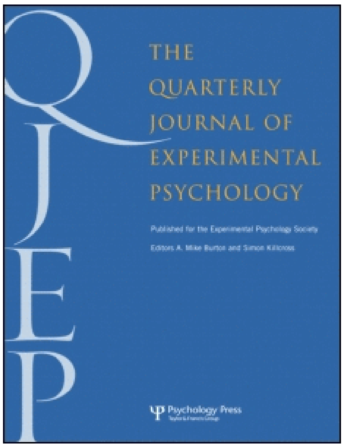

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

Effects of timbre and tempo change on memory for music

Andrea R. Halpern [a]; Daniel Müllensiefen [b]
[a] Bucknell University, Lewisburg, PA, USA [b] Department of Computing, Goldsmiths College, University of London, London, UK

First Published on: 07 October 2007

To cite this Article Halpern, Andrea R. and Müllensiefen, Daniel(2007)'Effects of timbre and tempo change on memory for music',The Quarterly Journal of Experimental Psychology,61:9,1371 — 1384

To link to this Article: DOI: 10.1080/17470210701508038

URL: http://dx.doi.org/10.1080/17470210701508038

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2008, 61 (9), 1371–1384

Psychology Press
Taylor & Francis Group

Effects of timbre and tempo change on memory for music

Andrea R. Halpern
Bucknell University, Lewisburg, PA, USA

Daniel Müllensiefen
Department of Computing, Goldsmiths College, University of London, London, UK

We investigated the effects of different encoding tasks and of manipulations of two supposedly surface parameters of music on implicit and explicit memory for tunes. In two experiments, participants were first asked to either categorize instrument or judge familiarity of 40 unfamiliar short tunes. Subsequently, participants were asked to give explicit and implicit memory ratings for a list of 80 tunes, which included 40 previously heard. Half of the 40 previously heard tunes differed in timbre (Experiment 1) or tempo (Experiment 2) in comparison with the first exposure. A third experiment compared similarity ratings of the tunes that varied in timbre or tempo. Analysis of variance (ANOVA) results suggest first that the encoding task made no difference for either memory mode. Secondly, timbre and tempo change both impaired explicit memory, whereas tempo change additionally made implicit tune recognition worse. Results are discussed in the context of implicit memory for nonsemantic materials and the possible differences in timbre and tempo in musical representations.

Memory for music, like that for other materials, can be revealed in a variety of ways. In everyday life people might *recall* music by singing or humming a tune that they recently heard, or that is familiar from long experience. In the lab, few researchers take on the formidable task of scoring sung outputs, and many participants are shy or uncertain singers, so the most common explicit memory task is *recognition*: Have you heard this tune recently? Again, this type of recognition is familiar to anyone who turns on the radio and immediately determines that a tune is familiar or not. Memory for music can also be revealed *implicitly*, or without a person consciously engaging in a memory task. Suddenly humming along to some music that a moment previously had been merely background sound is one example. We can study this kind of musical memory using some of the implicit techniques more commonly used in studies of verbal materials. Comparisons of implicit and explicit memory tests in music are worthwhile not only because both are used during everyday musical retrievals, but also because most studies of implicit and explicit memory concern verbal materials. The point of this article is to explore conditions under which implicit and explicit memory for music might be similar to or differ from memory for

Correspondence should be addressed to Andrea R. Halpern, Psychology Department, Bucknell University, Lewisburg, PA 17837, USA. E-mail: ahalpern@bucknell.edu

We thank research assistants Michelle Kanga, Brittney McGill, Chad Parker, and Ashley Wilhelm who participated in various phases of this research. Daniel Müllensiefen is funded by EPSRC Grant EP/D038855/1.

http://www.psypress.com/qjep
DOI:10.1080/17470210701508038

other kinds of materials. To what extent does musical memory adhere to findings from the verbal domain and to what extent might it differ?

An example of both similarity and difference between verbal and musical memory can be found in one direct predecessor of the current study. A commonly used implicit memory test for verbal materials is stem completion. In this task, people see a study list of words and later are presented with the opening few letters. They are asked to complete the word with "whatever comes to mind". Completion with study list words, in the absence of explicit memory, is a common paradigm for investigating preserved implicit memory (Brooks, Gibson, Friedman, & Yesavage, 1999). Word stem completion can also be presented in an auditory mode (Pilotti, Gallo, & Roediger, 2000). These tests are typically sensitive to changes in presentation format, implying reliance on a perceptual memory, but are less sensitive to changes in encoding task than are explicit memory tests (Roediger, Weldon, Stadler, & Riegler, 1992). These studies concur with many others in suggesting as a general memory principle a greater reliance in implicit tasks on perceptual memory, or memory for surface structure, and greater reliance on abstract codes in explicit tasks. This behavioural dissociation has also been found at the neural level, in studies of amnesics who retain some implicit memory abilities in the face of explicit memory deficits (e.g., Goshen-Gottstein, Moscovitch, & Melo, 2000) and the selective effects of midazolam, a benzodiazapine, on explicit versus implicit memory (Hirshman, Fisher, Henthorn, Arndt, & Passannante, 2003).

In contrast to this account, Warker and Halpern (2005) showed that musical stem completion does not respond to standard manipulations of perceptual form and encoding task in the same way as does verbal stem completion. The musical stem completion task matched the verbal counterpart as nearly as possible. The basic form of the task (Experiment 1 in Warker & Halpern, 2005) presented a set of unfamiliar melodies in a learning phase. Immediately afterwards, participants heard the beginning part of old or new tunes (the stem) and were asked, in the implicit version, to hum or sing the note that they thought *would come next musically*. The explicit version of the task followed, with a new set of tunes. The instructions were to complete the tune with the note *remembered* as having come next. Participants were screened to make sure they had scorable singing voices and had at least 2 years of instrumental music lessons to further ensure that actual singing would be adequate to the task. Both tasks showed more accurate pitch production for old than new items; performance was not correlated across the implicit and explicit tasks.

Once feasibility of the paradigm was demonstrated, Experiment 2 of Warker and Halpern (2005) added manipulations of encoding task and a perceptual change. Half the listeners were instructed to judge the regularity of the rhythm of each tune at encoding, which we deemed a shallow task, akin to vowel counting or other similar tasks used frequently in verbal levels of processing tasks (e.g., Craik & Lockhart, 1972). The other listeners rated the tunes on pleasantness, which we deemed a deep task that we likened to a semantic task, in that the complete pattern must be processed and then connected to some decision based on aesthetics (we acknowledge that "shallow" and "deep" cannot be mapped exactly from a verbal to a nonverbal domain; see General Discussion for more on this point). At test, half the old items were presented in the same timbre (implemented as the same synthesized instrument) as at learning, and half were played in a different timbre. Although memory was demonstrated in both tasks as a superiority of note production in old versus new items, neither the encoding task nor timbre change had an effect on either task. In one secondary analysis, tunes presented and tested in one of the two timbres was superior to memory for items presented in one timbre and tested in another, which was superior to items presented and tested in the other timbre. Surprisingly, this effect was seen not in the implicit, but in the explicit task. Thus we saw some evidence for a dissociation, but not in the direction predicted by the usual accounts of implicit/explicit memory differences.

This finding replicated, in certain key respects, those of an earlier study by Peretz, Gaudreau, and Bonnel (1998), using a different implicit and different explicit memory task. They chose to exploit the *mere exposure effect*, in which memory can be indexed by an increase in liking for a previously unfamiliar item after even only one exposure (Zajonc, 1968). This is a robust phenomenon seen in a wide variety of material, but is particularly suited to study memory for music, in which liking or pleasantness ratings are a natural response. Johnson, Kim, and Risse (1985) showed that even densely amnesic Korsakoff's syndrome patients preferred music in an unfamiliar style after only one previous presentation, compared to new pieces, while at the same time doing poorly in a recognition test. Halpern and O'Connor (2000) found the same dissociation in normal elderly listeners, who were at chance in recognizing just-presented melodies, but liked them better than new melodies.

Peretz et al. (1998) used both familiar and unfamiliar melodies, but results for the unfamiliar melodies are of most relevance to the current point. Different listeners were participants in the implicit (ratings of liking) and explicit (recognition memory) tasks. Encoding task was varied in one of the experiments (instrument decision task = shallow, and familiarity ratings = deep). The authors found no effect of encoding task on either implicit or explicit retrievals. Timbre change reduced memory performance, but only in the explicit task!

This result is surprising because of the assumed reliance of implicit memory retrieval on perceptual representations. Timbre is a typical example of a feature that, when changed, still keeps the identity of the melody intact. Melody is in fact generally defined in musicological terms as a sequence of pitched events in musical time without any reference to the timbral dimension (Grove Dictionary; Ringer, 2001). Busoni's (1922/1957) definition even explicitly addresses the neutral influence that instrumental timbre has on the identity of a melody when he said that the choice of instrument exercises "no change over its essence". Or expressed differently, a tune like *Happy Birthday* is the same tune in all essential respects whether sung, played on a violin, or played on a banjo.

The effect of timbre change only on the explicit task is also surprising because in the verbal analogue to this task, Schacter and Church (1992) presented auditory word lists for study and changed the voice speaking the words for some items at test. Voice change did reduce priming effects (but not cued recall performance), consistent with the proposed reliance of implicit memory measures on perceptual representations. Sheffert (1998) found that voice change effects in a perceptual identification task were most likely to be found when the words were embedded in noise, which encouraged data-driven processing. Peretz et al. (1998) presumably encouraged data-driven processing in their instrument identification encoding task, yet even in that condition, timbre change did not affect the magnitude of the mere exposure effect. Those authors speculated that timbre might function in a unique way to differentiate a given melody from all other melodies and thus increase the diagnostic value of timbre in an episodic task. This may parallel the effect that size and other "surface" attributes have on some tests of visual memory, in that size can affect explicit but not implicit memory measures in some visual object tasks (Biederman & Cooper, 1992; Jolicœur, 1987).

Because so few studies have been conducted on implicit memory for music, one overall motivation for the current study was to add to that literature. A more specific motivation was to examine possible reasons for the weak effect of timbre change in the Warker and Halpern (2005) study, when Peretz et al. (1998) found a robust effect. One possibility is that the mere exposure paradigm is more sensitive than the stem completion paradigm; to that end, we used a mere exposure paradigm here to check for replication. A second difference between the studies is that Peretz and colleagues used familiar, nameable timbres of flute and piano. Warker and Halpern deliberately used unfamiliar, hard-to-name timbres with the intent of reducing semantic coding of the timbres. It is possible that timbre change effects are more apparent with familiar timbres because the binding between the tune representation and timbre is facilitated by naming and/or familiarity. Finally, the two studies used different melodies, and perhaps item-specific effects could account for differences. Thus, in Experiment 1, we replicated the mere exposure procedure of Peretz, but using materials from Warker and Halpern (2005). We varied the familiarity of the timbres: Half the participants heard synthesized piano and organ tones, and half heard synthesized banjo and recorder tones, which fit criteria of unfamiliarity described below. In our second study, we varied tempo rather than timbre of tunes, to see whether effects of a completely different kind of "surface" change might produce effects similar to that of timbre. If so, that would question the uniqueness of timbre as a special memory attribute and instead allow us to speculate more generally about the effect of surface changes on implicit and explicit musical memory.

EXPERIMENT 1

The first experiment was conceived as a $2 \times 2 \times 2$ analysis of variance (ANOVA) design with the first factor being encoding task (shallow vs. deep, between groups) and the second factor being timbre change (same vs. different, within groups). Timbre familiarity (familiar vs. unfamiliar, within groups) served as the third factor. The difference between pleasingness ratings for old melodies and new melodies was employed as the dependent variable for measuring implicit memory performance. For measuring explicit memory performance, we used the difference between the recognition confidence ratings for old and new melodies as measured by the area under the receiver operating characteristic curve (AUC, see e.g., Swets, 1973).

We hypothesized that deep encoding should enhance explicit but not implicit memory for the melodies, if effects as predicted by the levels of processing theory (LOP, see Craik & Lockhart, 1972) act in music as they do in other materials. As a second hypothesis, different timbres for the same melody in encoding and test phase should decrease explicit memory performance, but have no effect on implicit memory, as was found by Peretz et al. (1998). Finally, timbre familiarity might show a selective effect on explicit memory with only a change in familiar timbres decreasing memory performance significantly, if this aspect were indeed responsible for the different results concerning timbre change in the Warker and Halpern (2005) and Peretz et al. (1998) studies.

Method

Participants

Participants were 63 undergraduate students at Bucknell University, between the ages of 18 and 22 years. They participated as part of a requirement for an Introductory Psychology class. The participants had a mean of 6.4 years of instrumental music lessons.

Materials

A total of 80 single-line (monophonic) melodies ranging in length from 5 to 10 s comprised the stimulus set. The melodies were not specifically quantized in time or amplitude, but allowed to vary naturally in amplitude and duration of accents. The melodies consisted of folk tunes and newly composed tunes; all were unfamiliar, as determined by ratings collected when these tunes were used in several previous experiments (an example may be seen in Warker & Halpern, 2005). We selected four timbres for the experiment, two familiar and two unfamiliar, but which were nevertheless equally pleasant. Preliminary testing suggested that the synthesized piano and organ voices on a Yamaha PSR-500 keyboard would be suitable for the familiar timbres, and the banjo and recorder voices would be suitable for the unfamiliar timbres. To confirm the timbre categories, we asked 11 Bucknell University students of varying musical backgrounds to listen to the same arpeggio played in each of the timbres, as synthesized by Cakewalk Pro Audio software, and to rate each timbre for pleasantness and familiarity on a 1 (low) to 7 (high) scale. Each timbre arpeggio was presented three times, and the scores averaged over the three listenings for each person. We combined the ratings of piano and organ, and ratings of recorder and banjo. The average familiarity rating was significantly higher for piano/organ (5.50) than for banjo/recorder (5.12); $t(21) = 1.98$; $p = .03$, one-tailed. However the average pleasantness ratings were equivalent for those pairs (4.53 vs. 4.18, *ns*).

The exposure phase consisted of 40 melodies, half of which were played in one timbre, and half of which were played in the other timbre from the same familiarity level. Those in the familiar timbre groups heard piano and organ, and those in the unfamiliar timbre groups heard recorder and banjo. The order of melodies and the assignment to a specific timbre were randomized. The melodies were played with a 5-s pause between each. The test phase consisted of all 80 melodies. A total of 40 of those melodies had not been heard in the exposure phase ("new melodies"), and 40 had been heard in the exposure phase ("old melodies"). Of the 40 old melodies, 20 were played in the other timbre from the same familiarity level as that in their original presentation. The melodies were played with an 8-s pause between each item to allow sufficient time for response. Old and new melodies were counterbalanced between the two participant groups, so that half of the melodies during the 80-melody playlist were new to half of the participants, and the other half were new to the other half of participants.

Procedure

Participants were tested in groups of one to three. Each participant was assigned to one of four groups. All groups were presented with one 40-melody playlist (exposure phase), were given a musical background questionnaire, and were then presented with the 80-melody playlist (test phase). Sessions were controlled by SuperLab software on a Macintosh G3 computer; playback was via Acoustic Research speakers at a comfortable listening level.

Two groups, the shallow task-familiar timbre (SF) and shallow task-unfamiliar timbre (SU) groups, were asked to identify the instrument with which the melody was played during the exposure phase. The SF group was composed of 15 participants; all other groups had 16. They first heard two short melodies to illustrate examples of both instruments involved (a piano and an organ, or banjo and recorder) and were asked whether they were able to identify the differences. All participants said the differences were clear. They were given an answer sheet and were asked to write the initial of the presented instrument (P or O; B or R) for each of the 40 melodies presented.

The other two groups, the deep-familiar timbre, and deep-unfamiliar timbre groups (DF and DU) were asked to rate the familiarity of the 40 melodies in the exposure phase based on a 3-point scale (unfamiliar, somewhat familiar, familiar). These groups were not given examples of the timbres, because we did not want to call attention to this characteristic.

After the exposure phase, and following the musical background questionnaire, participants were presented with all 80 melodies in the test phase of the experiment. They made two separate judgements about each melody during the 8-s pause following its presentation. They were asked to rate how well they liked the melody (which we refer to as "pleasingness") based on a 7-point scale ranging from 1 ("I do not like it") to 7 ("I like it a lot"). The participants also were asked to indicate whether they recognized the melody from the exposure phase, using a 6-point scale incorporating their recognition as well as their confidence. A score of 1 meant "I am very sure I did not hear this tune before", 2 meant "I am sure I did not hear this tune before", 3 meant "I don't think I heard this tune before", 4 meant "I think I heard this tune before", 5 meant "I am sure I heard this tune before", and 6 meant "I am very sure I heard this tune before". Half the participants in each of the four groups rated the pleasingness of the melody first, and half rated their recognition first.

Following the test phase, participants were debriefed and were given the opportunity to ask questions.

Results

Only one participant made one error in instrument classification over the entire experiment, verifying that the piano/organ and recorder/banjo sounds were easily discriminable. Familiarity ratings were scored as 0 = unfamiliar, 1 = somewhat familiar, and 2 = familiar. Average familiarity ratings ranged between 0.10 and 1.00 with a mean of 0.53 on the 3-point scale. Thus, as expected, the tunes used for testing were quite unfamiliar for the large majority of the participants.

Explicit scores were computed by converting the confidence ratings to AUC scores for each participant. This measure ranges from .50 (chance) to 1.0 (perfect discrimination between old and new items). Implicit scores were simply the difference in pleasingness ratings between old and new items. These scores were analysed in separate ANOVAs for the implicit and explicit tests, with encoding task and timbre familiarity as between-subjects factors. Same versus different timbre is only defined on the old items; for the purposes of calculating old–new discrimination or change in pleasingness, performance on same timbre and different timbre were each compared to scores on the complete set of new items.

The resulting explicit and implicit memory scores in all factor combinations are displayed in Table 1. Old–new implicit scores ranged from .20 to .25 in three of the four groups. These scores are in the same range as those found by Warker and Halpern (2005) using the stem completion paradigm. Only in the shallow encoding of familiar timbres were old–new scores lower, averaging .15 for same timbre and only .02 for different timbre (see Table 1). This suggests weaker implicit memory performance for familiar timbres. However, this trend was not confirmed by the ANOVA, as neither encoding task nor timbre familiarity or timbre change were shown to be significant effects in the ANOVA (mean for same timbre = .21 and for different timbre = .17), nor were any interactions (all Fs = 1.6 or less). For explicit memory, mean AUC scores ranged from .60 to .69, suggesting that the task was fairly difficult for participants. Nevertheless, the lower bound of the 95% confidence interval was above .50 for all groups and conditions. There was a strong main effect of timbre change: Same timbre items averaged an AUC of .67, and different timbre items earned .61, $F(1, 58) = 32.11$, $p < .001$, partial eta^2 = .365. No other main effects or interactions were significant.

We correlated implicit and explicit scores within each group, separately for same-timbre and different-timbre trials. Although all eight correlations were positive, only one was significant: in the deep encoding of familiar timbres, for same-timbre items, $r(14) = .53$, $r_{crit} = .50$. All other correlations were .42 and lower.

Table 1. *Mean explicit (AUC) and implicit (old–new pleasingness) memory scores by encoding task, difference in timbre, and timbre familiarity in Experiment 1*

	Same timbre				*Different timbre*			
	Familiar timbre		*Unfamiliar timbre*		*Familiar timbre*		*Unfamiliar timbre*	
Encoding task	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Explicit test								
Instrument	.67	.09	.69	.10	.60	.08	.63	.10
Familiarity	.68	.10	.66	.08	.61	.08	.60	.08
Implicit test								
Instrument	.15	.31	.24	.24	.02	.30	.24	.28
Familiarity	.24	.48	.20	.34	.17	.37	.25	.32

Note: AUC = area under the receiver operating characteristic curve.

Discussion

Experiment 1 replicated the mere exposure effect with unfamiliar music, as has been seen in several previous studies (Halpern & O'Connor, 2000; Johnson et al., 1985; Peretz et al., 1998; Warker & Halpern, 2005). Although the data hinted that this effect may depend on familiarity of timbre, this turned out not to be reliable, at least for the familiar versus unfamiliar timbres used in this experiment. Perhaps larger differences would accrue if the timbres differed more in familiarity than did our choices, although pilot testing showed that timbres rated as very unfamiliar from our particular keyboard also tended to be hard to identify, irritating, or both. Also similar to previous studies using unfamiliar melodies, we found no effect of encoding task in either type of memory.

Experiment 1 further replicated the finding of Peretz et al. (1998) that timbre change produces a reliable decrement in memory for unfamiliar items, but only in the explicit task. As this effect did not interact with timbre familiarity, this has now been shown with three sets of timbre pairs: piano/flute in the Peretz study, and piano/organ and banjo/recorder in the current study. In contrast to studies with verbal materials, the change of an apparently superficial aspect of timbre appears to affect explicit memory measures and not implicit measures. This seems to show an effect of encoding specificity, similar to the enhancement found for picture recognition when colour or monochrome is preserved between exposure and test (Spence, Wong, Rusan, & Rastegar, 2006). However, the absence of a superiority for perceptually matched stimuli at encoding and test during implicit tests contrasts with research cited above showing such enhancement for verbal materials.

We earlier noted that the discrepancy between the results for timbre change in the studies of Warker and Halpern (2005) and Peretz et al. (1998) might have been due either to timbre familiarity or to difference in memory paradigms. Because the same melody set was used here as in our previous study, and because timbre familiarity made no difference in the current results, we conclude that the discrepancy arises due to the difference in stem completion versus recognition paradigms. It may be the case, for instance, that a production task removes attention from timbre processing because the participant can only hum or sing in his or her own voice, which is far different from any timbres presented in the study. This might explain why Schacter and Church (1992) did find effects of voice change (in an implicit task): We all do sometimes imitate other voices and can at least imagine ourselves doing so. Imitating banjos and organs is beyond the experience or capability of all but the most adept jazz singers or sound effects experts.

The results with timbre change raise a question about the effects of changing other musical aspects on memory performance for melodies. The next experiment considers tempo. Like timbre, the tempo of a piece can be changed (within limits) and still maintain the piece's identity. This is certainly true from a musicological perspective and is reflected by the fact that in Western art music, often only relative indications of tempo ranges (e.g., andante, moderato, presto) for the performance of a piece are provided by the composer. Performers use tempo as a performance parameter that maintains the identity of a work while deliberately altering a performance feature that distinguishes different renditions of the same piece. On the perceptual side, tempo is not stringently associated with particular songs, across people. Halpern (1988) asked listeners to set the tempo of familiar children's, patriotic, and folk songs. Individuals varied in the tempos that they considered appropriate for the same tunes. For instance, among the 20 participants in Experiment 1 in that series, preferred tempo for *Happy Birthday* varied from 96 beats per minute (bpm) to 140 bpm, and in a slower song, *Hark the Herald Angels Sing*, the range was 48 to 92 bpm (note, these data were not published in that report). Thus, we were interested in whether tempo would also be bound with a tune's identity, as shown by worse memory performance when the tempo changed from study to test.

EXPERIMENT 2

The purpose of Experiment 2 was to evaluate the influence of encoding task and tempo change on implicit and explicit memory for short melodies. We also wanted to generalize the procedure by changing some of the timbres that we used and by slightly changing the implicit task from ratings of "pleasingness" (dislike to like) to ratings of "pleasantness" (very unpleasant to very pleasant). These differ in the emphasis on personal reaction versus aesthetic evaluation (one may, in theory, dislike a pleasant piece) and allowed us to expand the range of rating tasks that may be suitable to elicit evidence of implicit memory. We predicted that encoding task would continue to have null effects on memory of both types, that tempo change would reduce explicit memory, and that tempo change would not reduce implicit memory performance.

Method

Participants

A total of 32 undergraduates between the ages of 18 and 25 years from the University of Texas at Dallas were recruited as participants. They had received 4.4 years of instrumental music lessons on average and were paid or received course credit for their services.

Materials

As stimulus materials we used the same 80 short tunes, represented as single-line melodies as in Experiment 1.

For the tempo condition, 40 of these 80 tunes were altered in tempo. To minimize the influence of individual aesthetic judgements and, thus, to minimize an experimenter effect in the preparation of the materials the following rules were chosen to alter the tempo of the tunes.

The general aim was to make the tempo changes in altered tunes large enough to be perceptible, on the one hand, but not so large that recognition judgements could be based on the identification of extreme tempos per se. Therefore, the basic rule was to move the tempo of the altered tunes towards 120 bpm at the beat or crotchet level. The rate of 120 bpm is sometimes referred to as the spontaneous tempo, because people tend to tap around this rate, when not given any specific instruction about tempo in which to tap a regular series of beats (Fraisse, 1982). Therefore, every tune with an original tempo differing more than 15% from the target tempo (<102 bpm or >138 bpm) was altered in tempo by 15–20% in the direction of the target tempo. Pieces where the beat could be considered at the level of quavers or minims were altered accordingly. A musicologist from the University of Hamburg who was naïve to the aim of the experiment listened to each altered tune and confirmed that the chosen level of tempo alteration (15–20% around 120 bpm) generally preserved the character of the tunes. A total of 20 tunes were sped up, and 20 tunes were slowed down. This general rule was violated for a few tunes where the parameters could not be adhered to—for example, because of an unusually large range of durational values.

As our encoding task involved instrument classification, we chose three clearly different timbres that we assumed to have about the same degree of familiarity to our participants: cello, marimba, and clarinet. All three instruments are not among the most commonly studied and at the same time they are frequent instruments in Western concert orchestras. They belong to three different instrument families (idiophones, cordophones, aerophones), and they did not appear to differ too much in pleasantness when synthesized by our MIDI device. We rendered all MIDI files to audio (.wav files) using the FINALE score editor, which uses the Roland Virtual Sound Canvas that is well suited for reproducing naturalistic instrumental timbres. The tunes were randomly assigned to the timbre (cello, marimba, clarinet) and tempo (same, different) conditions, and the order of the tunes was eventually randomized. All audio versions of the tunes were burned to audio CDs inserting 4 s of silence between tunes. The tunes were played to the participants from the CDs from a Dell Precision M70 laptop computer over Harman-Kardon Multimedia speakers during the experimental session.

Procedure

The experiment used a 2 × 2 factorial design. The within-subject factor was a same versus different tempo, whereas the between-subject factor was the encoding task. Participants were assigned randomly to the two encoding task groups and were tested in groups of one to three. Half of the listeners were instructed to listen to a series of 40 short tunes and to mark the instrument in which they were played on an accompanying sheet, serving as the shallow task. One example for each of the three different instruments was played before the actual task. The remaining participants were instructed to rate their familiarity with the tunes on a 3-point scale where the tune could be indicated as being "unfamiliar", "somewhat familiar", or "familiar". This we considered to be a deep encoding condition, as in Experiment 1.

In the test phase of the experiment participants were asked to listen to a series of 80 tunes and had to indicate for each tune how pleasant it appeared to them (7-point scale, from "very unpleasant" to "very pleasant"). Of the 80 tunes in the test phase, 40 were new to the listeners. Of the 40 old tunes, 20 were altered in tempo, and 20 preserved the original tempo from the first phase. The pleasantness scale served as an indicator for implicit memory, with higher pleasantness ratings for old versus new items indicating (implicit) recognition of an item. Additionally, the participants had to rate for each of the 80 tunes whether they had heard the tune in the encoding session (6-point scale, from "very sure not heard it before" to "very sure heard it before"). This served as explicit memory ratings. The order of the two ratings scales on the sheet was reversed for half of the participants. After the test session, participants filled out a short questionnaire about their musical background.

Results

As in Experiment 1 a dependent measure for the explicit memory performance, the AUC, was computed for each participant and for same and different tempo items separately. For the implicit (pleasantness) ratings simply the mean difference between old and new items was calculated for each participant, separately for the same tempo and different conditions among the old items.

Instrument recognition, familiarity, and explicit memory performance

Recognizing the instruments was, once again, an easy task. The mean of the participants' correct answers was .99. Average familiarity ratings for the participants ranged between 0.00 and 1.23 with a mean of 0.63 on the 3-point scale (where 2 meant the tune was familiar).

The overall performance in the explicit memory task as measured by the AUC scores ranged from .51 to .83, with a mean of .68. The lower bound of the 95% confidence interval was above .50 for the mean performance in all four Groups × Conditions.

ANOVA results

We computed two different ANOVAs: one with the AUC for each participant as the dependent variable for the explicit task and another on the difference in pleasantness ratings as the dependent variables for the implicit task. The within-subject factor was same versus different tempo, which, of course, could only be calculated on the basis of the 40 old items present in the test phase of the experiment. The between-subject factor was the encoding task. The resulting means for all conditions are displayed in Table 2.

For the explicit task the within-subjects factor change in tempo proved to be significant, $F(1, 30) = 48.7$, $p < .001$, partial $\text{eta}^2 = .619$. Rating tunes with the same tempo as that during the encoding session yielded a mean AUC value of .72 over all participants, whereas tunes with different tempo were only recognized with an average performance level of .64. In contrast, encoding task as a between-subjects factor did not reach the usual significance level, $F(1, 30) = 1.4$, $p = .24$. The overall mean was .67 for the shallow task and .70 for the deep encoding task.

Table 2. *Mean explicit (AUC) and implicit (old–new pleasantness) memory scores by encoding task and difference in tempo in Experiment 2*

	Same tempo		*Different tempo*	
Encoding task	*Mean*	*SD*	*Mean*	*SD*
Explicit test				
Instrument	.71	.05	.62	.08
Familiarity	.73	.10	.66	.09
Implicit test				
Instrument	.30	.40	.04	.38
Familiarity	.23	.41	.06	.38

Note: AUC = area under the receiver operating characteristic curve.

We also reanalysed the data using years of musical training as a covariate. The general result did not change: Tempo did influence the participants' explicit memory ratings, but encoding task and years of lessons were far from reaching significance.

For the implicit measure of memory performance (difference in pleasantness ratings) we obtained very similar results. Tempo change was again significant in its influence on participants' ratings, $F(1, 30) = 8.7$, $p = .006$, partial $eta^2 = .225$. Tunes that were presented in the same tempo during encoding and test phase had a mean change in pleasantness ratings of 0.26 (on the 7-point scale), while the change in pleasantness ratings was only 0.05 for tunes that were presented in a different tempo. Encoding task, again, made no difference to the memory performance, $F(1, 30) = 0.5$, $p = .822$; the means for the two groups were .17 for the shallow and .14 for the deep encoding. Like with the explicit memory measure, this general result did not drastically change when years of music lessons was added as a covariate. Encoding task and years of training had no significant effect, whereas tempo change came very close to the usual .05 significance level, $F(1, 29) = 4.0$, $p = .054$.

Correlating implicit and explicit measures of memory performance

Although the analysis of implicit and explicit memory measures yielded essentially the same results when considering the influence of tempo change and encoding task, implicit and explicit memory were uncorrelated in the individual participants. The correlation of AUC scores and old–new difference in pleasantness scores of the 32 participants aggregated over all experimental conditions yielded, $r = .11$, which did not reach significance, $p = .290$. We computed correlations between implicit and explicit memory scores for each of the four groups from the 2 × 2 experimental design and found values for r ranging from −.24 to .34, with none of these correlations getting even close to the usual significance level of .05.

We also correlated mean AUC scores and pleasantness difference scores over items, to see whether the same melodies were eliciting high or low performance in the explicit and implicit tasks. Medians of explicit memory ratings for the individual tunes varied from 2.0 to 5.5 even within the same experimental condition ("old" tunes with same tempo). A Kruskal–Wallis test confirmed this difference in memorability between tunes in the same experimental condition, $\chi^2(32, N = 640) = 105.4$, $p < .001$, and a similar significant difference was found between pleasantness ratings for individual tunes. But the correlation between implicit and explicit ratings over old tunes with the same tempo was very low, $r(32) = .08$. Thus the two tasks, even though they were both sensitive to tempo change, were dissociated both on the participant and the item level.

EXPERIMENT 3

Experiments 1 and 2 results differed in one major respect: Timbre change only affected memory in the explicit task in Experiment 1, but tempo change affected memory tested in both ways in Experiment 2. Tempo and timbre were selected to be similar musical dimensions in the sense that both of them can be considered surface features of melodies as opposed to pitch interval, rhythm, or melodic contour. While even highly noticeable changes in tempo and timbre (within limits) do not affect a tune identity (see e.g., Radvansky, Fleming, & Simmons, 1995; Tunks, Bowers, & Eagle, 1993), comparatively few changes in pitch interval and contour can obfuscate a melody's identity greatly as has been shown widely in the past (e.g., Deutsch, 1972; White, 1960). We therefore wondered whether the more extensive effects in Experiment 2 might be due to tune identity being more affected by our particular tempo change than our timbre changes. Of course, we could not put these on the same metric, so we asked a new set of participants to rate the similarity of pairs of tunes that differed in timbre only, tempo only, both timbre and tempo, or neither.

Method

Participants

A total of 16 students at the University of Texas at Dallas participated from the same pool as that in Experiment 2. They were required to have had a minimum of 2 years of private instrument lessons; mean years of lessons = 7.38.

Materials

We used 48 tunes from Experiment 2. On each trial, participants heard two repetitions of the same tune. Each tune was randomly assigned to each of the four conditions: same timbre and tempo, same timbre and different tempo, different timbre and same tempo, and different timbre and different tempo. The two timbres were piano and organ (given that timbre did not affect results in Experiment 1, we only tested one pair). These two timbres were clearly distinguishable, as instrument classification performance in both previous experiments was nearly perfect for piano versus organ. Each timbre appeared equally often as the first and second timbres over trials. The two tempos for a given tune were those used in Experiment 2. In different trials, the faster and slower tempo appeared equally often in the first and second positions. Because we were mostly interested in the conditions where only timbre or only tempo varied, we had 16 trials of each of these, and 8 trials of the other two types.

Procedure

Participants were asked to rate each pair for similarity, using a 1 (least similar) to 7 (most similar) scale. They were encouraged to use the whole range of the scale and were told that tunes might vary in their instrument or speed. Playback used the same equipment as that in Experiment 2.

Results

The average ratings for the four groups of tunes conformed to our expectations on a rank level: Identical tunes of course garnered the highest average rating (6.39, $SD = 0.58$) and tunes varying on both dimensions elicited the lowest average rating (3.54, $SD = 1.02$). Of more interest were the other two conditions that received an intermediate average similarity rating. Tunes with the same timbre and different tempo yielded an average rating of 5.28 ($SD = 0.59$), and tunes with different timbre but same tempo were considered less similar, mean = 4.29 ($SD = 0.82$). This difference was significant, two-tailed test, $t(15) = 3.44$; $p = .004$. Thus we concluded that for the timbres and tempos used here, having the same timbre made two tunes sound more similar than did having the same tempo.

GENERAL DISCUSSION

The two major experiments together reinforced previous findings that memory for music can be measured implicitly using both a production task (Warker & Halpern, 2005) and an affective rating task. The implicit memory results generalized in this set of studies over two variants of the rating task (pleasingness and pleasantness) and with several kinds of timbre. The implicit memory success was also independent of encoding conditions, which is typical of other experiments using nonmusical materials. We also showed that memory performance on explicit and implicit tests were uncorrelated whether measured with participants as the unit of analysis or with items (Experiment 2). Thus we can conclude that at one broad level, these techniques can capture the everyday experience of processing music more fluently after only one previous exposure, even without definite recollection of the encoding experience.

We also replicated the somewhat paradoxical effect that we found earlier (as did Peretz et al., 1998): that a change in an apparent surface structure affects an explicit task of recognition memory. Particularly in the case of unfamiliar music, as that used here, "surface" qualities may be more bound with item identity than is the case with verbal materials, because there are no semantic anchors to allow an elaborated code, and thus memory must be encoded in literals. Peretz et al. did find an effect of timbre change on explicit memory even for familiar music, but those melodies were originally songs, and their presentation without words in flute or piano timbres might have encouraged a less abstract and more literal memory representation than is true for verbal materials that have a true semantic code.

Peretz et al. (1998) had speculated that timbre might enjoy a "special status" as a surface aspect that is nevertheless used for recollection (p. 899). We showed that timbre may have a special, but not unique, status. Although both timbre and tempo refer to a dynamic aspect of the musical information, they certainly differ in their realization. Timbre includes spectral cues, is categorical, and can be abstracted with minimal information, whereas tempo is only durational, is a continuous dimension, and at least a few events must occur before a tempo can be perceived. However, both appear to play a part in forming an episodic memory trace of at least unfamiliar music (we did not test familiar music), and thus we have even stronger evidence that memory for music, like memory for some visual representations, differs from verbal memory.

Also consistent with our previous study (Warker & Halpern, 2005) and that of Peretz et al. (1998) is the failure to find levels of processing effects with unfamiliar music. Studies with familiar music may reduce to verbal tasks, if in fact participants code the titles of the tunes (familiar music is usually highly nameable as well). Thus using unfamiliar music is a stronger test of the generality of LOP effects. Very few instances of musical LOP effects occur in the literature; a recent search of *PsychInfo* using search terms "levels of processing" or "depth of processing" and "music" revealed only one traditional LOP study (Segalowitz, Cohen, Chan, & Prieru, 2001). Although interpretations of absence can be tricky, we think it safe to assume that many people have tried this sort of study. Several explanations occur to us. Specific to our studies, perhaps timbre recognition and ratings of familiarity do not differ in the depth of processing that they require, nor did ratings of rhythmic regularity and pleasantness in our previous study (Warker & Halpern, 2005), and consequently the memory traces for the items processed by the two procedures do not differ. These sets of tasks do appear to be analogous to standard LOP manipulations with verbal materials, but perhaps simply exploring more tasks is necessary before concluding that music is not amenable to these encoding manipulations.

The second interpretation for the missing LOP effects rejects the notion of a hierarchical system or a possible one-dimensional ordering of encoding tasks for music objects. As different aspects of music might have high or low importance depending on the context and the goal of music processing in a particular setting, different encoding tasks could direct participants' attention to different aspects of the musical stimuli, and thus the encoding of the items might not differ in depth nor the memory traces differ in strength. Given the example of the present study, instrument recognition and familiarity ratings might result in abstracting different cues from the presented stimuli, but this might not make any difference in terms of the retrieval processes required by the implicit and explicit memory procedure in the test phase. Encoding specificity might be a more important principle for remembering music than it is for more semantically codable information, overriding any LOP effects.

One puzzling aspect of our results was that timbre change did not affect implicit memory in Experiment 1, but tempo change did in Experiment 2. In the studies we have been reviewing, that result is an outlier. We tested one explanation in Experiment 3, on the thought that tunes that differed in the tempos we used here might be considered as more dissimilar than tunes that differed in the timbres we selected. However, we found the opposite result. Our tunes seemed to be categorized more by timbre than by tempo, in that a change of timbre was perceptually more salient than a change in tempo.

Another explanation for the difference in the effect of timbre and tempo might stem from the way the participants relate to the musical items. Although this study did not require any musical production, in the real world implicit memory is often instantiated by singing along with a tune, and this experience may have guided the participants in our task. Among the most important cues that need to be abstracted from a music object in order to reproduce it is tempo, because tempo is crucial for synchronizing musical objects (e.g., voices or parts) that are produced by different people. In contrast, timbre is largely irrelevant if the task consists in reproducing a melody from aural memory, because timbre may very well be determined by the instrument available: usually one's voice but sometimes by an instrument. So in terms of a possible reproduction during implicit retrieval, tempo might be a more integral and eventually important feature than timbre.

But if tempo was indeed a more integral feature of melodies, why did the participants in Experiment 3 rate melodies that differed in tempo more similar to each other than those differing in timbre? First, the task of deciding on the identity of instrumental timbres is obviously easier than a judgement about tempos, because first the timbre decision can be made after hearing the first note, while a different tempo can only be detected when time evolves. Secondly, differences in tempo would be conceptualized on a quasi-continuous scale having a larger number of possible realizations. In contrast, there were only two timbre categories between which to decide. So, a distinction that can be made more quickly and that requires only two categories to choose from is obviously easier and might in turn lead the participants to reason that it is the greater difference between the objects that makes the task easier.

CONCLUSION

To summarize, the results of this study show that implicit memory for melodies can be reliably demonstrated using the mere exposure paradigm with only a single prior exposure, and that implicit memory seems less susceptible to changes in at least certain surface aspects (i.e., timbre) than is explicit memory. However, neither experiment showed an effect of encoding task, contrary to the levels of processing-theory framework. Taken together with the scant evidence of LOP effects with music in the literature, we conjecture that memory for music—as a domain that mostly lacks a clear semantic level—is more easily explained using other principles of encoding and retrieval. The types of encoding that facilitate or inhibit tune memory are an area for interesting future studies.

Original manuscript received 5 October 2006
Accepted revision received 6 June 2007
First published online 7 October 2007

REFERENCES

Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 121–133.

Brooks, J., Gibson, J., Friedman, L., & Yesavage, J. (1999). How to drastically reduce priming in word stem completion—and still present the words. *Memory & Cognition*, *27*, 208–219.

Busoni, F. (1957). *Von der Einheit der Musik* [The essence of music]. London: Rockliff (Original work published 1922).

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics*, *11*, 411–412.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *Psychology of music* (pp. 149–180). New York: Academic Press.

Goshen-Gottstein, Y., Moscovitch, M., & Melo, B. (2000). Intact implicit memory for newly formed verbal associations in amnesic patients following single study trials. *Neuropsychology*, *14*, 570–578.

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

Halpern, A. R., & O'Connor, M. G. (2000). Implicit memory for music in Alzheimer's disease. *Neuropsychology*, *14*, 391–397.

Hirshman, E., Fisher, J., Henthorn, T., Arndt, J., & Passannante, A. (2003). The effect of midazolam on conscious, controlled processing: Evidence from the process-dissociation procedure. *Memory & Cognition*, *31*, 1181–1187.

Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 22–36.

Jolicœur, P. (1987). A size congruency effect in memory for visual shape. *Memory & Cognition*, *15*, 531–543.

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

Pilotti, M., Gallo, D., & Roediger, H. L., III (2000). Effects of hearing words, imaging words, and reading on auditory implicit and explicit memory tests. *Memory & Cognition*, *28*, 1406–1418.

Radvansky, G. A., Fleming, K. J., & Simmons, J. A. (1995). Timbre reliance in nonmusicians' and musicians' memory for melodies. *Music Perception*, *13*, 127–140.

Ringer, A. L. (2001). Melody. In L. Macy (Ed.), *Grove music online*. Retrieved March 1, 2007, from http://www.grovemusic.com

Roediger, H. L., III, Weldon, M. S., Stadler, M. L., & Riegler, G. L. (1992). Direct comparison of two implicit memory tests: Word fragments and word completions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1251–1269.

Schacter, D. L., & Church, B. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 915–930.

Segalowitz, N., Cohen, P., Chan, A., & Prieru, T. (2001). Musical recall memory: Contributions of elaboration and depth of processing. *Psychology of Music*, *29*, 139–148.

Sheffert, S. M. (1998). Voice-specificity effects on auditory word priming. *Memory & Cognition*, *26*, 591–598.

Spence, I., Wong, P., Rusan, M., & Rastegar, N. (2006). How color enhances visual memory for natural scenes. *Psychological Science*, *17*, 1–6.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science*, *182*, 990–1000.

Tunks, T. W., Bowers, D. R., & Eagle, C. T. (1993). The effect of stimulus tempo on melodic error detection. *Psychomusicology*, *12*, 41–51.

Warker, J. A., & Halpern, A. R. (2005). Musical stem completion: Humming that note. *American Journal of Psychology*, *118*, 567–585.

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

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, Monograph Supplement*, *9*, 1–27.