Journal of Experimental Psychology:
Human Perception and Performance
2011, Vol. 37, No. 5, 1512–1526

0096-1523/11/$12.00 DOI: 10.1037/a0023507

Specific Previous Experience Affects Perception of Harmony and Meter

Sarah C. Creel
University of California, San Diego

Prior knowledge shapes our experiences, but which prior knowledge shapes which experiences? This question is addressed in the domain of music perception. Three experiments were used to determine whether listeners activate specific musical memories during music listening. Each experiment provided listeners with one of two musical contexts that was presented simultaneously with a melody. After a listener was familiarized with melodies embedded in contexts, the listener heard melodies in isolation and judged the fit of a final harmonic or metrical probe event. The probe event matched either the familiar (but absent) context or an unfamiliar context. For both harmonic (Experiments 1 and 3) and metrical (Experiment 2) information, exposure to context shifted listeners' preferences toward a probe matching the context that they had been familiarized with. This suggests that listeners rapidly form specific musical memories without explicit instruction, which are then activated during music listening. These data pose an interesting challenge for models of music perception which implicitly assume that the listener's knowledge base is predominantly schematic or abstract.

Keywords: encoding specificity, auditory representation, music perception, context recall

Does music listening activate generic knowledge, or do listeners preferentially activate very context-specific representations? Domains of inquiry as diverse as language processing (e.g. Goldinger, 1998; Hare, McRae, & Elman, 2003; Trueswell, Tanenhaus, & Kello, 1993), social stereotyping (e.g. Gill, 2004; Locksley, Borgida, Brekke, & Hepburn, 1980), and music perception (Justus & Bharucha, 2001) all seek to understand the roles of generic and highly specific knowledge in informing the perceptions of present events. In regards to music perception, we know that listeners activate musical knowledge gleaned from previous experiences when hearing music (e.g., Hannon & Trehub, 2005a, 2005b). What is not clear is what set or subset of musical knowledge is invoked by the listener.

Music perception is thought to involve the implicit formation of expectations for pitch patterns and metrical patterns (Huron, 2006; Meyer, 1967). Expectations are estimates of the probability that particular upcoming events will occur based on activated memory information (e.g., DeLong, Urbach, & Kutas, 2005). Listeners can be "surprised" by unexpected musical events, such as a deceptive cadence (where a highly schematically expected tonic chord [C in C major] is replaced with a different chord) or an "oddball" event (as in Franz Joseph Haydn's *Surprise* Symphony, where a quiet melody is interrupted by a loud chord). Listeners' expectations are exploited by composers and performers, and it is theorized that they contribute to musical enjoyment; that the right amount of fulfilled and foiled expectations are maximally enjoyable (Meyer, 1967). One puzzle is exactly what set of musical knowledge generates these implicit expectations under what circumstances: A surprising turn of phrase for Mozart might well sound trite coming from The Beatles. In order to have different expectations in different situations, listeners would need to have distinct memory representations contributing to expectations in each situation. That is, they would need context-specific memory.

Classic memory research suggests that context influences memory processes (e.g., Godden & Baddeley, 1975), and this is true of musical memory. Dowling (1986) played a melody with one set of harmonies for listeners and then asked listeners to detect a change in the melody when the underlying harmonies were changed or unchanged. Listeners with a moderate level of music experience had difficulty detecting a melodic change when the harmonic contexts changed. Povel and Essens (1985) similarly found that listeners had difficulty identifying familiar rhythms when the metrical context—the underlying beat—was changed (see Essens & Povel, 1985, for similar results in rhythm encoding). This suggests that musical stimuli were better recognized in the contexts in which they were learned.

The question addressed here is whether specific musical memories are reactivated during later listening. It is instructive to consider an analogous question that has been posed in sentence processing: Do comprehenders deploy specific information about how individual words function, or do they deploy information about how words of a class (e.g., verbs) function generally? For instance, transitive verbs generally only have a subject and a direct object. However, the transitive verb *put* is unusual in that it requires a subject, object, and location. You cannot say, *She put the car*, you must say, *She put the car in the garage*. Do comprehenders activate just a subject and object for this idiosyncratic verb, or do they also activate the location, even though other verbs do not require it? Psycholinguistic research suggests that yes, rather than using a general category of "transitive verb," comprehenders activate fine-grained information about the particular use of that verb (Hare et al., 2003; Trueswell et al., 1993). More specifically, comprehenders expect just an object for most verbs

This article was published Online First May 9, 2011.

Correspondence concerning this article should be addressed to Sarah C. Creel, Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515. E-mail: creel@cogsci.ucsd.edu

(e.g., *eat*), but they expect an object and a location when the verb is *put*. That is, the most accurate account of human sentence processing contains specific information about particular verbs, even when other verbs share some structural overlap.

A similar case is made here about musical representations. The hypothesis is that musical memory contains detailed contextual information, which is activated during listening to the extent that it is similar to the current event. Such graded activation would allow listeners to keep track of the multiple musical styles that they are often exposed to (e.g., Rentfrow & Gosling, 2003; Wong, Roy, & Margulis, 2009). Such memory might consist of individual exemplars (e.g., Carpenter & Grossberg, 2003; Goldinger, 1998; Hintzman, 1986) or numerous style-specific schemas. Either alternative would allow more similar memories to more strongly affect processing.

Levels of Specificity: Specific and Generic Information in Music

Existing research paints a mixed picture of musical knowledge. On the one hand, the knowledge that constrains music processing often appears fairly generic or stylized (e.g., Justus & Bharucha, 2001; Krumhansl, 1990). On the other hand, listeners demonstrably possess specific musical memories (Palmer, Jungers, & Jusczyk, 2001; Schellenberg, Iverson, & McKinnon, 1999; Schellenberg & Trehub, 2003).

Justus and Bharucha (2001) distinguish generalized knowledge of common musical patterns, or *schematic* knowledge, from context-specific, or *veridical*, knowledge. In the current paper, *schematic* information is used synonymously with *generic*—musical knowledge that glosses over *veridical* (here, synonymous with *specific*) knowledge. Justus and Bharucha presented listeners with pairs of chords (PRIME . . . TARGET), and asked them to make speeded in tune and out of tune judgments of each target. Target chords were either closely harmonically related to the prime chord (and thus schematically expected) or distantly harmonically related to the prime chord (not schematically expected). The authors boosted within-experiment (veridical) expectations of the schematically-unexpected chord either by giving a preview of it on each trial (PRIME TARGET PRIME TARGET) or by presenting it more frequently than the schematically-expected chord (75% vs. 25%). Listeners' reaction times were still faster to schematically-expected in-tune chords than unexpected ones. The authors suggested that, even when recent experience suggested a specific chord, generalized musical knowledge was activated. Justus and Bharucha's veridical and schematic distinction is taken to represent a distinction of detail; that is, a continuum from more-detailed (fully veridical, analog) to less-detailed (schematic, abstract) representations, though the reader should not take this as an assertion that listeners can access single musical episodes in memory.

Schematic Knowledge in Processing

Abundant evidence suggests that listeners activate schema-like knowledge in processing the two major dimensions of music, pitch and time (e.g. Hannon & Trehub, 2005a, 2005b; Krumhansl, 1990). For instance, listeners know what tones are likely to occur together, which Krumhansl (1990) calls the "tonal hierarchy." This knowledge seems to be gleaned from one's native musical system between infancy and adulthood (Lynch, Eilers, Oller, & Urbano, 1990). A listener hearing tones drawn from the C major scale (C, D, E, F, G, A, B) would gauge the note E to fit well (E is contained in the C major scale) but not the note D-sharp (D#, not contained in the scale; Krumhansl & Shepard, 1979; see Krumhansl, 1990 for an overview). Western listeners also know implicitly what chords are likely to follow other chords (Justus & Bharucha, 2001; Patel, Gibson, Ratner, Besson, & Holcomb, 1998): If listeners hear a G dominant chord, they strongly expect a C tonic triad to follow. This dominant-to-tonic chord sequence is highly schematized as it is the formulaic chord sequence that ends most Western musical pieces (Kostka & Payne, 1995).

Recognition of meter is also affected by knowledge that could be described as schematic: Notes are expected to occur at equal time intervals (e.g., Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999). No single piece of music conforms completely to a single repeating time interval (unless one is listening to a ticking watch, metronome, or perhaps a fugue by Johann Sebastian Bach), but across a listener's experience, musical events tend to fall at evenly-spaced time points (Palmer & Krumhansl, 1990). This accrued knowledge may underlie listeners' metrical expectations. Research by Hannon and Trehub (2005a, 2005b) is consistent with the idea that particular meters are encoded over extensive experience. These authors presented Western listeners and Eastern European listeners with "uneven" meters (meters with unevenly spaced beats). Western adults had difficulty recognizing these meters, which they had almost no experience with, even after two weeks of training. Eastern European adults, who had vast cultural exposure to uneven meters, recognized them easily. One can interpret this as evidence that listeners abstract metrical schemas from lengthy experience (i.e., over developmental time) but not from substantial specific experience over a span of two weeks. If this account is correct, then metrical schemas may be relatively resistant to new experience.

Numerous models of tonality perception and beat detection implicitly assume schematic knowledge. For instance, the Krumhansl-Schmuckler key-finding algorithm (described in Krumhansl, 1990; see also Temperley, 1999) guesses the key of a musical work by comparing a vector of frequencies-of-occurrence of each pitch class (C, C#, D, D#, etc.) in the piece of music to vectors of ratings of tone importance of each of the 24 possible keys (C major, C minor, C# major, etc.; from Krumhansl, 1990). The key with the highest correlation is the algorithm's guess as to the key of the piece. Similarly, models of beat detection such as Large and Jones (1999) posit that listeners determine meter via internal oscillators that become entrained to periodicity (temporal regularity) in the signal. These key vectors and oscillators are essentially schematized knowledge representations. Though these researchers may not assume that the listener's knowledge is wholly schematic, the models constitute assumptions that schematic knowledge is a good approximation of listeners' knowledge.

Knowledge of Specific Musical Contexts

In addition to evidence for schematic knowledge, other data suggest that listeners possess detailed musical knowledge. For instance, listeners are above chance at identifying the correct absolute pitch of familiar songs (Schellenberg & Trehub, 2003). Listeners can identify familiar songs (Schellenberg, Iverson, &

McKinnon, 1999) and musical genres (Gjerdingen & Perrott, 2008) from 200–250 ms of exposure; both sets of authors suggest that this is based on spectral (timbral) information. Finally, listeners can recognize familiar melodic fragments based on subtle details of musical expression (amplitude and timing variation; Palmer et al., 2001).

These studies suggest that listeners are capable of encoding musical information in great detail. However, we do not know whether listeners routinely invoke such specific memory representations when processing music in everyday listening contexts. Recent work on music learning suggests that listeners do sometimes invoke representations derived from recent experience. For instance, listeners can gauge the goodness-of-fit of a tone (Lantz & Cuddy, 1998; see also Creel & Newport, 2002) or a chord (Jonaitis & Saffran, 2009) to a context as greater when it occurs often during the experiment, even if it does not resemble Western harmony. However, these studies focused on learning one representation across multiple musical examples rather than multiple representations. Even less work (with the exception of Hannon & Trehub, 2005b) has explored veridical memories of metrical information. Thus, it is not known how strongly specific musical memories influence music processing.

The Current Study

This study investigated whether listeners activate context memory for a particular melody when processing basic musical properties (harmony and meter). That is, when they hear a melody that previously occurred in a specific context, do they form musical expectations that are shaped by the previous specific context (*specific* expectations), expectations that are shaped only by the physically present context (*generic* expectations), or both? Each experiment had two phases. In the familiarization phase, listeners heard a small set of melodies. Each melody occurred in one of two musical contexts (Table 1) that was played simultaneously with the melody. Each listener only heard one context per melody. In the probe phase, all listeners heard each melody in isolation (without its context), followed by a probe event drawn from either the familiarized context or the other context. Probes assessed either chord goodness-of-fit (Experiments 1 and 3) or meter goodness-of-fit (Experiment 2).

Probe ratings following isolated melodies should uncover whether schematic memory, veridical memories, or both are activated during listening. If listeners' processing (activation of memories, leading to some set of musical expectations) is mediated mostly by schematic representations, then previously-heard contexts should not influence probe ratings. For instance, if a melody on its own activates C major more than A minor, all listeners should prefer the C major probe based on schematic memory activation whether or not they had previously experienced the melody in a C major context or an A minor context. If processing is affected by more detailed representations—that is, the memory of the familiarized C major versus A minor—then listeners' probe judgments should vary depending on what context they were familiarized with. In short, probe judgments consistent with the melody alone imply the activation of schematic memory, and probe judgments consistent with preceding context imply the activation of veridical memory.

Experiment 1

Experiment 1 was used to determine the effects of the familiarity with particular musical contexts on the perception of tonal goodness of fit. During familiarization, listeners heard a set of melodies (described in Table 2). Each melody occurred in one of two contexts. After repeated familiarization to each melody in a

Table 1
Sample Stimuli for Experiments 1–3

Experiment	Listening phase context		Probe phase	
	Experiment 1			
	Melody	Context	Probe A	Probe B
Listener 1	Melody X	F# major	**F# major**	D# minor
	Melody Y	G# major	E major	**G# major**
Listener 2	Melody X	D# minor	F# major	**D# minor**
	Melody Y	E major	**E major**	G# major
	Experiment 2			
	Melody	Context	3/4 Probe	6/8 Probe
Listener 1	Melody X	3/4	**3/4**	6/8
	Melody Y	6/8	3/4	**6/8**
Listener 2	Melody X	6/8	3/4	**6/8**
	Melody Y	3/4	**3/4**	6/8
	Experiment 3			
	Melody probe	Context probe	Consistent	Inconsistent
Listener 1	Melody X	C major	**C major**	A minor
	Melody Y	D# minor	F# major	**D# minor**
Listener 2	Melody X	A minor	C major	**A minor**
	Melody Y	F# major	**F# major**	D# minor

Note. **Boldface** indicates the familiar probe for the listener in the probe phase.

Table 2
Experiment 1: Stimulus Melodies

Melody	Time signature	Tempo (quarter)	Context A	Context B	Melody instrument	Context instrument(s)
1	6/8	110	F# major	D# minor	acoustic guitar	electric grand piano
2	3/4	140	E major	G# major[b]	guitar (nylon strings)	piano (A); church organ (B)
3	3/4	96	F major	A minor[b]	vibraphone	acoustic grand piano
4	4/4	130	D major	F# minor	vibraphone	pizzicato strings & piano (B only)
5	4/4	80	G# major[a]	D# minor	percussive organ	strings (A), tremolo strings (B)
6	4/4	108	B minor	G# minor	fretless bass	contrabass & voice oohs
7	4/4	80	D major[a]	D minor	English horn	French horns
8	6/8	120	Eb major	Eb minor	flute	drawbar organ
9	6/8	180	Eb major	F# major	acoustic guitar	acoustic grand piano
10	6/8	80	G major	E minor	music box	shakuhachi
11	3/4	120	Ab major[a]	Ab minor	clarinet	reed organ (A); contrabass (B)
12	3/4	80	D major[a]	A minor	flute	synthesized strings

[a] Lowered 7th scale degree. [b] Lowered 2nd scale degree.

particular context, there was a probe phase. In the probe phase, listeners heard each melody in isolation, which was followed by a probe chord. During some trials the probe chord was the tonic (most stable) triad from the context that was familiar to the listener, and during other trials the probe was the tonic triad from the unfamiliar context (the one that that listener had not heard during familiarization). Listeners were asked to rate how well the probe chord fit the preceding melody. The measure of interest was the difference in listeners' goodness-of-fit ratings for familiar contexts (the ones heard with that melody during the listening phase) versus unfamiliar contexts. If listeners' harmonic fit judgments were not influenced by melody-specific exposure, ratings would be identical regardless of what their familiarization experience was. That is, if listeners were depending mostly on long-term schematic knowledge, then they would have a chord preference for each melody, but that preference would not be changed by recent exposure. If listeners were influenced by general properties within the experiment (e.g. Jonaitis & Saffran, 2009)—particularly the experiment-wide distribution of major versus minor chords—they would show little preference for either, as contexts were roughly evenly divided between major and minor chords (54% major). Finally, if listeners were influenced by veridical memory of melody-specific contexts, then they would give higher chord-fit ratings to chords that matched the contexts that they previously heard with the melodies.

Method

Participants. Twenty-four undergraduate students from the University of California, San Diego, psychology human participant pool took part in the experiment for course credit. The musical background of each participant was assessed by a brief questionnaire. No participants reported hearing difficulties.

Stimuli. Twelve melodies were composed by the author in Finale software (2009.r2, MakeMusic, Inc.), ranging from 9–18 s long. Each melody was presented simultaneously with one of two harmonic contexts: Context A or Context B. The Context A harmonies (Table 2) were generally major-key tonalities (11/12), and the Context B harmonies were generally minor-key (10/12) or more exotic tonalities (1/12). However, the Context A for one melody was not necessarily similar to the Context A for another melody. For nine melodies, one context centered on a major triad, and the other context centered on a minor triad, usually in related musical keys. The remaining three contexts centered on two minor triads (1) or two major triads (2). Contexts spanned a wide range of major and minor keys, to reduce potential harmonic carryover effects (e.g., a preceding melody's C major context biasing a listener toward a C major interpretation of the following A minor context of a different melody). The intensity of the melodic line was on average 18.1 points higher (in MIDI velocity) than the accompaniment (context) voices (SD = 17.2), though intensity was adjusted for individual melodies to make the melodic line prominent in that particular instrumentation. After composition, melodies with contexts, melodies without contexts, and probe chords were exported from Finale as .aiff files. The identical melody sound file was used on both Context A and Context B probe trials, with a different probe chord presented at the end.

Procedure. Participants were located in a sound-treated room and used a Macintosh Mini computer with PsyScope X Build 51 software (Cohen, MacWhinney, Flatt, & Provost, 1993; downloaded from http://psy.ck.sissa.it/ on 9/25/2007). Sounds were presented via Sennheiser HD 280 Pro headphones adjusted to a comfortable volume.

During familiarization, each participant heard eight melodies in particular harmonic contexts (examples are in Figure 1). Each melody was heard in context 10 times. For each participant, four melodies were heard in Context A and four melodies were heard in Context B. The remaining four melodies were left out for a given listener until the probe phase (see Table 3 for example lists). The pairing of melodies with Context A versus Context B as well as what melodies were left out were counterbalanced across participants (Table 3). The reason for this matched design—identical melodies with nonidentical contexts—was to circumvent the problem that in naturally occurring music melodies tend to contain properties related to their contexts. By counterbalancing exposure, the degree of influence of context exposure alone (rather than the properties of the melody itself) could be determined. The harmonic information present in melodies alone and melodies in context is presented in Table 4. In particular, this table presents correlations of note counts for each melody-context combination, with vectors of tone goodness-of-fit (Krumhansl, 1990). The correlation in each

Figure 1. Experiment 1. (a) Example melody in Context A (natural minor). (b) Example melody in Context B (Mixolydian, similar to major). During probe trials, only the melody line (top of each) was played. Note that because the 3rd scale degree was omitted, the melody was consistent with either a major or minor interpretation.

cell is between the note count and key of the A probe or B probe. Higher correlations verify that contexts provided the intended key information.

As a cover task, to maintain participants' attention without asking them to explicitly encode any musical attributes, the participants were asked to rate their affective interpretation of the melody (on a continuum from sad to happy) and of the degree to which they liked the melody (on a continuum from dislike to like a lot) by mouse clicking a two-dimensional display with axes labeled by the rating dimensions. These ratings were not analyzed.

After there were 10 presentations of each melody in its context, each listener proceeded to a probe phase. In this phase, listeners heard all 12 melodies devoid of context—that is, without accompaniment—followed after 500 ms of silence by a single probe chord. The probe chord was either the tonic triad from Context A or the tonic triad from Context B. Each probe chord occurred at a time point after the melody in which listeners never heard a chord, which prevented the listeners from expecting a musical event at that time point. For a given melody, the top notes of both A and B probe chords were identical, to account for melodic continuity effects across contexts (i.e., the top note of one chord being closer in pitch height to the last note of the melody). During each trial participants were asked to rate the goodness-of-fit of the probe chord to the preceding melody.

Each melody was presented with two probe chords (on different trials): one that matched Context A for that melody and one that matched Context B for that melody. Ratings were made by clicking a mouse cursor on a left-to-right graphical ruler denoting goodness of fit. The x-value of the mouse click was converted into a deviation from the central point of the rating scale: The center x-coordinate of the ruler was subtracted from the x-value to yield distance from the central ("Okay") rating, and was then divided by half the length of the ruler (in pixels). Thus, this measure could vary from -1 to $+1$, with 0 being the neutral point. Each melody and probe combination was tested twice.

Design. Familiarization (A, B, or none) and probe chord (A or B) were both within-participants factors. During familiarization, each participant heard four Context-A-familiarized melodies and four Context-B-familiarized melodies. Melody and context pairings were distributed across six lists so that across participants each melody and context pairing was equally likely, and each listener heard roughly the same number of major- and minor-tonality contexts. The probe phase was identical for all participants (except for different randomizations): Each participant heard every isolated melody four times, twice followed by an A Probe and twice followed by a B Probe. The presentation of unfamiliarized melodies in the probe phase afforded an assessment of whether Context A's or Context B's were generally more preferable, allowing for a clearer assessment of preference shifts.

Results

Probe rating reliability was calculated as the correlation between the first and second instances of melody-probe combinations for each participant. This ranged from .28–.81, with median reliability at .53.

Listeners' chord fit ratings for each melody (displayed in Figure 2) were strongly influenced by the contexts that they had been exposed to: Listeners gave a probe chord higher ratings if it matched the context that they had been familiarized with than if it did not match the context. For melodies heard in Context A, listeners rated A Probes higher than B Probes, and for melodies

Table 3
Experiment 1: Example Exposure to Contexts for Different Listeners

	Tonality		Example familiarization		
Melody	Context A	Context B	Listener 1	Listener 2	Listener 3
1	F# major	D# minor	Context A	—	ContextB
2	E major	G# major	Context B	Context A	—
3	F major	A minor	—	Context B	ContextA
4	D major	F# minor	Context A	—	ContextB
5	G# major	D# minor	Context B	Context A	—
6	B minor	G# minor	—	Context B	ContextA
7	D major	D minor	—	Context B	ContextA
8	Eb major	Eb minor	Context A	—	ContextB
9	Eb major	F# major	Context B	Context A	—
10	G major	E minor	Context B	Context A	—
11	Ab major	Ab minor	—	Context B	ContextA
12	D major	A minor	Context A	—	ContextB

Table 4
Correlations of Each Melody, Alone and in Contexts A and B, With Krumhansl-Kessler Tonal Profiles (Krumhansl, 1990)

	Melody only		In context A		In context B	
Melody	A	B	A	B	A	B
1	0.70	**0.75**	**0.91**	0.79	0.65	**0.89**
2	0.27	**0.49**	**0.82**	0.11	0.30	**0.69**
3	0.49	**0.51**	**0.83**	0.65	0.61	**0.80**
4	0.34	**0.81**	**0.88**	0.66	0.42	**0.89**
5	0.61	**0.73**	**0.86**	0.42	0.55	**0.91**
6	0.68	0.44	**0.89**	0.29	0.29	**0.89**
7	0.32	**0.81**	**0.80**	0.69	0.63	**0.86**
8	**0.86**	0.74	**0.77**	0.44	0.62	**0.77**
9	**0.41**	0.21	**0.51**	0.16	0.26	**0.45**
10	**0.33**	0.29	**0.78**	0.39	0.44	**0.61**
11	**0.77**	0.53	**0.88**	0.54	**0.79**	0.76
12	0.46	**0.87**	**0.77**	0.66	0.23	**0.94**
Mean	0.52	0.60	0.81	0.48	0.48	0.79
SD	0.30	0.22	0.11	0.22	0.18	0.15

Note. Correlations are between note frequencies in the melody with the tonal profile of the key of Context A or Context B. For example, the correlation of 0.70 in the top left corner represents the correlation between the note frequency-of-occurrence vector for melody 1 and the F# major tonal profile. Bold indicates the higher correlation in each context (e.g. in context A).

heard in Context B, listeners rated B Probes higher than A Probes. For melodies withheld until the probe phase, there was a small preference for A Probes.

An ANOVA on chord ratings for previously heard melodies with Probe Chord (Context A probe, Context B probe) and Familiarization (Context A, Context B) as within-participants and within-items factors confirmed the assessments mentioned in the previous paragraph. Effect sizes are reported as generalized eta-squared (η^2_G), which corrects for inflated estimates of partial η^2 (Bakeman, 2005), and Cohen's *d* for *t*-tests.

There was an effect of Probe Chord, $F1(1, 23) = 20.9$, $p = .0001$; $F2(1, 11) = 4.68$, $p = .05$, $\eta^2_G = .05$, demonstrating an overall preference for Context A probes. This likely reflects schematic preferences for major contexts. There was no effect of Familiarization alone, $F1F2(1, 11) = 2.12$, $p = .17$, $\eta^2_G = .00$.

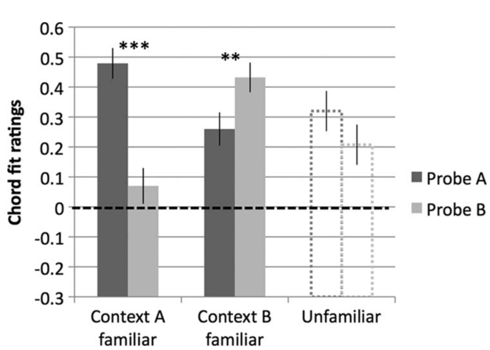

Figure 2. Experiment 1. Listeners' ratings of chords following isolated melodies. Context A probes (dark gray bars) and Context B probes (light gray) are shown. Bars with dotted outlines denote trials where the melody had not been heard before the test phase. The dashed black line indicates the midpoint of the rating scale. Error bars are standard errors. ** $p < .01$. *** $p < .001$.

There was a large interaction of Familiarization × Probe Chord, $F1(1, 23) = 39.67$, $p < .0001$; $F2(1, 11) = 56.96$, $p < .0001$, $\eta^2_G = .24$. This resulted from higher ratings for Context A probe chords when the melody was familiarized in Context A, $t1(23) = 8.38$, $p < .0001$; $t2(11) = 5.51$, $p = .0002$, $d = 1.5$, but higher ratings for Context B probe chords when the melody was familiarized in Context B, $t1(23) = 3.04$, $p = .006$; $t2(1,11) = 2.93$, $p = .01$, $d = .67$. Thus, chord preferences were strongly shaped by the prior listening context. For melodies that were left out of the familiarization for a given listener, there was a nonsignificant tendency to prefer Context A probes, $t1(23) = 1.74$, $p = .09$; $t2(11) = 1.5$, $p = .16$; $d = .34$.

Effects of musical experience. One might wonder whether listeners with more musical experience would demonstrate better memory in a listening context, which would affect the generalizability of these results. The number of years in which listeners had participated in individual or group music-making was calculated ($M = 3.6$, $SD = 4.0$, range = 0–12). One participant's questionnaire data were lost, so that participant was omitted. The correlation between years of musical experience and the size of the context-match effect (match rating minus mismatch rating) was then calculated. The correlation was low and nonsignificant, $r = .19$, $t(21) = .87$, $p = .21$, suggesting that musical experience did not strongly affect listeners' memories for previous contexts.

Discussion

Listeners' ratings of chord fits to isolated melodies were influenced by previously hearing the melodies in particular harmonic contexts. That is, listeners gave higher ratings to chords that matched the specific harmonic context that they had previously heard with the melody. This is consistent with the hypothesis that listeners were reinstating harmonic contexts upon hearing the isolated melody and that this reinstated context modulated their perceptions of chord fit. The results substantiate the role for veridical representations (not just schematic ones) in music processing.

However, the influences of veridical knowledge of harmonic context on processing may be particularly easy to learn. That is, we know that listeners are good at keeping track of harmonic probabilities in artificial contexts (Jonaitis & Saffran, 2009), that listeners can readily accommodate to an unfamiliar tonal system (Castellano, Bharucha, & Krumhansl, 1984; Krumhansl et al., 2000), and that listeners readily integrate melodic information with other types of information such as sentences (e.g., Serafine, Crowder, & Repp, 1984; Serafine, Davidson, Crowder, & Repp, 1986). Perhaps harmonic information is something that is easily learned or relearned in an experimental context. In contrast, other musical properties such as meter may still be driven predominantly by schematic knowledge.

Meter, even more than tonality, seems to be strongly linked to schematic knowledge. For instance, even trained musicians have more difficulty finding a beat in music that does not have a large proportion of events at a particular phase and period (Snyder & Krumhansl, 2001), and multiple days of exposure to a new metrical pattern does not facilitate processing in adults (Hannon & Trehub, 2005b). These results suggest that there is resistance to the encroachment of veridical information. This fits with the notion that listeners calculate the metrical properties of a stimulus with regard to metrical schemas, with veridical knowledge playing little to no role in metrical processing. That is, the meter of a stimulus that is presented immediately may be so compelling on its own that veridical knowledge is useless. However, if listeners routinely encode the metrical properties of the melody's context, they may be able to derive metrical as well as harmonic information from veridical memory. To address this possibility and extend the results of Experiment 1 to another aspect of musical patterning, Experiment 2 tested the effects of prior listening experiences on listeners' encodings of meter.

Experiment 2

This experiment tested listeners' metrical apprehensions of melodies after hearing each melody simultaneously with one of two metrical contexts (accompaniments). As in Experiment 1, listeners heard several melodies in particular contexts. Each listener heard only one of two metrical contexts for each melody. Then their metrical encoding of each melody without context was tested by presenting the melody in isolation followed by a metrical probe. To make the task—presumably a rather difficult one—as easy as possible, listeners were asked to verify a metrical continuation rather than tapping along, which eliminated motor control as a source of difficulty. The two metrical possibilities were also made highly discriminable (3/4 meter vs. 6/8 meter, which differed in both beat duration and beat subdivisions). Finally, only eight melodies were used (each with two contexts). Unfamiliarized baseline judgments were provided by a separate set of participants to keep the experimental session under one hour.

The question of interest was whether listeners' judgments for a metrical continuation of the melody would be influenced by the metrical properties of a familiarized context. If listeners were influenced only by schematic expectations or preexisting biases, then they would not show effects of familiarity in their metrical goodness-of-fit judgments. If they were influenced by the general metrical tendencies within the experiment, their metrical judgments would not favor either 3/4 or 6/8, since each is equally frequent during familiarization. Finally, if they were influenced by veridical memory for the meter for particular melodies, then they would favor 3/4 probes for melodies heard in 3/4 contexts and 6/8 probes for melodies heard in 6/8 contexts.

Method

Participants. Forty-eight undergraduates from the same pool as Experiment 1 participated in the experiment. Of these participants, 24 were in the experimental condition and 24 were in the control condition. Note that the use of a separate control conditions meant that experimental and control participants could have schematic knowledge that was not identical. There was extra statistical noise introduced by using two sets of participants, but this noise should not be systematic.

Stimuli. Eight melodies (described in Table 5) were constructed in Finale that could reasonably be construed in either 3/4 or 6/8 time. Music in 6/8 time and music in 3/4 time contain six eighth notes per bar (marked by the lowest rows of x's in Figure 3a and 3b). However, music in 3/4 time was more likely to have note onsets on the first, third, and fifth eighth notes (middle row of x's, Figure 3a), and music in 6/8 time was more likely to have note onsets on the first and fourth eighth notes (middle row of x's, Figure 3b). To create melodies that could fit with either time, the frequencies of onsets on eighth notes three and five (indicative of 3/4) and four (indicative of 6/8) were similar.

Two metrical contexts were constructed for each melody, one in 3/4 (Figure 3a) and one in 6/8 (Figure 3b). Contexts, unlike the

Table 5
Experiment 2: Stimuli

Melody	Key	Tactus, 3/4	Tactus, 6/8	Melody instrument	Context instrument(s)
a	C major[a]	120	80.0	soprano recorder	strings
b	C major	100	66.7	oboe	piano & strings
d	A major[b]	140	93.3	acoustic guitar	church organ
e	E minor	120	80.0	flute	harp
f	G major	140	93.3	xylophone	pizzicato strings
g	C major[c]	120	80.0	acoustic guitar	tango accordion
h	F# minor	100	66.7	alto sax	trombone, bass trombone, & cymbal
i	C major	140	93.3	flute	harpsichord

[a] Lowered 7th scale degree. [b] lowered 2nd scale degree. [c] raised 4th scale degree.

Figure 3. Experiment 2. The first four measures of an example melody with a 3/4 context (a) and a 6/8 context (b). Metrical grids are marked in x's at the top.

melodies, were constructed to be metrically unambiguous, with most note onsets falling on eighth notes 1, 3, and 5 in 3/4 contexts, and 1 and 4 in 6/8 contexts. Correlations of numbers of note onsets to idealized 3/4 patterns (2 0 1 0 1 0) and 6/8 patterns (2 0 0 1 0 0) are provided in Table 6. These correlations follow Palmer and Krumhansl's (1990) instantiation of Lerdahl and Jackendoff's (1983) metrical grid. Note that these correlations were much lower for melodies out of context than melodies in context, suggesting that the melodies alone were relatively amenable to either meter. Also notice that the melodies in context correlate strongly with particular meters.

Expressive playback options in Finale were turned off, so that Finale did not impose different sets of metrical emphases on the two different versions of a melody. All isolated melodies and a series of drumbeats (described in the next section) were generated from Finale files that were identical (all in 6/8 meter), except for the timing of the drumbeats. After composition, melodies were exported from Finale as .aiff files.

Procedure. As in Experiment 1, the experiment was run in PsyScope X. Also as before, there was a familiarization phase followed by a probe phase. In the familiarization phase, each participant heard four melodies with a 6/8 context and four melodies with a 3/4 context. Each melody was presented eight times during familiarization as the participants completed the same mood and preference rating tasks as in Experiment 1. Control participants did not take part in the familiarization phase. In the probe phase, all participants heard all melodies four times, twice with a 3/4 probe and twice with a 6/8 probe.

Prior to the probe phase, listeners were instructed that they would hear melodies followed by drumbeats and that the drumbeats were either correct or incorrect. To provide a frame of reference, they were given two examples of "Happy Birthday", one followed by correct drumbeats and one by incorrect drumbeats, and two examples of "Greensleeves" followed by correct and incorrect drumbeats. They could play these examples as many times as needed before continuing. After this, they proceeded to the probe phase, where they made goodness-of-fit judgments of metrical probes for the actual melodies that they had heard in the familiarization phase.

During the probe phase, each melody was heard without accompanying context, followed by a metrical probe in one of the two possible meters. The metrical probe was a series of 9 (in 6/8) or 13 (in 3/4) pulses in woodblock timbre. Pulses began on the start of the measure after the last tone onset of the melody and continued for four full measures plus the downbeat of the fifth measure. This way, the metrical probe did not temporally coincide with the melody itself. The only way that listeners could show a preference for a probe was by calculating a period and phase during the melody, and continuing to "feel" that period and phase long enough to compare to the metrical probe. Note that this measure predominantly assessed period (time between beats), not phase (temporal locations of beats). This was deliberate in that early tests of the author (a classical musician with over 20 years of training and practice) suggested that phase differences in metrical probes

Table 6
Experiment 2: Metrical Template Correlations for Isolated Melodies and Melody and Context Combinations

	Melody alone		Melody + 3/4 context		Melody + 6/8 context	
Melody	3/4 meter	6/8 meter	3/4 meter	6/8 meter	3/4 meter	6/8 meter
A	0.35	**0.55**	**0.90**	0.47	0.40	**0.90**
B	0.48	**0.62**	**0.85**	0.42	0.69	**0.94**
D	**0.51**	0.37	**0.96**	0.60	0.64	**0.98**
E	**0.70**	0.34	**0.94**	0.28	0.79	**0.92**
F	0.56	**0.76**	**0.93**	0.48	0.44	**0.94**
G	0.58	**0.82**	**0.95**	0.51	0.45	**0.95**
H[a]	**0.85**	0.60	**0.95**	0.67	0.54	**0.93**
I	**0.62**	0.39	**0.91**	0.29	0.57	**0.92**
Mean	0.58	0.56	0.92	0.47	0.56	0.94
SD	0.15	0.18	0.04	0.14	0.14	0.02

Note. The **boldface** correlation is the higher correlation with idealized metrical patterns in each pair.
[a] Naïve listeners preferred 3/4 continuations of all melodies except this one.

were more difficult to detect than period differences. Ratings were made by clicking a mouse cursor on a graphical ruler (the same as Experiment 1) denoting goodness of fit.

Design. Familiarization (3/4 or 6/8) and probe meter (3/4 or 6/8) were within-participants factors. This differed from Experiment 1 only in that a different set of listeners completed a no-exposure condition (hearing and rating probes only after isolated melodies, without familiarization). The assignment of each melody to each context was counterbalanced over participants, so that across eight conditions each melody occurred equally often with a 3/4 context or a 6/8 context, and each participant heard the same number of 3/4 and 6/8 contexts.

Results

Probe rating reliability was calculated for each participant. For control participants, median reliability was .41, ranging from –.24 to .97. For experimental participants, reliability was somewhat higher, with a median of .71, ranging from .20 to .99.

As shown in Figure 4, listeners provided higher ratings for metrical continuations of isolated melodies that matched the contexts that they had previously heard. It is interesting that the unfamiliar tempos in each case were below the midpoint of the rating scale, implying that listeners seemed to regard them as actually erroneous, rather than simply less good. For listeners in the no-exposure condition, there was an overall preference to hear the isolated melodies in 3/4 meter. One potential reason for this 3/4 bias is that subdivisions of the beat into groups of two notes are more common than subdivisions of the beat into groups of three notes in Western music. Hannon, Snyder, Eerola, and Krumhansl (2004), who had listeners report whether they were perceiving 6/8 meter or 3/4 meter, also found a general bias toward 3/4. In any case, the 3/4 preference of naïve listeners suggests that listeners derived reasonably strong ideas of each melody's meter based on the melody by itself, but that listeners who had heard melodies in context were more strongly influenced by context than by a priori preferences.

An ANOVA on meter-fit ratings with Probe Meter (3/4, 6/8) and Familiarization (3/4 context familiar, 6/8 context familiar) as within-participants and within-items factors confirmed the assessment in the previous paragraph. The effect of Probe Meter did not reach significance, $F1(1, 23) = 2.26$, $p = .15$; $F2(1, 7) = 4.45$, $p = .07$, $\eta^2_G = .04$, suggesting that there was no overall preference for either meter. The effect of Familiarization did not approach significance (both Fs < 1). The only significant effect was the interaction of Probe Meter × Familiarization, $F1(1, 23) = 40.31$, $p < .0001$; $F2(1, 7) = 47.62$, $p = .0002$, $\eta^2_G = .46$. This interaction resulted from higher goodness-of-fit ratings of the 3/4 probes for melodies familiarized in a 3/4 context, and higher ratings of 6/8 probes for melodies familiarized in a 6/8 context. Overall, ratings for the "wrong-meter" continuations were significantly below the midpoint of the rating scale, $t1(23) = 3.39$, $p = .003$; $t2(7)=3.74$, $p = .007$, $d = .98$, suggesting that listeners regarded these metrical continuations not just as less preferable but as significantly less than "okay." Finally, a t-test confirmed that for no-context listeners 3/4 continuations were preferred to 6/8 continuations, $t1(23) = 3.30$, $p = .003$, $t2(7) = 3.06$, $p = .02$, $d = 1.28$, suggesting that there was a preexisting bias toward 3/4 time.

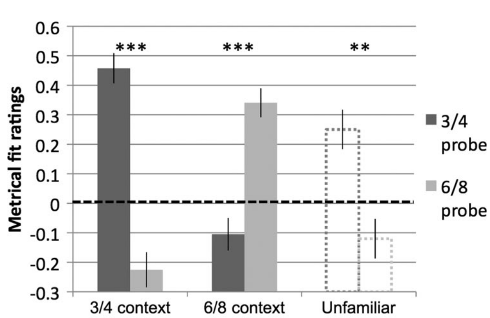

Figure 4. Experiment 2. Listeners' ratings of metrical continuations of isolated melodies. Listeners who heard a melody familiarized in a 3/4 context preferred the 3/4 probe (dark gray) over the 6/8 probe (light gray), and the reverse was true for melodies familiarized in a 6/8 context. ** $p < .01$. *** $p < .001$.

Effects of musical experience. As in Experiment 1, one might wonder whether more musically-experienced participants were better at recalling matches to previous meter. Accordingly, the correlation between years of musical experience ($M = 6.8$, $SD = 5.6$, range $= 0–16$) and the size of the context-match effect (ratings for familiar metrical context minus ratings for unfamiliar metrical context) was computed. Unlike Experiment 1, this correlation reached significance, $r = .47$, $t(22) = 2.49$, $p = .01$. This implies that more musically-experienced listeners made stronger distinctions between familiar metrical continuations and unfamiliar metrical continuations. Nonetheless, even the participants in the lowest third of musical experience (0–3 years) showed a significant preference for familiar metrical continuations, $t(7) = 2.39$, $p < .05$, $d = 1.48$, suggesting that the effect is present even in those with relatively little musical experience.

Discussion

Listeners' ratings of metrical fit to melodies were influenced by their memories of particular melody-meter combinations. They seemed to incorporate memory of previous musical experiences into their apprehension of meter. To the author's knowledge, this is the first result of its kind: A strong influence of perceptual memory on the apprehension of what is often thought to be a basic perceptual property of musical experience (though see Phillips-Silver & Trainor, 2005, 2007, for the effects of motor encoding on meter perception in a repeating rhythmic pattern).

A curious aspect of these results is the negative ratings for the unlearned meters. There were no such negative ratings for chords in Experiment 1. Why might this be the case? Are listeners more certain about their meter judgments than their tonality judgments? Though this may be so, a more likely explanation is that the two meters in Experiment 2 were mutually exclusive, whereas the two chord choices in Experiment 1 were less exclusive of one another, in most cases containing related harmonic material. Listeners in harmonic probe-type experiments provide the highest ratings for tonic notes or tonic triads, but they still provide above-mean ratings of related or in-key material (e.g. Krumhansl, 1990). Thus, listeners in the two experiments were accurately reflecting the graded harmonic relations in the first experiment and the more stark metrical oppositions in Experiment 2. It is also the case that the probe events in Experiment 1 were timbre-appropriate to the melodies and contexts, and the probe in Experiment 2 was not timbre-appropriate. This issue is raised again in the General Discussion.

Experiments 1 and 2 show that two basic properties of musical experience seem to be strongly influenced by specific memory constructed from previously hearing a melody. This suggests that veridical knowledge has a robust role in processing musical structures whereby hearing music preferentially activates the most similar memories of music, which then influences interpretation.

Of course, one might reasonably argue that these first two experiments have demonstrated that preceding context can affect listeners' musical processing intuitions when no intuitions exist. That is, the melodies used in Experiments 1 and 2 were engineered to be tonally or metrically ambiguous. It is interesting that in Experiment 2 listeners seemed to have a priori intuitions about the melodies' meters, favoring 3/4 meter despite attempts to make the melodies ambiguous. Experiment 1 was more successful in creating tonal ambiguity. This ambiguity, which failed to activate schematic knowledge strongly, might have led listeners to infer tonal or metrical properties based on the only other memory source that they had, the context of that earlier melody.

There are two things to be said about this objection. First, it is consonant with similarity-based memory activation models (Goldinger, 1998; Hintzman, 1986): If a signal (the melody) has low similarity to existing memory traces (the sum of musical knowledge), then it will not activate those traces strongly. Second, another prediction from similarity-based models is that if a signal is highly consistent with a schema—meaning that it resembles much of musical experience—then many traces will be activated in addition to the familiarized context traces, resulting in something that looks like stronger activation of a schema. This is different from a strong-schema perspective in that schema activation would negate the effects of veridical context.

The aforementioned prediction was tested in a final experiment. In Experiment 3, listeners were exposed to strongly schematic melodies, each designed to evoke a single musical key. Each melody was heard by some listeners in a schema-consistent context and by other listeners in a schema-inconsistent context. Much like Experiment 1, ratings of chords from schematic contexts and nonschematic contexts were compared to assess whether previous context affected harmonic fitness judgments. If listeners were influenced solely by schematic knowledge, their chord ratings would be higher for schema-consistent chords, regardless of familiarized contexts. However, if listeners were also influenced by previous context, ratings would favor probe chords matching the context in which the melody had been heard in the familiarization phase.

Experiment 3

Participants. Thirty-six participants from the same pool as the first two experiments took part in the experiment. This number was increased relative to previous experiments in order to detect what was potentially a smaller effect. Music experience was comparable to previous experiments (M = 4.3 years, SD = 3.9 years, range = 0–14).

Stimuli. Twelve melodies (described in Table 7) were constructed so that each melody strongly implied a particular tonal center. This was accomplished by emphasizing the first and fifth scale degrees, frequently presenting notes from the tonic triad (such as C, E, and G in C major) in sequence, and ending the melody on the tonic note (C). Half of the melodies were constructed to center around a major tonality, and the other half were constructed to center around a minor tonality. This manipulation was verified by computing correlations with Krumhansl's (1990) tonal profiles of scales (Table 8) and by listener ratings of isolated melodies. Half of the melodies were in 4/4 meter, and the rest were in triple meters (3/4 or 6/8). Meters were evenly divided between major and minor tonalities. Two harmonic contexts were composed for each melody (examples in Figure 5): one that was schema-consistent (i.e., matching the strongly implied key in the melody) and one that was schema-inconsistent. Unlike Experiment 1, probe chords were presented in an organ timbre, which did not occur in any of the contexts or melodies. This eliminated timbral familiarity, which was present in Experiment 1, as a source of goodness-of-fit judgments.

Procedure. As in Experiment 1 each listener was exposed to 8 of the 12 melodies, with the remaining four occurring for the first time during the probe phase. These novel melodies allowed for verification that the melody's intended tonality was in fact the chord that the participants preferred. Melodies that were novel and melodies that were heard in typical or atypical contexts were counterbalanced across participants. Four of the melodies that each participant heard had a major melody and four had a minor

Table 7
Experiment 3: Melodies and Their Harmonic Contexts

Melody	Time signature	Tempo (quarter)	Consistent key	Inconsistent key	Melody instrument	Context instrument(s)
1	44	90	C major	A minor	pan flute	piano
2	44	90	D major	D minor	trumpet	French horn & trombone
3	44	90	E major	C# minor	piano	piano & violin
4	44	120	F# minor	F# major	alto saxophone	piano & electric bass
5	44	120	C minor	Eb major	piano	piano
6	44	120	G# minor	E major	harp	violin, viola, & cello
7	68	100	E minor	G major	French horn	marimba
8	34	200	Bb minor	Db major	oboe	harpsichord
9	68	60	Bb major	G minor	accordion	strings
10	34	120	D minor	Bb major	piccolo	flute, oboe, & bassoon
11	68	120	F# major	D# minor	flute	oboe & bassoon
12	34	140	Ab major	F minor	clarinet	guitar (nylon strings)

Table 8
Correlations of Melodies and Contexts With Tonal Profiles (Krumhansl, 1990)

	Melody alone		Typical context		Atypical context	
Melody	Typical key	Atypical key	Typical key	Atypical key	Typical key	Atypical key
1	**0.85**	0.32	**0.92**	0.46	0.72	**0.84**
2	**0.94**	0.64	**0.86**	0.69	0.72	**0.85**
3	**0.87**	0.49	**0.96**	0.51	0.32	**0.83**
4	**0.79**	0.61	**0.97**	0.50	0.51	**0.71**
5	**0.80**	0.55	**0.80**	0.53	0.72	**0.76**
6	**0.79**	0.16	**0.92**	0.26	0.52	**0.77**
7	**0.76**	0.23	**0.84**	0.63	0.55	**0.85**
8	**0.86**	0.42	**0.97**	0.50	0.86	0.86
9	**0.87**	0.58	**0.94**	0.55	0.73	**0.90**
10	**0.65**	0.12	**0.83**	0.25	0.64	**0.89**
11	**0.92**	0.64	**0.93**	0.65	0.59	**0.94**
12	**0.87**	0.66	**0.96**	0.68	0.73	**0.89**
Mean	0.83	0.45	0.91	0.52	0.63	0.84
SD	0.08	0.20	0.06	0.15	0.14	0.07

Note. Boldface in Table 8 marks the higher correlation in a particular context (e.g. in the typical context).

melody: Half of the melodies were presented in a context that was consistent with the tonality emphasized in the melody, and half of the melodies were presented in what will be referred to as the "atypical" context. In other respects the procedure was the same as in Experiment 1.

Design. The design was identical to the design in Experiment 1 except that instead of Context A and Context B familiarizations and probes, listeners heard schema-consistent and schema-inconsistent familiarizations and probes.

Results

Due to a glitch in PsyScope X, the first test trial for each participant recorded the presentation of the wrong melody and had to be discarded. Fortunately, the affected stimulus was random and was measured elsewhere for each participant since each trial was presented twice. One participant was accidentally run in the wrong condition, but the participant was included nonetheless because the results did not pattern differently when this participant was excluded. As before, probe rating reliability was calculated for each participant. Median reliability was .37, ranging from −.16 to .95.

First, the probe ratings of novel melodies were examined to verify that the intended tonality was in fact the one that was preferred by participants. A *t*-test on probe ratings for typical chord probes versus atypical chord probes for novel melodies (Figure 6, right) showed that typical probes were strongly preferred over atypical ones, $t1(35) = 5.63$, $p < .0001$; $t2(11) = 6.86$, $p < .0001$; $d = 1.02$. Further, each of the 12 melodies, when it had not been presented in a context during familiarization, showed higher ratings for the typical probe than the atypical probe.

The effects of exposure context were assessed using an ANOVA on probe ratings with Familiarization (schema-consistent, schema-inconsistent) and Probe Chord (schema-consistent, schema-inconsistent) as within-participants factors. There was an effect of Probe Chord, $F1(1, 35) = 36.03$, $p < .0001$; $F2(1, 11) = 40.92$, $p < .0001$; $\eta^2_G = .174$, with consistent probes receiving higher ratings overall. This reflects influences of activation of schematic knowledge. The effect of Familiarization did not reach significance (*F*s < 1), but the Familiarization × Probe Chord interaction was significant, $F1(1, 35) = 7.5$, $p < .01$; $F2(1, 11) = 18.92$, $p = .001$; $\eta^2_G = .027$. In particular, the difference between ratings for

Figure 5. Experiment 3. Sample melody in a schema-consistent context (a) and a schema-inconsistent context (b).

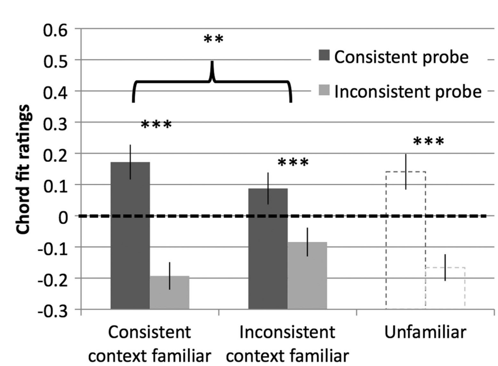

Figure 6. Experiment 3. Schema-consistent and schema-inconsistent probe ratings by exposure context. ** $p < .01$. *** $p < .001$.

consistent and inconsistent probes was diminished when listeners were familiarized with an inconsistent context. The *t*-tests suggested that this resulted from an increase in inconsistent probe ratings, $t1(35) = 2.41$, $p = .02$; $t2(11) = 3.06$, $p = .01$; $d = .40$, and possibly a decrease in consistent probe ratings, $t1(35) = 2.09$, $p = .04$; $t2(11) = 1.21$, $p = .25$; $d = .26$.

One possible explanation of this pattern of results might be that only some listeners were swayed by the inconsistent exposure, and others were unaffected by previous exposure. If this were the case, then there should be a negative correlation between the effects of exposure context and the effects of schema consistency. To assess whether this correlation existed, two scores were calculated for each participant: a veridical-fit score (ratings for context-matching probes minus context-mismatching ones, regardless of schema consistency) and a schematic score (ratings for consistent minus inconsistent probes, regardless of the context that was heard). The correlation between these two scores (Figure 7) was then calculated. The correlation was positive and significant, $r = .37$, $t(34) = 2.34$, $p = .03$, suggesting that both sources were used more often by listeners who made finer discriminations. Figure 7 also reveals two outliers, which lay more than three standard deviations from the mean of each score. The correlation is stronger when these outliers are eliminated, $r = .59$, $t(32) = 4.16$, $p = .0002$. Overall, it does not seem that schematic and contextual knowledge negate each other, but rather that they work together to determine the tonal percept, which is similar to Justus and Bharucha's (2001) third experiment.

Effects of musical experience. As with the first two experiments, one wonders whether listeners with more musical experience might be more sensitive to the musical exposure manipulation. A counterprediction for the experiment was that those with musical experience might be less sensitive to the manipulation due to their greater attentive exposure to Western music. To assess these possibilities each listener's interaction score was calculated as the difference score between consistent and inconsistent responses for consistent exposure minus the difference score between consistent and atypical responses after inconsistent exposure. The calculation indexes the size of the change in schematic preference as a result of exposure to schema-inconsistent material. The result was correlated with the number of years each participant had experience playing music. Two participants' musical experience data were missing, so these were excluded. The correlation was weak and not significant, $r = .16$, $t(32) = .91$, $p = .37$, though it approached significance when the two outliers were removed, $r = .33$, $t(32) = 1.92$, $p = .06$. The results suggest that the relationship between musical experience and the sensitivity to new contextual information was not strong.

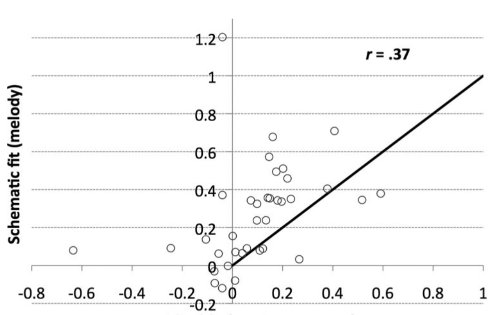

Figure 7. Experiment 3. Familiarization context (veridical) effect on ratings × melodic content (schematic) effect on ratings. The black line indicates equivalent effects of each.

Discussion

Listeners heard melodies designed to activate harmonic schemas. Half of the melodies that each listener heard were familiarized in a schema-consistent context and half were familiarized in a schema-inconsistent context. Although listeners preferred the schema-consistent probe chords overall, this preference was diminished when listeners were familiarized with schema-inconsistent contexts. This outcome suggests that schematic knowledge—or something that looks like schematic knowledge—shapes music processing, and newly formed context-specific representations shape processing as well.

General Discussion

Three experiments showed that listeners' memories of particular musical contexts were activated during music listening. In Experiment 1, listeners who were familiarized with a melody in a particular harmonic context stored information about that harmonic context along with the melody. That is, when listeners heard the melody in isolation followed by a single chord, they gave higher ratings to a chord when it fit the particular harmonic context that they were familiar with. Listeners in Experiment 2 preferred metrical continuations of isolated melodies when the metrical continuations fit the meter in which the melody had been familiarized. Finally, listeners in Experiment 3 who heard strongly harmonically-schematic melodies in schema-consistent or schema-inconsistent contexts showed the effects of the schematic properties of the melody, but these properties were moderated by experimental exposure to schema-inconsistent contexts. These results suggest that listeners not only encoded veridical contextual information about harmony and meter, but they also reactivated that information during processing when the melody was heard alone. In sum, for both harmony and meter, listeners appeared to activate melody-specific contexts from their prior listening experiences during music listening.

This study suggests that music processing, much like language processing, is guided by relatively specific knowledge. Listeners not only form highly specific musical memories, but they also employ those memories during processing. This means that musically-specific representations have a functional role in music understanding. Further, this suggests that highly specific memories complement schematic biases in processing, and when schematic memories are weak, specific memories may override them. This suggestion is itself consistent with the prediction of exemplar-style memory activation models (Carpenter & Grossberg, 2003; Goldinger, 1998; Hintzman, 1986) that detailed memories are more likely to shape processing when there are relatively few instances in memory that match the input. It also seems true that listeners activate specific contexts even when there are strong influences of schema-like memory.

Comparison to past work suggesting schematic expectations. The current study found strong effects of veridical musical memory in addition to the role of schematic knowledge. As such, this study may seem to be somewhat at odds with prior work demonstrating strong schematic expectations. In particular, Justus and Bharucha (2001), Hannon and Trehub (2005a, 2005b), and Krumhansl (1979) found strong effects of listeners' schemas. Why were listeners' expectations so easily modified in the current study? There are several reasons that these seeming discrepancies may have occurred, which center on differences in the measures used across studies and differences in the types of musical material that was used and the extent to which it was used. The following discussion is intended for reconciling these results.

One difference between the current work and previous work showing schematic activation is the measure used: The authors of the current study used goodness-of-fit judgments, and others used reaction times (Justus & Bharucha, 2001). Those authors found strong schematic effects in their data. In regards to reaction time, it may be the case that the listeners in the current study experienced schematic processing effects prior to making their musical goodness-of-fit judgments, but since responses were not speeded, veridical knowledge had time to emerge (see Bharucha & Todd, 1991, for a model consistent with this possibility). A parallel argument for the role of word representations was made by McLennan and Luce (2005). They theorized that abstract word knowledge (phonemes) is activated more rapidly than more specific knowledge (who has recently spoken a word). Their results from a lexical decision task were consistent with this suggestion: Listeners showed the effects of specific word knowledge only when reaction times were relatively late (due to task difficulty). If the effects of specific representations in music are similarly slow to emerge, it may explain why researchers using reaction times as a measure tend to find strong schematic effects.

A second difference between the current study, which supports a role for veridical knowledge, and previous studies, which support a role for schematic knowledge, is that in the current work both contexts were reasonably well supported by prior experience. That is, listeners had experienced music with both major- and minor-key harmonic settings and both 3/4 and 6/8 metrical settings. Specific associations that listeners had made with melodies were supported by previous similar experiences, meaning that listeners could be easily swayed in either direction. This was not the case for the "schematic" studies cited above (Hannon & Trehub, 2005a, 2005b; Justus & Bharucha, 2001; Krumhansl & Shepard, 1979; Krumhansl, 1990). In those cases, listeners seemed to take material unsupported by previous experience (distant chord progressions, out-of-key notes, atypical metrical patterns) and fit it to preexisting schemas (or composites of previous experience).

A third possible divergence between the current study and Justus and Bharucha (2001), in particular, is that the melodies used here were much longer and musically richer (several measures with melodic lines and accompaniment figures) than the other authors' contexts (two to four chords). As shown by Bigand, Madurell, Tillmann, and Pineau (1999), longer musical sequences build up stronger contextual expectations. In the current study, the length of the melodies may have built up more specific activation of context, thus allowing more specific information to be accessed than the veridical contexts of Justus and Bharucha.

Listeners' representations of musical context. One question of great interest is how listeners in the current study represented musical context. Two likely possibilities are that they associated melodies with particular tonal centers or metrical patterns and that they formed acoustically-detailed memories of musical elements and then coactivated those elements during listening.

Listeners may have encoded somewhat generic harmonic or metrical information along with the melody. That is, a melody might be encoded as centering around a particular major or minor triad or as sounding fast or slow. Upon hearing the isolated melody, these associations would also be activated. Of course, listeners would need to have sufficient memory to identify particular melodies in order to determine the associated contextual information. Thus, at the very least, listeners seem to form moderately detailed representations that integrate melodic properties with harmonic or metrical information. Such moderate detail might facilitate the extension of specific musical knowledge to stylistically similar melodies.

Another possibility is that listeners are storing veridical (analog) traces of the music that they hear, encoding melody and context as a single event. That whole event is later activated during music listening, which shapes processing. Given that listeners can store detailed traces in their memory (Levitin, 1994; Levitin & Cook, 1996; Palmer et al., 2001; Schellenberg et al., 1999; Schellenberg & Trehub, 2003), it seems plausible that the results in the current study are subserved by the activation of such memories. This account of the data is consonant with exemplar models (e.g., Carpenter & Grossberg, 2003; Goldinger, 1998; Hintzman, 1986) in which memory consists of a collection of experienced traces that are activated to the extent that they match the input. Listeners might be seen as simulating or reinstating context, which is similar to notions of prediction or expectation in music (Bigand et al., 1999; Huron, 2006; Meyer, 1967; Patel et al., 1998). However, prediction of previously heard events is not quite sufficient to explain these results: Listeners did not hear any events during familiarization that were at the temporal locations of the probes. That is, they could not have been veridically predicting one probe event over another, since the probe events never happened during familiarization. It is more likely that they compared the probe to the activated memory of the surrounding musical elements.

What elements might have contributed to the activation of moderately or extremely specific memories? Two candidates are melodic form (the patterns of ups and downs in a melody) and timbral information. Although melodic form was not altered in any of the current experiments, timbre was altered. Experiment 1 presented timbre-matching probes, and Experiment 3 presented timbre-mismatching probes. A *t*-test comparing context-matching trials for the two experiments suggests that timbral match affects listeners' fit judgments: Ratings were higher in Experiment 1, $t1(58) = 6.45$, $p < .0001$; $t2(22) = 6.17$, $p < .0001$; $d = 1.75$. This implies that timbral characteristics influenced listeners' assessments of harmonic fit. This accords with data from Krumhansl and Shepard (1979) that showed that less musically-experienced listeners incorporated nonharmonic characteristics (pitch height) into their harmonic-fit judgments. The perspective of the current paper is that such nonharmonic characteristics may be a natural part of musical memory and music listening.

Implications for music processing. The current study has several interesting implications for understanding music processing. First and foremost, it suggests that highly specific knowledge is a substantial determinant of online processing of even moderately familiar music. When apprehending metrical and tonal information in a melody, listeners activate aspects of previous instances of hearing that melody, or previous instances of hearing similar melodies, which then guides processing. This provides an interesting challenge to models of music perception which try to predict the listener's key or meter judgment based directly on musical input, with sensitivity only to schema-like information rather than diverse previous experiences (Krumhansl, 1990; Large & Jones, 1999; Temperley, 1999). In particular, the current study suggests that richer contextual memory should be incorporated into models.

One way that this might be accomplished is by activating knowledge weighted by the degree of similarity of the present musical input to the contents of memory. This resembles the solution to a similar challenge faced by models of word recognition: How do I match what I am hearing to the set of all heard patterns that I know? The solution in that domain posits graded activation of multiple representations based on similarity (e.g., Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994). The same may be true of music in that listeners are implicitly comparing what they are hearing to previous experience and activating multiple similar instances which shape their perceptions (e.g., Carpenter & Grossberg, 2003; Goldinger, 1998; Hintzman, 1986). It may be argued that these models are more geared toward predicting listener judgments of unfamiliar musical material than dealing with familiar music. The counterargument is that even unfamiliar material may be differentially similar to certain specific instances in memory and that those specifics guide processing more than a general schema does. Nonetheless, more work will be needed to illuminate this matter.

A second implication for music processing is that this work demonstrates a plausible mechanism for learning musical style: Listeners encode large amounts of veridical musical information, which then groups itself by similarity. Similarity-based groupings then shape implicit musical prediction. For instance, listeners might generate different harmonic expectations for a jazz band than for a Baroque orchestra. The current work demonstrates that such differentiated expectations can be learned fairly rapidly. It is of interest to determine how rapidly and implicitly such information might be learned and how far reaching it might be: Does it persist for days rather than minutes, and does it extend to similar but nonidentical musical material? Studies are planned to address these questions.

Conclusion

Music processing, like processing in other cognitive domains, seems to activate contextually-specific memories as well as generalized knowledge. Moreover, the context does not have to be present to affect processing. Three experiments demonstrated that listeners activate specific musical memories during listening. These data suggest that music processing is influenced by the activation of specific memories in proportion to the input's similarity to those memories. This account helps explain how two listeners can hear an identical musical event and have differing impressions of basic structural properties of the music, and it raises the question of whether current models of music perception, which predominantly embody schematic memory, can adequately account for human musical behavior.

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37,* 379–384.

Bharucha, J. J., & Todd, P. M. (1991). Modeling the perception of tonal structure with neural nets. In P. M. Todd & D. G. Loy (Eds.), *Music and Connectionism.* (pp. 128–137). Cambridge: MIT Press.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 184–197.

Carpenter, G. A., & Grossberg, S. (2003). Adaptive resonance theory. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks,* Second Edition, Cambridge, MA: MIT Press, 87–90.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113,* 394–412.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers, 25,* 257–271.

Creel, S. C., & Newport, E. L. (2002). Tonal profiles of artificial scales: Roles of frequency and recency. Paper presented at the *International Conference on Music Perception and Cognition.* Adelaide, Australia: Causal Productions.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117–1121.

Dowling, W. J. (1986). Context effects on melody recognition: Scale-step versus interval representations. *Music Perception, 3,* 281–296.

Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics, 37,* 1–7.

Gill, M. (2004). When information does not deter stereotyping: Prescriptive stereotyping can foster bias under conditions that deter descriptive stereotyping. *Journal of Experimental Social Psychology, 40,* 619–632.

Gjerdingen, R., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research, 37,* 93–100.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66,* 325–31.

Goldinger, S. D. (1998). Echoes of echoes?: An episodic theory of lexical access. *Psychological Review, 105,* 251–279.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 956–974.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16,* 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences (USA), 102,* 12639–12643.

Hare, M., McRae, K., & Elman, J. L. (2003). Memory and language sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language, 48,* 281–303.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93,* 411–428.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation.* Cambridge, Massachusetts: MIT Press.

Jonaitis, E. M., & Saffran, J. R. (2009). Learning harmony: The role of serial statistics. *Cognitive Science, 33,* 951–968.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13,* 313–319.

Justus, T. C., & Bharucha, J. J. (2001). Modularity in musical processing:

The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1000–1011.
Kostka, S., & Payne, D. (1995). *Tonal harmony.* New York: McGraw-Hill, Inc.
Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11,* 346–374.
Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* NY: Oxford U Press.
Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context, *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–594.
Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition, 76,* 13–58.
Lantz, M. E., & Cuddy, L. L. (1998). Total and relative duration as cues to surface structure in music. *Canadian Acoustics, 26,* 56–57.
Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review, 106,* 119–159.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.
Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception and Psychophysics, 56,* 414–423.
Levitin, D. J., & Cook, P. R. (1996). Absolute memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics, 58,* 927–935.
Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology, 39,* 821–831.
Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing, 19,* 1–36.
Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1,* 272–276.
McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.
McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31,* 306–321.
Meyer, L. B. (1967). *Music, the arts, and ideas: Patterns and predictions in twentieth-century culture.* Chicago: U Chicago Press.
Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52,* 189–234.
Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language, 45,* 526–545.
Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 728–741.
Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10,* 717–733.
Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308,* 1430.
Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition, 105,* 533–546.
Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2,* 411–440.
Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84,* 1236–1256.
Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review, 6,* 641–646.
Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14,* 262–266.
Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition, 16,* 285–303.
Serafine, M. L., Davidson, J., Crowder, R. G., & Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language, 25,* 123–135.
Snyder, J., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse-finding. *Music Perception, 18,* 455–489.
Temperley, D. (1999). What's key for key? The Krumhansl-Schmuckler key-finding algorithm reconsidered. *Music Perception, 17,* 65–100.
Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology Learning, Memory, and Cognition, 19,* 528–53.
Wong, P. C. M., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception, 27,* 81–88.

Received July 7, 2010
Revision received January 25, 2011
Accepted January 30, 2011 ■