & Halpern, A. R. (1998). Identification of speeded and slowed melodies by younger, middle-aged, and older musicians and non-musicians. *Psychology and Aging, 13*, 462–471.

Bain, J. D., & Humphreys, M. S. (1988). Relational context: Independent cues, meanings, or configurations?. In G. M. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 97–137). John Wiley and Sons.

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*, 501–515.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*, 129–148.

Chalmers, K., & Humphreys, M. S. (1998). Role of generalized and episodic specific memories in the word-frequency effect in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 610–632.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review, 3*, 37–60.

Crowder, R. G., Serafine, M. L., & Repp, B. H. (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory and Cognition, 18*, 469–476.

Dalton, P. (1993). The role of stimulus familiarity in context-dependent recognition. *Memory and Cognition, 21*, 223–234.

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

DeWitt, L. A., & Samuel, A. G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General, 119*, 123–144.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*, 341–354.

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology, 1*, 30–49.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America, 49*, 524–531.

Fernandez, A., & Glenburg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory and Cognition, 13*, 333–345.

Gardiner, J. M., & Radomski, E. (1999). Awareness of recognition memory for Polish and English folk songs in Polish and English folk. *Memory, 7*, 461–470.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.

Glenburg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory and Cognition, 2*, 95–112.

Godden, D. R., & Baddeley, A. D. (1980). When does context influence recognition memory? *British Journal of Psychology, 91*, 99–104.

Halpern, A. R. (1984). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 496–512.

Halpern, A. R., Bartlett, J. C., & Dowling, W. J. (1995). Aging and experience in the recognition of musical transpositions. *Psychology and Aging, 10*, 325–342.

Hebert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory and Cognition, 25*, 518–533.

Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory and Cognition, 4*, 221–232.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory of episodic, semantic, and procedural tasks. *Psychological Review, 97*, 208–233.

Humphreys, M. S., Dennis, S., Chalmers, K. A., & Finnigan, S. (2000). Dual processes in recognition: Does a focus on measurement operations provide a sufficient foundation. *Psychonomic Bulletin and Review*, 593–603.

Huppert, F. A., & Piercy, M. (1978). The role of trace strength in recency and frequency judgments by amnesic and control subjects. *Quarterly Journal of Experimental Psychology, 30*, 347–354.

Jacoby, L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Jones, M. R., & Ralston, J. (1991). Some influences of accent structure on melody recognition. *Memory and Cognition, 19*, 8–20.

Peretz, I., Gaudrea, D., & Bonnel, A. M. (1998). Exposure effects on music preference and recognition. *Memory and Cognition, 26*, 884–902.

Russo, R., Ward, G., Geurts, H., & Sheres, A. (1999). When unfamiliarity matters: Changing environmental context between study and test affects recognition memory for unfamiliar stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 488–499.

Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody for songs after unilateral temporal lobe lesions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 793–804.

Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition, 16*, 285–303.

Snodgrass, J., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50.

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (in press). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–403). New York: Oxford University Press.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Wagner, A. D., & Gabrieli, J. D. (1998). On the relationship between recognition familiarity and perceptual fluency: Evidence for distinct mnemonic processes. *Acta Psychologica, 98*, 211–230.

White, B. W. (1960). Recognition of distorted melodies. *American Journal of Psychology, 73*, 100–107.