

Tonal hierarchy representations in auditory imagery

Dominique T. Vuvan · Mark A. Schmuckler

Published online: 9 November 2010
© Psychonomic Society, Inc. 2010

Abstract Two experiments investigated psychological representations of musical tonality in auditory imagery. In Experiment 1, musically trained participants heard a single tone as a perceptual cue and built an auditory image of a specified major tonality based on that cue; participants' images were then assessed using judgments of probe tones. In Experiment 2 participants imaged a minor tonality rather than a major one. Analysis of the probe tone ratings indicated that participants successfully imaged both major and minor tonal hierarchies, demonstrating that auditory imagery functions comparably to auditory perception. In addition, the strength of the major tonal image was dependent upon the pitch and tonal relations of the perceptual cue and the to-be-imagined tonality. Finally, representations of minor tonal hierarchies were less robust than those of major ones, converging with perceptual evidence that minor tonalities are less psychologically stable than major tonalities.

Keywords Imagery · Music cognition · Auditory perception

Introduction

Mental imagery enjoys a longstanding history as a valuable tool to study human information processing. Because imagery draws upon memory processes and information manipulation, studying such images yields knowledge about the properties of the representations one employs. As humans are predominantly visual creatures, it is no surprise that research from vision dominates mental

imagery work (e.g., Finke, 1980, 1985; Finke, 1989; Kosslyn, 1980; Kosslyn, Thompson & Ganis, 2006).

Behavioral studies of visual imagery have investigated its function in learning, memory, reasoning, and spatial judgments (Canellopoulou & Richardson, 1998; Cooper & Shepard, 1973; Currie, 1995; Dirx & Craik, 1992; Finke & Shepard, 1986; Gordon & Hayward, 1973; Hartley, 1977, 1981; Kerst & Howard, 1978; Knauff & Johnson-Laird, 2002; Paivio, 1965, 1975; Paivio & Foth, 1970; Shepard & Chipman, 1970; Shepard, Kilpatrick & Cunningham, 1975), as well as identified the elements comprising mental images, how these elements can be manipulated, and whether they veridically represent real world perceptions. These studies have revealed that mental images contain accurate depictions of basic properties (Finke & Schmidt, 1977; Kosslyn, 1973; Kosslyn, Ball & Reiser, 1978; Pinker, 1980; Shepard & Metzler, 1971; Watkins & Schiano, 1982), and, importantly, that they can be manipulated by transformations such as rotation (Cooper, 1975; Cooper & Shepard, 1973, 1975; Shepard & Metzler, 1971) and scanning (Denis, Goncalves & Memmi, 1995; Denis & Kosslyn, 1999; Finke & Pinker, 1982; Kosslyn, 1973; Kosslyn et al., 1978; Pinker, 1980). For instance, Kosslyn (1973) showed participants drawings and then later asked them to verify pictorial features of those drawings from memory. The time it took for participants to verify a feature was linearly related to the spatial distance of that feature from an initial focus point, which demonstrated that participants were mentally scanning between the focus point and the feature of interest, just as they would in a perceptual context.

This focus on vision ignores the fact that humans are also linguistic and musical creatures, and that auditory imagery is an important part of mental life (Reisberg, 1992). Whereas research in the visual domain has generally

D. T. Vuvan (✉) · M. A. Schmuckler
University of Toronto Scarborough,
Toronto, Canada
e-mail: dominique.vuvan@utoronto.ca

taken for granted participants' abilities to imagine basic visual structure (e.g., color and shape), auditory imagery research has focused on examining the veridicality of basic auditory dimensions in auditory images. These dimensions include pitch (Crowder, 1989; Halpern, 1989; Schellenberg & Trehub, 2003), timbre (Crowder, 1989), loudness (Intons-Peterson, 1980, 1992), and time (Halpern, 1988), and studies have confirmed participants' abilities to accurately image these auditory perceptual attributes.

One consequence of this focus in auditory imagery is that some fundamental aspects of psychological processes investigated in visual imagery have received scant attention within the auditory domain. One such issue involves what many identify as an important degree of isomorphism between perceptual processing and imagery processing (Finke, 1980, 1985, 1989; Kosslyn, 1980; Shepard, 1978, 1982a). Researchers over the past decades have suggested that these two domains are highly similar, constrained by comparable factors, and even make use of the same neurological substrates (see Kosslyn, 1994, for a review).

Interestingly, Hubbard (2010), in a review of auditory imagery, highlights investigating the relation between auditory imagery and auditory perception as one (of many) as yet unanswered questions in this field. In this regard, specific questions might involve whether imagery interferes with or facilitates perception, whether imagery and perceptual processes are comparably instantiated neurologically, and so on. Thus, the relation between imagery and perceptual processes represents an important domain of inquiry.

One way to pursue this issue involves investigating whether auditory images capture complex structural organizations (as opposed to basic auditory dimensions) that are inherent in auditory perception. In this regard, musical structure represents a compelling candidate for study in that there have been extensive efforts devoted to investigating listeners' apprehension of complex structural organizations within musical contexts. Within the music cognition literature, probably the best place to start is the organization of the most salient musical dimension—pitch (Prince, Schmuckler, & Thompson, 2009; Prince, Thompson, & Schmuckler, 2009; Schmuckler, 2004, 2009). Within the pitch domain, the most thoroughly studied of such organizations involves “musical tonality” (Krumhansl, 1990, 2000; Schmuckler, 2004, 2009; Schmuckler & Tomovski, 2005). Simply described, tonality refers to the organization of the 12 chromatic pitches (the complete set of Western musical tones) around a reference pitch, such that the other tones are heard in relation to this focal pitch. These tones then form a hierarchy of importance, with some tones fitting well and other tones fitting poorly with this central pitch.

In Western music, this hierarchy has four levels (Schmuckler, 2004). The top level contains the tonal centre

or *tonic* and is the point of greatest stability within the key. The rest of the tones in the hierarchy are defined by *scale degrees*, which measure their position in semitones (the smallest unit of pitch distance employed in Western music) above the tonal centre. Thus, if the tonic is scale degree 0, the second hierarchical level contains scale degrees 4 and 7, and the third level contains scale degrees 2, 5, 9, and 11. Together, these first three levels comprise the scale tones in the key of interest. Finally, the lowest level of the hierarchy (notes outside the scale) contains the remaining five tones (scale degrees 1, 3, 6, 8, and 10), which are the least important tones of the chromatic set. This tonal hierarchy is shown in Table 1.

Musical tonality contains two additional important properties. First, there are two categories of tonality in Western music: the major tonality, which follows the hierarchical structure just described, and the minor tonality, which embodies a modified version of the hierarchical positions of the tones (Table 1). There are actually three forms of the minor—natural, harmonic, and melodic—each with slightly different tonal hierarchies. Past research has largely ignored these distinctions, typically employing either the natural or harmonic minor; Table 1 presents the harmonic minor form, which was used in these experiments. The second property of tonality is that this pitch hierarchy can be built with any of the chromatic tones as the tonic. Combining the 12 chromatic tones with the two different forms produces 24 tonal organizations within Western music.

Over the years research has investigated the psychological reality of this theoretical hierarchy of stability. The most thorough of such work is by Krumhansl and colleagues (see Krumhansl, 1990; 2000 for reviews). In a now classic study, Krumhansl and Shepard (1979) presented listeners with tonality defining musical contexts, followed by a probe tone drawn from the 12 chromatic tones, and asked listeners to judge how the probe fit with

Table 1 Summary of the tonal hierarchies for A major and A minor

Tonal hierarchy level	Semitone interval from tonic	Pitch class
Major		
	A major	
Tonic	0	A
Tonic triad members	4 7	C# E
Diatonic tones	2 5 9 11	B D F# G#
Non-diatonic tones	1 3 6 8 10	A# C D# F G
A minor		
Tonic	0	A
Tonic triad members	3 7	C E
Diatonic tones	2 5 8 11	B D F G#
Non-diatonic tones	1 4 6 9 10	A# C# D# F# G

the tonal contexts. Listeners' ratings for these probes matched the hierarchy just described, with the tonic receiving the highest rating, followed by tones at the second and third levels, and finally the tones at the bottom of the hierarchy. Krumhansl and Kessler (1982) extended these findings to minor tonalities (the first property described earlier) and generalized across different tonal centers (the second property). The average ratings for the 12 probe tones following a major or minor context appear in Fig. 1, and are called the *tonal hierarchy*.

Tonality has been observed to have far-reaching effects on musical processing, with studies demonstrating tonal influences on perception (Bigand, Parncutt & Lerdahl, 1996; Krumhansl & Schmuckler, 1986; Marmel, Tillmann & Dowling, 2008; Warrier & Zatorre, 2002), memory (Krumhansl, 1979; Krumhansl & Castellano, 1983; Leman, 2000), and musical expectation (Bharucha, 1994; Bharucha & Stoeckig, 1986; Huron, 2006; Schmuckler, 1989, 1990). Tonality has also been the focus of work in auditory imagery. Hubbard and Stoeckig (1988), for instance, found

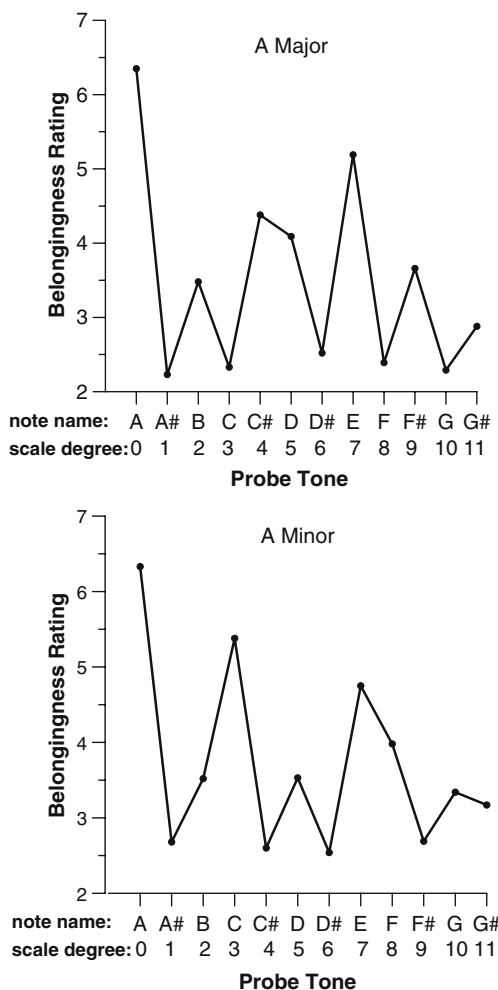


Fig. 1 (a) A major and (b) A minor tonal hierarchies (adapted from Krumhansl & Kessler, 1982)

that participants demonstrated priming results based on imagined musical stimuli that were comparable to the priming observed with actual sounded musical contexts, with the pattern of priming predictable from the tonal relatedness of the stimuli. Similarly, Janata and Paroo (2006) found that imagining a prototypical tonal context influenced participants' detection of pitch mistuning in target notes, with accuracy related to the hierarchical stability of the to-be-judged tone with reference to the imagined tonal context.

These previously findings, then, suggest that imagined tonal contexts have comparable effects on auditory and/or musical processing as do sounded contexts. These results are limited, however, in that they do not explore the full structural organization of these images. In this case, the question is whether tonal images induce a comparable hierarchical organization on the perception of tones as observed with actual sounded tonal contexts. An affirmative answer to this question would provide strong evidence for the structural equivalence of musical imagery and perception, and would bring our understanding of auditory imagery into line with our understanding of visual imagery. Investigating these issues was the goal of this project.

Experiment 1: auditory imagery of major tonalities

Experiment 1 investigated the degree to which an imagined tonal context induces similar processing of tones as observed with sounded tonal contexts. This goal was accomplished by providing listeners with a cue (a single tone) to a given tonality and then instructing them to image this tonality. Success in imagery generation was determined by assessing the degree to which hierarchical structure could be observed in this image.

This study also addressed a series of secondary questions, including the flexibility of imagery generation and the role of additional factors on imagery generation. Flexibility was explored in two ways. First, flexibility was examined by providing listeners with different cues for imagery generation, with these cues varying in their presumed effectiveness in producing a tonal image, ranging from highly effective to very ineffective. Accordingly, this possibility addresses the ease of imagery generation as a function of cue validity.

If imagery formation does vary with the validity of the cue, then it becomes an interesting question to determine what additional factors drive this variation. Consideration of this issue reveals an array of possibilities. First, it may be that imagery variation is based on basic perceptual characteristics of the cue such as its actual physical pitch (in frequency), either on its own or in relation to the imaged tonality. The former suggests that imagery is driven solely

by the pitch of the cue tone itself, whereas the latter implies that the physical difference (in frequency) between the pitch of the cue and the generated tonality would be critical. Intriguingly, this latter situation provides an auditory analogue for the visuospatial mental scanning effects described earlier (e.g. Kosslyn, 1973).

Second, the effectiveness of the cue for imagery generation may vary in accordance with the musical relation between the cue and imaged tonality. Once again, there are multiple ways in which such a possibility might be evident. For instance, rather than acting as a signal to an experimentally-produced tonality, this cue might instead give rise to a tonal percept based on the cue itself. Or alternatively, it might be that the more musically related the cue is to the imagined tonality, the more effective it is as a cue. In both cases such an effect would indicate a role for long-term, internalized tonal representations in auditory imagery. Interestingly, such effects would be analogous to previous findings on tonal relatedness called “key distance effects,” which have been observed in the music cognition literature across a range of perceptual and memory tasks (Bartlett & Dowling, 1980; Bharucha & Krumhansl, 1983; Krumhansl, Bharucha, & Castellano, 1982; Krumhansl & Castellano, 1983). Accordingly, positive results along these lines would also tie together perceptual and imagery processing.

A second way in which flexibility was examined involved assessing listeners’ abilities to shift the tonal image they produced. Along these lines, one could ask listeners to either maintain a consistent tonal image across the experimental session, or to modify their tonal image across the session such that they start by imaging one tonality, then change to imaging a different tonality, and so on. Generally, such a manipulation investigates the ease with which listeners can shift their imagined tonal representation, a question that is analogous to the perception of what is called “key modulation” in music. Previous work in music cognition (Schmuckler & Tomovski, 2005; Smith & Cuddy, 2003; Toiviainen & Krumhansl, 2003) has found that listeners can rapidly shift perceived tonal orientations, moving effortlessly between different tonalities. Accordingly, manipulating the consistency of the to-be-judged tonal image represents yet another point of contact in assessing the relation between perceptual and imagery processing.

Methods

Participants

Twenty-four individuals ($M_{age} = 21.2$ years, $SE = 0.5$ years) from the University of Toronto Scarborough community

participated in this study. Because this study required knowledge of Western musical structure all participants were required to have a minimum of five years of formal music instruction ($M = 9.5$ years, $SE = 0.8$ years). Participants also had an average of 3.9 years of musical theory ($SE = 0.8$ years), listened to music for 16.9 h/week ($SE = 4.8$ h/week), and played music for 7.7 h/week ($SE = 1.6$ h/week). One participant had participated in a music psychology experiment before, and two of the participants reported having absolute pitch.

Stimuli

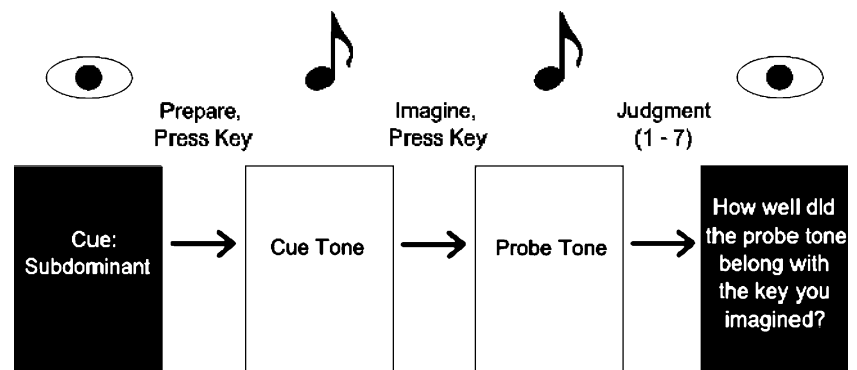
The stimuli consisted of one second long piano tones ranging from A#₃ (233.08 Hz) to G#₅ (830.61 Hz) in semitone steps, synthesized using Finale 2005.r2 (Author, 2004). These stimuli were presented to participants using an Intel Pentium 4 PC running MATLAB 7.0 (Moler, 2004). The visual stimuli appeared on an LG Flatron L1710S monitor, while the auditory stimuli were heard through Sennheiser HD 280 pro headphones plugged into a Creative Sound Blaster Audigy 2 ZS soundcard.

Experimental task

All listeners participated in a modified version of the probe tone task (Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979), which appears schematically in Fig. 2. In this task, listeners heard a cue tone, which was used to generate the to-be-imagined tonality (henceforth called the *imagined tonality*), followed by a to-be-rated probe tone. The cue tones varied in their *cue function* and *cue type*. Cue Function refers to the fact that the cue tone could correspond to any of the seven scale tones (levels 1 to 3; see Table 1). Because these tones lie at different levels of the hierarchy, they differ in their relatedness to a given tonality, and thus represent cues of varying strength when instantiating a given tonality. As such, cue function operationalized the first means by which flexibility in imagery variation was tested. Cue Function was a within-subjects variable.

Cue Type was comprised of two possible values; these values, with respect to cue function, appear in Table 2. In the *cue varied* (CV) condition, the pitch (in Hz) of the cue changed in parallel with changes in cue function, taking on values consistent with the A major scale. For instance, when the cue tone was A₄ it was always designated as scale degree 0, and when the cue tone was B₄ it was designated as scale degree 2. Accordingly, although the pitch of the cue tone fluctuated, the imaged tonality was always A major. In the *cue constant* (CC) condition, the pitch of the cue was always A₄ (440 Hz), with the tonal function of this tone varying in accordance with the seven major scale

Fig. 2 Schematic diagram of the imagery probe tone procedure



tones. Thus, when the cue tone (A_4) was labeled as scale degree 0, the imaged tonality was A major. When this same cue was labeled as scale degree 2, however, the imaged tonality was G major. Cue Type, then, operationalized the second means of assessing the flexibility of tonal image generation, with the CC condition requiring listeners to shift their tonal image with each different cue, whereas the CV condition required a constant tonal image across cues (see Table 2). Cue type was a between-subjects variable, with listeners randomly assigned to either CC or CV conditions.

Finally, the probe tones could be any of the 12 tones of the chromatic set, with the range of the set determined by the imaged tonality (see Table 2). In the CV block, because the imaged tonality was always A major, the probe tones ranged from A_4 to $G\#_5$. In the CC block, the probe tones varied in accordance with the imaged tonality. Thus, when the cue tone (A_4) was labeled as scale degree 0 (imaged tonality = A major), the probe tones ranged from A_4 to $G\#_5$. When the cue tone was labeled as scale degree 2 (imaged tonality = G major) the probes ranged from G_4 to $F\#_5$. Probe tone was a within subjects variable, and enabled the primary goal of this study—to determine whether imaged tonalities produce a comparable hierarchical organization of the chromatic set as found with sounded tonalities.

All listeners received seven randomized blocks of experimental trials corresponding to the seven cue functions. Within each block listeners heard a random ordering of the 12 probe tones. All participants received two repetitions of each cue function block. Overall, listeners received 168 trials in all (7 cue functions \times 12 probe tones \times 2 repetitions).

Procedure

Initially, participants were visually presented with the cue function for the upcoming cue tone (e.g., scale degree 5),¹ and then pushed a key to hear the cue tone. Participants

¹ In this experiment the cue tone was actually labeled using musical terminology (e.g., scale degree 5 is the “subdominant”). For simplicity of explication, and to reduce the amount of musical terminology, we have chosen not to employ this terminology throughout this paper. The musical name does, however, appear in Fig. 1.

were instructed to form an image of a musical tonality based on this cue and indicated cue function. For instance, if the cue function was “scale degree 5” listeners were to consider the cue as the fifth scale degree of a major tonality. Listeners were told that the image could be produced via any means they chose, short of singing aloud to themselves. After a self-determined interval of key imagining, participants pressed a key to hear the probe tone, and then rated how well the probe fit with the imagined tonality using a 1 (fits very poorly) to 7 (fits very well) scale.

Prior to beginning the experimental trials, participants received sample trials, and the experimenter remained in the room during the first five experimental trials. Upon finishing the experiment participants completed a music background questionnaire. The entire experiment lasted approximately one hour.

Results and discussion

To assess the stability of probe tone ratings across participants, intersubject correlations (ICs) were calculated separately for the CC and CV groups by averaging, for each listener, across the two repetitions, and then amalgamating across the 7 cue functions and 12 probe tones. Both groups produced significant ICs, with an average $r(82) = 0.29$ ($SE = 0.02$), $p < 0.01$, and $r(82) = 0.45$ ($SE = 0.02$), $p < 0.01$, respectively. To assess differences in ratings stability between groups, ICs were averaged for each group and then compared using Fisher’s z' transformation; the ICs for the CV group were not significantly different from those for the CC group, $z = 1.17$, ns .

Because two participants in this study reported having absolute pitch (AP), it was important to determine whether the AP participants performed the experiment differently from the rest of the participants, based on the idea that AP possessors process pitch differently from non-AP possessors (Levitin & Rogers, 2005). Accordingly, the averaged ICs for the AP participants (both of whom were in the CC condition) were compared to the averaged ICs for the remaining non-AP participants (from the CC condition).

Table 2 Levels of the Cue Function, Cue Type, and Probe Tone factors, and Pitch and Key Distance measures for Experiment 1

Cue Function	Cue Type (Imaged Tonality)	Probe Tones	Pitch Distance		Key Distance	
Scale degree	Cue varied	Cue constant	Cue varied	Cue constant	Ascending	Overlap
0	A ₄ (A major)	A ₄ (A major)	A ₄ - G# ₅	A ₄ - G# ₅	0	7
2	B ₄ (A major)	A ₄ (G major)	A ₄ - G# ₅	G ₄ - F# ₅	2	5
4	C# ₅ (A major)	A ₄ (F major)	A ₄ - G# ₅	F ₄ - E ₅	4	3
5	D ₅ (A major)	A ₄ (E major)	A ₄ - G# ₅	E ₄ - D# ₅	5	6
7	E ₅ (A major)	A ₄ (D major)	A ₄ - G# ₅	D ₄ - C# ₅	7	6
9	F# ₅ (A major)	A ₄ (C major)	A ₄ - G# ₅	C ₄ - B ₄	9	3
11	G# ₅ (A major)	A ₄ (Bb major)	A ₄ - G# ₅	Bb ₃ - A ₄	11	2

This analysis failed to reveal any significant differences between these groups (mean IC_{non-AP} for the CC group = 0.28; mean IC_{AP1} = 0.34, $z = 0.42$, *ns*; mean IC_{AP2} = 0.42, $z = 1.01$, *ns*), indicating that the AP participants performed the task comparably to the non-AP possessors.

To facilitate subsequent analyses, all probe tone ratings were transposed to a common imaged tonality of A major; this procedure affected ratings for the CC block but not for the CV block. Participants' ratings were analyzed using a four-way analysis of variance (ANOVA), with the within-subjects factors of Cue Function (7 levels – see Table 2), Probe Tone (12 levels – see Table 2), and Repetition (1 versus 2), and the between-subjects factor of Cue Type (CV versus CC). Of the four main effects, the only significant finding occurred for Probe Tone, $F(11, 242) = 33.18$, $MSE = 281.24$, $p < 0.001$, $\eta_p^2 = 0.60$. This effect indicates that different probe tones were rated differentially by participants. Of all the interactions produced by this design, the only significant effect occurred between Cue Function and Probe Tone, $F(66, 1452) = 7.80$, $MSE = 22.52$, $p < 0.001$, $\eta_p^2 = 0.26$, suggesting that cue functions moderated changes in probe tone ratings. Interestingly, the fact that no significant effects involving Cue Type were observed is evidence for flexibility in image generation, with listeners able to effectively shift their tonal representations throughout the experiment.

Although this analysis demonstrates that the different probe tones elicited different ratings, it does not address whether the imaged tonality produced a comparable tonal hierarchy to a perceptual context. To answer this question, ratings were averaged across listeners and cue function, repetition, and cue type to produce a single set of ratings for the 12 probe tones. These averaged imagery ratings were then correlated with the 12 major and 12 minor tonal hierarchies (Krumhansl & Kessler, 1982). These ratings showed a better correspondence with the tonal hierarchy for the imaged key of A major, $r(10) = 0.88$, $p < 0.001$, than with any other tonal hierarchy. Figure 3 shows these

averaged ratings in relation to the A major tonal hierarchy. Probe tone ratings were not strongly correlated with A minor, $r(10) = 0.23$, *ns*.

Subsequent analyses focused on understanding the aforementioned Cue Function x Probe Tone interaction. The goal was to examine how the various cue functions differentially influenced listeners' abilities to image a given tonality, along with assessing possible influences on tonal imagery; Table 3 provides an example of three potential factors when the cue function was scale degree 2 (B₄). These factors align with the various possibilities described earlier, and are based on characteristics of the cue tone itself. The first factor is the *imaged tonality*, and represents the tonality that should have been produced if listeners were accurately generating the requested tonality. Accordingly, this factor consisted of the tonal hierarchy for A major (Krumhansl & Kessler, 1982). The second factor was a *cue pitch* factor, and assessed whether listeners simply gave a higher rating to the probe tone having the same pitch as the cue. This factor consisted of a dummy profile with the pitch of the cue tone receiving a value of 1 and the remaining

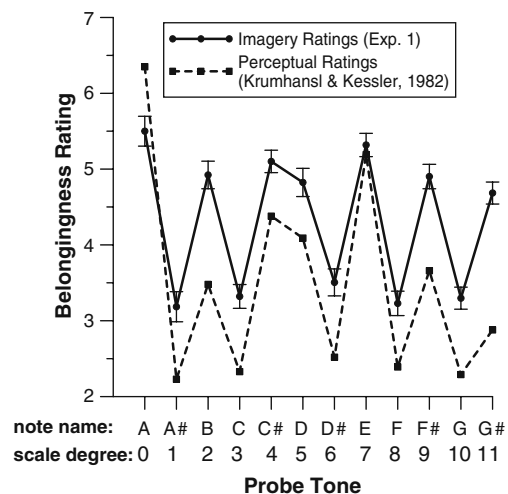


Fig. 3 Mean ratings for the probe tones of Experiment 1, along with the A major tonal hierarchy

Table 3 Imaged Tonality, Cue Pitch, and Cue Tonality Predictors for Scale Degree 2 (Cue Tone B) in Experiments 1 and 2

Scale Degree	Note Name	Experiment 1 (major)			Experiment 2 (minor)		
		Imaged Tonality	Cue Pitch	Cue Tonality	Imaged Tonality	Cue Pitch	Cue Tonality
0	A	6.35	0	2.29	6.33	0	3.34
1	A#	2.23	0	2.88	2.68	0	3.17
2	B	3.48	1	6.35	3.52	1	6.33
3	C	2.33	0	2.23	5.38	0	2.68
4	C#	4.38	0	3.48	2.60	0	3.52
5	D	4.09	0	2.33	3.53	0	5.38
6	D#	2.52	0	4.38	2.54	0	2.60
7	E	5.19	0	4.09	4.75	0	3.53
8	F	2.39	0	2.52	3.98	0	2.54
9	F#	3.66	0	5.19	2.69	0	4.75
10	G	2.29	0	2.39	3.34	0	3.98
11	G#	2.88	0	3.66	3.17	0	2.69

Imaged Tonality is based on the Krumhansl and Kessler (1982) values for A major (Experiment 1) or minor (Experiment 2). Cue Pitch is a dummy variable describing the expected probe rating if listeners are simply rating the cue pitch higher than all other pitches. Cue Tonality is based on the Krumhansl and Kessler (1982) values for the major (Experiment 1) or minor (Experiment 2) key based on the cue (B major/minor in this example)

tones receiving values of 0. The third factor was a *cue tonality* factor, and evaluated whether the cue tone itself was perceived as the tonic of a major tonality. For instance, if the cue tone was B₄, listeners might have generated a tonal hierarchy based on B major. This factor consisted of the major tonal hierarchy having the cue tone as its tonic (Krumhansl & Kessler, 1982).

To assess the impact of these factors, 12-element probe tone profiles for each cue function were created for each participant, and were then correlated with each of the three predictors. This procedure produced 21 correlations for each listener. These correlations were then normalized using Fisher’s *z'* and entered into a repeated-measures ANOVA with Predictor and Cue Function as within-subjects factors. This analysis revealed no main effect for Predictor, $F(2, 46) = 1.93, ns$. In contrast, there was a significant main effect for Cue Function, $F(6, 138) = 20.00, MSE = 1.68, p < 0.001, \eta_p^2 = 0.47$, demonstrating that different cue functions generally produced different correlations with these predictors. Most important was the significant interaction between Predictor and Cue Function, $F(12, 276) = 6.03, MSE = 0.44, p < 0.001, \eta_p^2 = 0.21$, which illustrates that the relative strengths of each of the three predictors varied depending upon the tonal function of the cue. Figure 4 depicts the nonstandardized correlations, averaged over participant, as a function of these two factors.

This interaction was further probed using simple effects analyses to assess the effect of Predictor at each level of Cue Function. The effect of Predictor was significant for scale degree 0 [A], $F(1, 23) = 34.71, MSE = 0.89, p < 0.001, \eta_p^2 = 0.60$, indicating that Imaged and Cue Tonality

(which are identical for this cue function) were better predictors than Cue Pitch. The effect of Predictor was also significant for scale degree 7 [E], $F(1, 23) = 5.06, MSE = 0.21, p = 0.01, \eta_p^2 = 0.18$, with Imaged Tonality a better predictor than Cue Pitch, $t(23) = 3.25, p < 0.01$ and marginally better than Cue Tonality, $t(23) = 2.03, p = 0.055$. Cue Tonality was significantly better than Cue Pitch, $t(23) = 3.04, p < 0.01$. None of the other cue functions were significant, all $F_s < 2.50, ns$. One reason this pattern of results is noteworthy is that the two scale degrees (0 [A] and 7 [E]) that produced the most differentiation between the predictors are scale degrees lying at the top of the tonal hierarchy (see Table 1). In contrast, those cues lying at lower

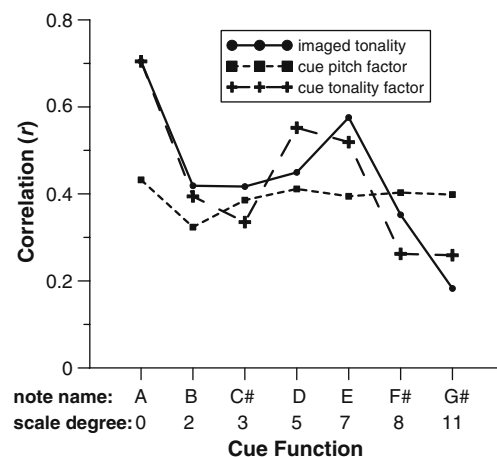


Fig. 4 Averaged correlations between the probe tone ratings and imaged tonality, cue pitch factor, and cue tonality factors as a function of cue function for Experiment 1

levels of the tonal hierarchy (e.g., scale degrees 9 [F#] and 11 [G#]) produced non-significant correlations between the imaged tonality and the probe tone ratings. Thus, when the cue was psychologically stable and hierarchically important, listeners were better able to produce the specified musical image; when the cue was psychologically unstable and hierarchically unimportant, imagery suffered. In other words, good cues make good images in a musical context.

Subsequent analyses addressed additional explanations for why these cue functions may have varied in their imagery induction, again based on influences already identified. For instance, one previously discussed possibility was based on *pitch distance*, with the greater the pitch spread between the cue and the tonic of the imaged tonality the less accurate the image of that tonality. For instance, if the cue was scale degree 2 [B] (two semitones above the tonic) listeners should produce a more accurate tonal image than if the cue tone was scale degree 4 [C#] (four semitones above the tonic). Additionally, there is the possibility that the musical relatedness between the cue tonality and the imaged tonality played a role in auditory imagery, consistent with previously identified ideas of *key distance*. Accordingly, one might predict that the closer the tonal relation between the imaged and cue tonality the more successfully participants would produce the imaged tonality.

To assess these possibilities, pitch distance and key distance measures were created (Table 2). Pitch distance was quantified in terms of the number of ascending semitones separating the tonic of the imaged tonality and the cue tone. Ascending semitone distance is important because the cue tone was always heard at or above the pitch of the tonic of the imaged tonality.² Key distance was quantified by counting the number of shared tones in the major scales of any two tonalities. For instance, the major scale built on scale degree 0 [A] consists of scale degrees <0, 2, 4, 5, 7, 9, 11>, or the notes <A B C# D E F# G#>, whereas the major scale built on scale degree 7 [E] consists of scale degrees <0, 2, 4, 6, 7, 9, 11>, or the notes <A B C# D# E F# G#>; these two scales share six of the seven tones. In contrast, the major scale built on scale degree 4 [C#] consists of scale degrees <1, 3, 4, 6, 8, 9, 11>, or the notes <A# C C# D# F F# G#>; thus, these scales share three tones. Accordingly, the tonality based on scale degree 7 [E] is more related in key distance to the imaged tonality than is the tonality based on the scale degree 4 [C#].

² One can also quantify pitch distance using the minimum number of semitones between the imaged tonality tonic and the cue tone. Minimum pitch distance is of interest given the circular nature of musical tones (Revesz, 1954; Shepard, 1964, 1982a, 1982b) and based on the notion of octave equivalence (Deutsch, 1972, 1979; Sergeant, 1983). Because analyses of minimum pitch distance in this project did not produce any significant results, they will not be discussed.

To predict imagery performance across cue function, the correlations between averaged probe tone ratings and imaged tonality were themselves correlated with pitch distance and key distance across the seven cue functions. This analysis revealed that imagery performance was significantly related to ascending pitch distance, $r(5) = -0.76$, $p < 0.05$, and key distance, $r(5) = 0.90$, $p < 0.01$. Thus, it was easier to generate a tonal image when the cue tone was close in pitch to the tonic of the imaged tonality and when the cue tonality was more musically related to the imaged tonality. A multiple regression predicting tonal imagery correlations from the pitch distance and key distance measures revealed a significant relation, $R = 0.93$, $p < 0.05$, with key distance contributing significantly, $\beta = 0.71$, $p < 0.05$, but not pitch distance, $\beta = -0.30$, *ns*.

In sum, three critical findings emerge from Experiment 1. First, and most importantly, listeners successfully produced an image of a major tonality, with this tonal image influencing ratings of subsequently sounded musical tones. This result provides an answer to the fundamental question underlying this project, namely, that imagery processing is strongly analogous to perceptual processing. Second, listeners demonstrated flexibility in imagery generation, showing little difference in producing a tonal image that remained constant across the experimental session versus a tonal image that varied across the experiment. This result is also consistent with findings in perceptual processing, arising from studies on key modulation (Schmuckler & Tomovski, 2005; Smith & Cuddy, 2003; Toiviainen & Krumhansl, 2003). Third, the accuracy of the tonal images were multiply determined, based on both pitch and key distance factors. This final result also confirms the relation between perceptual and imagery processes, in that key distance effects have been observed in perceptual tasks involving both rating scales and memory for musical materials (e.g., Bartlett & Dowling, 1980; Bharucha & Krumhansl, 1983; Krumhansl, Bharucha, & Castellano, 1982; Krumhansl & Castellano, 1983; Krumhansl & Kessler, 1982). Thus, perceptual and imagery processes do appear to be tapping similar mechanisms in musical processing.

It is important to note, however, that Experiment 1 tested the strongest possible situation for auditory imagery – the production of a major tonality. Previous research in music cognition distinguishes the major tonality in that it is a highly psychologically stable musical structure (Krumhansl, 1990; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979). Accordingly, it is of interest to examine the production of an auditory image of a less stable musical organization.

Experiment 2: auditory imagery of minor tonalities

Experiment 2 investigated listeners' abilities to generate an auditory image that is more difficult than the image used in

Experiment 1—a minor tonality. Previous work suggests that the perception of minor keys is less psychologically stable than the perception of major keys (Delzell, Rohwer & Ballard, 1999; Harris, 1985; Krumhansl, Bharucha, & Kessler, 1982). Accordingly, investigation of imagery for minor keys provides an important test of the idea that auditory perception and imagery are comparable by assessing whether participants’ abilities to produce such images will suffer from the same instability as their perceptual counterparts.

Methods

Participants

Fifteen musically trained participants ($M_{age} = 18.8$ years, $SE = 0.3$ years) were recruited from the University of Toronto Scarborough community. All participants had a minimum of five years of formal music instruction, with a mean training of 8.9 years ($SE = 0.6$ years), had 3.4 years of musical theory training ($SE = 1.0$ years), listened to music for 15.6 h/week ($SE = 3.1$ h/week), and played music for 5.7 h/week ($SE = 1.3$ h/week). One participant had previously participated in a music psychology experiment and two participants reported having absolute pitch.

Stimuli, experimental task, and procedure

The hardware, software, and sound stimuli were identical to Experiment 1 and the experimental task was almost identical. Again, listeners were asked to image a tonality based on a cue tone, and to rate probe tones relative to this imaged tonality. The cue tone varied corresponding to the minor scale degrees, although given the results of Experiment 1 trials were restricted to the cue varied (CV) condition, with a constant imaged key of A minor. Listeners once again heard all 12 chromatic tones as probes, with this set matched to the range of the imaged tonality (see Table 4).

All participants received seven randomized blocks of trials, corresponding to each cue function. Each block consisted of the full set of randomized probe tones. All participants heard two repetitions of each cue function block, producing 168 trials in all (7 cue functions x 12 probe tones x 2 repetitions).

The procedure for Experiment 2 was the same as Experiment 1, except that participants were instructed to imagine a harmonic minor rather than a major tonality. The harmonic minor was chosen due to its general familiarity to listeners, and because it has been most commonly employed by researchers. After completing the experimental trials, listeners filled out a music background questionnaire. The experiment lasted approximately one hour.

Results and discussion

Intersubject correlations were calculated as in Experiment 1. As a group, participants produced low ICs, with an average $r(82) = 0.11$ ($SE = 0.03$), *ns*. To compare Experiment 1 and 2, ICs were transformed using Fisher’s z' . Experiment 2 ICs were significantly less than ICs for the CV trials of Experiment 1, $z = 2.35$, $p < 0.05$, indicating far less consistency across participants when imagining minor keys. Once again, because two listeners reported absolute pitch, z' -transformed ICs for the absolute pitch participants were compared to those for all other participants. As before, the mean intersubject correlations for the AP possessors were not significantly different from those for the other participants (mean $IC_{non-AP} = 0.10$; mean $IC_{AP1} = 0.18$, $z = 0.51$, *ns*; mean $IC_{AP2} = 0.15$, $z = 0.32$, *ns*).

Probe tone ratings were analyzed using a 3-way ANOVA, with Cue Function, Probe Tone, and Repetition as within-subjects factors. This analysis produced one significant main effect, that of Probe Tone, $F(11, 154) = 3.79$, $MSE = 20.0$, $p < 0.001$, $\eta_p^2 = 0.21$. The only significant interaction was between Cue Function and Probe Tone, $F(66, 924) = 3.81$, $MSE = 11.57$, $p < 0.001$, $\eta_p^2 = 0.17$, suggesting that varying cue functions produced different probe tone ratings.

Table 4 Levels of the Cue Function, Cue Type, and Probe Tone factors, and Pitch and Key Distance measures for Experiment 2

Cue Function (Scale Degree)	Cue Tone (Imaged Tonality)	Probe Tone	Pitch Distance	Key Distance
Scale Degree	—	—	Ascending	Overlap
0	A ₄ (A minor harmonic)	A ₄ - G# ₅	0	7
2	B ₄ (A minor harmonic)	A ₄ - G# ₅	2	3
3	C ₅ (A minor harmonic)	A ₄ - G# ₅	3	4
5	D ₅ (A minor harmonic)	A ₄ - G# ₅	5	4
7	E ₅ (A minor harmonic)	A ₄ - G# ₅	7	4
8	F ₅ (A minor harmonic)	A ₄ - G# ₅	8	3
11	G# ₅ (A minor harmonic)	A ₄ - G# ₅	11	3

Subsequent analyses explored whether the imaged minor tonality produced a comparable organization on the probe tones as a sounded tonal context. Accordingly, data were averaged across all participants and factors to produce a single set of 12 probe tone ratings. These averaged ratings, shown in Fig. 5, were then correlated with the 12 major and 12 minor tonal hierarchies. The ratings matched the tonal hierarchy of the imaged tonality of A minor better than any other tonal hierarchy, $r(10) = 0.79$, $p < 0.001$, demonstrating global success at generating a minor tonal image. This result is qualified, however, by the strong relation with the A major hierarchy, $r(10) = 0.66$, $p < 0.05$. Despite the fact that a subsequent multiple regression analysis revealed that the A minor hierarchy significantly predicted probe tone ratings, $\beta = 0.61$, $p < 0.05$, whereas the A major hierarchy did not, $\beta = 0.36$, *ns*, the correlation with A major underscores a level of interference from this tonality that was not observed when imagining a major tonality.

Further analyses investigated the significant Cue Function by Probe Tone interaction by correlating each listener's averaged probe tone ratings for each cue function with the imaged tonality (A minor), cue pitch, and cue tonality factors employed in Experiment 1. These correlations were then used as dependent variables for a repeated-measures ANOVA with Predictor and Cue Function as within-subjects factors. This analysis revealed a significant effect of Predictor, $F(2, 28) = 4.42$, $MSE = 1.28$, $p < 0.05$, $\eta_p^2 = 0.24$, with planned contrasts showing a difference between Cue Pitch ($M = 0.37$, $SE = 0.09$) and Imaged Tonality ($M = 0.16$, $SE = 0.05$), $F(1, 14) = 5.20$, $p < 0.05$. Cue Tonality ($M = 0.30$, $SE = 0.07$) was a marginally better predictor than Imaged Tonality, $F(1, 14) = 3.72$, $p = 0.07$. These findings are important in demonstrating that pitch-based representations dominated over tonal hierarchy representations in this task, further evidence for

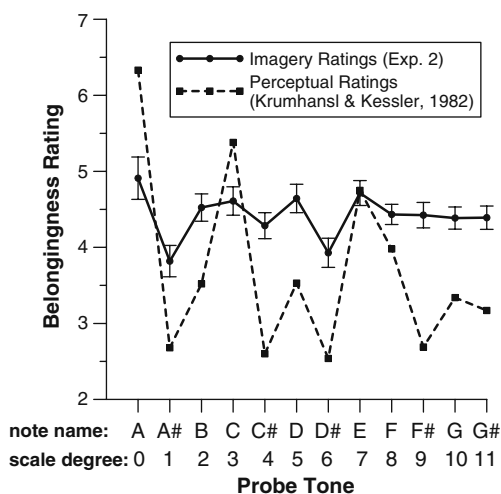


Fig. 5 Mean ratings for the probe tones of Experiment 2, along with the A minor tonal hierarchy

difficulties in imaging a minor key. The only remaining noteworthy effect from this analysis was the significant Predictor by Cue Function interaction, $F(12, 168) = 2.26$, $MSE = 0.14$, $p < 0.05$, $\eta_p^2 = 0.14$. Figure 6 presents the non-transformed correlations, averaged over participant, for the three predictors with respect to cue function, and shows that, converging with Experiment 1, the relative roles of the three predictors varied with the hierarchical stability of the cue.

Subsequent simple effects analyses revealed that effect of Predictor was non-significant for scale degrees 0 [A], 3 [C], and 7 [E], all $F_s < 1.36$, *ns*, but was significant or marginally significant for the remaining scale degrees, all $F_s(2, 28) > 2.39$, $p < 0.11$. Interestingly, this pattern also dissociates those scale degrees occupying the top two levels of the tonal hierarchy from the scale degrees at the next lower level (see Table 1), although it does so in an inverse fashion relative to Experiment 1. Additionally, Cue Pitch and Cue Tonality consistently out-predicted Imaged Tonality when the cue was at a lower level in the tonal hierarchy, but not at upper levels (see Fig. 6). Thus, these findings actually converge with the results for Experiment 1, with tonally important cues leading to better (albeit not great) imagery performance. In this case, however, the impact of tonally important cues was to reduce the influence of pitch information.

Finally, the impact of the pitch distance and key distance factors (described earlier) on imagery variation was assessed. In contrast to Experiment 1, neither measure correlated with the imaged tonality correlations.

This experiment expanded the results of Experiment 1 in two important ways. First, as evidenced by the correlation between the averaged probe tone ratings and Krumhansl and Kessler's (1982) minor tonal hierarchy for A minor, participants can imagine a minor tonality, although collapsing across all of the experimental factors was critical to this

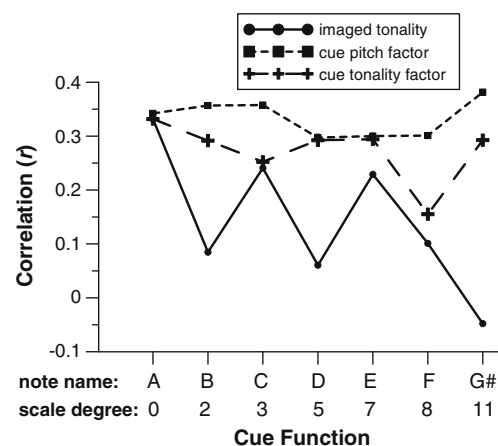


Fig. 6 Averaged correlations between the probe tone ratings and imaged tonality, cue pitch factor, and cue tonality factors as a function of cue function for Experiment 2

finding. This caveat highlights the second principal finding of this study, i.e., that imaged minor tonal hierarchies were less stable psychologically than major tonal hierarchies. This result is important for two reasons. First, and relative to the fundamental aim of this project, this finding demonstrates again a close relation between perceptual and imagery processes, with this evidence taking the form of a limitation in the operation of these two processes.

The second reason this finding is noteworthy is that it actually provides some much needed confirmatory evidence that minor tonalities are less psychologically stable than major tonalities. The idea that major versus minor tonalities are processed differently is common folklore in the music cognition literature, and is based on a number of studies demonstrating differences in perception of these two musical modes (e.g., Cook, 2009; Cook & Hayashi, 2008; Crowder, 1984, 1985; Crowder, Reznick & Rosenkrantz, 1991; Hunter, Schellenberg & Schimmack, 2010; Kastner & Crowder, 1990), most typically in the form of musical emotion (see Hunter et al., 2010, for a review). Very few studies, however, have actually demonstrated that the minor tonality is literally less psychologically stable than the major tonality. Thus, although Krumhansl and colleagues (1982) and Harris (1985) found that major tonics act as stronger cognitive reference points than minor tonics, and Delzell et al. (1999) observed that minor key melodic patterns were more difficult to play by ear than major key melodic patterns, these few studies form the bulk of the evidence for this assertion. Accordingly, these findings are important within the music cognition literature in supporting this idea.

General discussion

Two experiments examined listeners' abilities to produce images of a complex musical object—tonality. Both studies found that after being provided with a cue to this tonality, listeners produced the requested image as demonstrated by predictable variation in the perceived goodness-of-fit of a set of sounded probe tones. More importantly, this variation in the ratings following an imagined context mirrored what would typically be found when such probe tones were heard following a sounded tonal context. Accordingly, tonal imagery induced a comparable perceived hierarchy of stability as tonal perception.

Before considering the implications of this work it is important to address what appears to be an important limitation in these studies, namely, that this work did not explicitly include a sounded tonal context for direct comparison with the imaged tonal context. Given the continual emphasis in this work on the equivalence between perceptual and imagery processing, failing to include a

sounded condition seems a curious oversight. However, it is important to recognize that throughout this project the “gold standard” ratings of Krumhansl and Kessler (1982) were employed for comparison. Because these ratings have been replicated many times in a plethora of contexts, they are generally taken as definitive indicators of perceived tonal organization. Moreover, it is not clear what sounded context would have been most appropriate to employ in this work. Given that participants were not constrained as to how to best generate imagined tonalities, the choice of a specific tonal context in a “perception” condition would by definition have failed to be equivalent to the participant generated “imagery” context. As such, any context would have been suspect in its “match” to the imagery condition, meaning that we would have been left simply employing the classic Krumhansl and Kessler (1982) ratings as comparators. Given that these ratings were themselves derived from multiple key-defining contexts, they are clearly the best indicator of the presence of tonal organization.

Moving beyond this concern, these studies have a variety of implications for our understanding of music processing specifically, and auditory and visual processing more generally. Beginning with the more specific, probably the most important issue in this work for music cognition involves the previously discussed evidence for the psychological instability of the minor tonality.

Of course, the demonstration of a difference in stability between major and minor keys begs the question of what factors underlie this differential stability in the first place. Although a definitive answer to this question cannot be determined, the current results do provide some insight into this issue. One possible explanation stems from the fact that, as described earlier, there are actually three different forms of the minor tonality, each with slightly different theoretical patterns of hierarchical importance. Given the simultaneous existence of multiple, highly similar cognitive structures, the instability of the minor might simply be due to an inherent ambiguity amongst these different representations.

Our studies provide mixed evidence for this explanation. On the one hand, inspection of the ratings from Experiment 2 (see Fig. 5) reveals that of the 12 probes, the ratings are the most differentiated for the first eight scale degrees, and are relatively flattened for the final four scale degrees. Importantly, it is these four final scale degrees that differentiate the different minor forms. Thus, this flattened pattern supports this ambiguity in the representation of these scale degrees, one that could contribute to the psychological instability of the minor.

On the other hand, if these four tones truly underlie this psychological instability, one might expect to see increased stability when looking at only the first eight scale degrees. Unfortunately, this prediction was not

borne out in this study. Specifically, a series of analyses (not formally reported) focusing on the probe tone ratings of just the first eight scale degrees found the same evidence for psychological instability as reported for the complete set of probes, indicating that this instability did not emanate solely, or even primarily, from the final four scale degrees. Moreover, other research suggests that listeners have no difficulty in simultaneously representing the distinctions between the three forms of the minor (Vuvan, Prince & Schmuckler, *under review*), undermining the idea that the mere existence of these highly similar tonal structures produces this psychological instability. Accordingly, the locus of this effect remains undetermined at the moment.

These studies also have implications for more general aspects of psychological processing. First, and most fundamentally, these results demonstrate a strong parallel between imagery and perceptual processes in the auditory domain, comparable to what has been observed in the visual domain (Finke, 1980, 1985, 1989; Kosslyn, 1980; Shepard, 1978, 1982a). According to Hubbard (2010), insight into the relation between auditory perception and imagery is one area in which research has been lacking, and these results provide an important start in addressing this question. Along these lines, noteworthy results include the instantiation of a complex hierarchy of perceived/imagined stability, the importance of factors such as key distance in the perception/imagination of tonality, and the ability of listeners to perceive/image movement between tonalities or key modulation. Clearly, perception and imagery are tapping common psychological processes and mechanisms.

A second, and similarly noteworthy, implication of these results is that our findings also highlight a number of points of convergence between visual and auditory imagery literature. One such point of contact involves the idea that auditory images can contain sophisticated structural elements. As described earlier, previous work on auditory imagery has largely emphasized basic perceptual features, and has only sporadically investigated higher-order structural relations. In contrast, work in visual imagery has routinely demonstrated that visual images incorporate relatively complex structural features. Thus, the finding that participants could form pitch structures containing hierarchically organized information not only provides an important extension of work in auditory imagery, but also brings our conceptualization of the sophistication of auditory images more into line with our understanding of visual images.

Another way in which our studies highlight a convergence between auditory and visual imagery involves the mental transformations of these images. Again, earlier work on auditory imagery has only rarely investigated the

viability of mental transformations of musical images (Cupchik, Phillips & Hill, 2001; Zatorre, Halpern & Bouffard, 2009), although this work does support the idea that visual and auditory mental transformations are related (Cupchik et al., 2001). Our studies add to this literature in demonstrating that increasing frequency distance away from a cue made it increasingly harder to produce an accurate major tonality image. Intuitively, this finding seems related to the idea of mental scanning (Denis et al., 1995; Denis & Kosslyn, 1999; Finke & Pinker, 1982; Kosslyn, 1973; Kosslyn et al., 1978; Pinker, 1980) in which the time required to scan between two points in a mental image is related to the physical distance between these points in a visual scene. And in fact, supporting this mental scanning idea, participants in these studies did report scanning strategies such as mentally “singing from the cue to the tonic of the scale” to accomplish the tonal imaging task. Accordingly, this project provides the first evidence that mental scanning can occur in pitch space, and suggests an analogy between visuospatial and auditory pitch distance (Gentner, 1983; Kubovy, 1988; Kubovy & Schutz, 2010; Kubovy & Van Valkenburg, 2001; McDermott, Lehr & Oxenham, 2008; Prince et al., 2009). Of course, this interpretation must be seen as tentative given that the pitch distance effect was only observed for one of the two experiments. Clearly, though, this finding represents an intriguing avenue for future research.

In sum, our experiments have highlighted a neglected area of study by researchers, i.e., the investigation of higher-order auditory representations in imagery. This work has demonstrated that such higher-order auditory images do exist in listeners, and that they operate according to similar constraints as those based on perception. Moreover, these studies have identified some factors underlying the formation of these images. These studies, however, have also generated at least as many questions as they have answered, questions that have important implications for understanding auditory processes in particular, and basic imagery processes (cutting across modality) in general. Accordingly, future work on this topic has the potential of adding significantly to our understanding of a very basic component of mental life.

Acknowledgments This research and preparation of this manuscript were supported by a doctoral Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada awarded to the first author, and a grant from the Natural Sciences and Engineering Research Council of Canada to the second author. This research was submitted in partial fulfillment of a Masters of Arts degree in Psychology at the University of Toronto by the first author, with portions of this work presented at the 48th Annual Meeting of the Psychonomic Society. The authors would like to thank Jon Prince, Marlene Behrmann, David Waller, Martina Rieger, and two anonymous reviewers for their helpful comments on this work.

References

- Author (2004). *Finale 2005.r2* [Computer software]. Eden Prairie, MN: Makemusic, Inc.
- Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 6(3), 501–515.
- Bharucha, J. J. (1994). Tonality and expectation. In R. Aiello (Ed.), *Musical Perceptions* (pp. 213–239). Oxford: Oxford University Press.
- Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, 13(1), 63–102.
- Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, 12(4), 403–410.
- Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, 58(1), 125–141.
- Canellopoulou, M., & Richardson, J. T. E. (1998). The role of executive function in imagery mnemonics: Evidence from multiple sclerosis. *Neuropsychologia*, 36(11), 1181–1188.
- Cook, N. D. (2009). Harmony perception: Harmoniousness is more than the sum of interval consonance. *Music Perception*, 27(1), 25–41.
- Cook, N. D., & Hayashi, T. (2008). The psychoacoustics of harmony perception. *American Scientist*, 96(4), 311–319.
- Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, 7(1), 20–43.
- Cooper, L. A., & Shepard, R. N. (1973). The time required to prepare for a rotated stimulus. *Memory & Cognition*, 1(3), 246–250.
- Cooper, L. A., & Shepard, R. N. (1975). Mental transformations in the identification of left and right hands. *Journal of Experimental Psychology: Human Perception and Performance*, 10(1), 48–56.
- Crowder, R. G. (1984). Perception of the major/minor distinction: I. Historical and theoretical foundations. *Psychomusicology*, 4(1–2), 3–12.
- Crowder, R. G. (1985). Perception of the major/minor distinction: II. Experimental investigations. *Psychomusicology*, 5(1–2), 3–24.
- Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 472–478.
- Crowder, R. G., Reznick, J. S., & Rosenkrantz, S. L. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society*, 29(3), 187–188.
- Cupchik, G. C., Phillips, K., & Hill, D. S. (2001). Shared processes in spatial rotation and musical permutation. *Brain and Cognition*, 46(3), 373–382.
- Currie, G. (1995). Visual imagery as the simulation of vision. *Mind and Language*, 10(1), 25–44.
- Delzell, J. K., Rohwer, D. A., & Ballard, D. E. (1999). Effects of melodic pattern difficulty and performance experience on ability to play by ear. *Journal of Research in Music Education*, 47(1), 53–63.
- Denis, M., Goncalves, M., & Memmi, D. (1995). Mental scanning of visual images generated from verbal descriptions: Towards a model of image accuracy. *Neuropsychologia*, 33(11), 1511–1530.
- Denis, M., & Kosslyn, S. M. (1999). Scanning visual mental images: A window on the mind. *Cahiers de psychologie cognitive*, 18(4), 409–616.
- Deutsch, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics*, 11(6), 411–412.
- Deutsch, D. (1979). Octave generalization and the consolidation of melodic information. *Canadian Journal of Psychology*, 33(3), 201–205.
- Dirkx, E., & Craik, F. I. (1992). Age-related differences in memory as a function of imagery processing. *Psychology and Aging*, 7(3), 352–358.
- Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, 87(2), 113–132.
- Finke, R. A. (1985). Theories relating mental imagery to perception. *Psychological Bulletin*, 98(2), 236–259.
- Finke, R. A. (1989). *Principles of mental imagery*. Cambridge: MIT Press.
- Finke, R. A., & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(2), 142–147.
- Finke, R. A., & Schmidt, M. J. (1977). Orientation-specific color aftereffects following imagination. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 599–606.
- Finke, R. A., & Shepard, R. N. (1986). Visual functions of mental imagery. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (Vol. 2). New York: Wiley-Interscience.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155–170.
- Gordon, I. E., & Hayward, S. (1973). Second-order isomorphism of internal representations of familiar faces. *Perception & Psychophysics*, 14(2), 334–336.
- Halpern, A. R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 434–443.
- Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, 17(5), 572–581.
- Harris, R. W. (1985). Perceived relatedness of musical tones in major and minor tonal contexts. *The American Journal of Psychology*, 98(4), 605–623.
- Hartley, A. A. (1977). Mental measurement in the magnitude estimation of length. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 622–628.
- Hartley, A. A. (1981). Mental measurement of line length: The role of the standard. *Journal of Experimental Psychology: Human Perception and Performance*, 7(2), 309–317.
- Hubbard, T. L. (2010). Auditory imagery. *Psychological Bulletin*, 136(2), 302–329.
- Hubbard, T. L., & Stoeckig, K. (1988). Musical imagery: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(4), 656–667.
- Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts*, 4(1), 47–56.
- Huron, D. (2006). *Sweet anticipation*. Cambridge: MIT Press.
- Intons-Peterson, M. J. (1980). The role of loudness in auditory imagery. *Memory & Cognition*, 8(5), 385–393.
- Intons-Peterson, M. J. (1992). Components of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 45–72). Hillsdale: Lawrence Erlbaum Associates.
- Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics*, 68(5), 829–844.
- Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, 8(2), 189–202.
- Kerst, S. M., & Howard, J. H. (1978). Memory psychophysics for visual area and length. *Memory & Cognition*, 6(3), 327–335.
- Knauff, M., & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition*, 30(3), 363–371.
- Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception & Psychophysics*, 14(1), 90–94.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge: Harvard University Press.

- Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge: MIT Press.
- Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of imagery scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 4(1), 47–60.
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. New York: Oxford University Press.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11(3), 346–374.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, 126(1), 159–179.
- Krumhansl, C. L., Bharucha, J. J., & Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics*, 32(2), 96–108.
- Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, 8(1), 24–36.
- Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition*, 11(4), 325–334.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89(4), 334–368.
- Krumhansl, C. L., & Schmuckler, M. A. (1986). The Petroushka chord: A perceptual investigation. *Music Perception*, 4(2), 153–184.
- 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(4), 579–594.
- Kubovy, M. (1988). Should we resist the seductiveness of the space: Time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 318–320.
- Kubovy, M., & Schutz, M. (2010). Audio-visual objects. *Review of Philosophy and Psychology*, 1(1), 41–61.
- Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, 80(1–2), 97–126.
- Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, 17(4), 481–509.
- Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences*, 9(1), 26–33.
- Marmel, F., Tillmann, B., & Dowling, W. J. (2008). Tonal expectations influence pitch perception. *Perception & Psychophysics*, 70(5), 841–852.
- McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science*, 19(12), 1263–1271.
- Moler, C. (2004). MATLAB (Version 7.0) [Computer software]. Natick, MA: The Mathworks, Inc.
- Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal Learning & Verbal Behavior*, 4(1), 32–38.
- Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, 3(6), 635–647.
- Paivio, A., & Foth, D. (1970). Imaginal and verbal mediators and noun concreteness in paired-associate learning: The elusive interaction. *Journal of Verbal Learning & Verbal Behavior*, 9(4), 384–390.
- Pinker, S. (1980). Mental imagery and the third dimension. *Journal of Experimental Psychology: General*, 109(3), 354–371.
- Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). The effect of task level and pitch structure on pitch-time interactions in music. *Memory & Cognition*, 37(3), 368–381.
- Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance*, 35(5), 1598–1617.
- Reisberg, D. (Ed.). (1992). *Auditory imagery*. Hillsdale: Lawrence Erlbaum Associates.
- Revesz, G. (1954). *Introduction to the psychology of music*. Norman: University of Oklahoma Press.
- Schellenberg, E. G., & Trehub, S. E. (2003). Accurate pitch memory is widespread. *Psychological Science*, 14(3), 262–266.
- Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, 7(2), 109–150.
- Schmuckler, M. A. (1990). The performance of global expectations. *Psychomusicology*, 9(2), 122–147.
- Schmuckler, M. A. (2004). Pitch and pitch structures. In J. G. Neuhoff (Ed.), *Ecological psychoacoustics*. New York: Academic.
- Schmuckler, M. A. (2009). Components of melodic processing. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Handbook of Music Psychology*. Oxford: Oxford University Press.
- Schmuckler, M. A., & Tomovski, R. (2005). Perceptual tests of an algorithm for musical key-finding. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5), 1124–1149.
- Sergeant, D. (1983). The octave: Percept or concept. *Psychology of Music*, 11(1), 3–18.
- Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, 36(12), 2346–2353.
- Shepard, R. N. (1978). The mental image. *The American Psychologist*, 33(2), 125–137.
- Shepard, R. N. (1982a). Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89(4), 305–333.
- Shepard, R. N. (1982b). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic.
- Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology*, 1(1), 1–17.
- Shepard, R. N., Kilpatrick, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology*, 7(1), 82–138.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701–703.
- Smith, N. A., & Cuddy, L. L. (2003). Perceptions of musical dimensions in Beethoven's *Waldstein* sonata: An application of Tonal Pitch Space theory. *Musicae Scientiae*, 7(1), 7–31.
- Toivianen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception*, 32(6), 741–766.
- Vuvan, D. T., Prince, J. B., & Schmuckler, M. A. (under review). Probing the minor tonal hierarchy. *Music Perception*.
- Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, 64(2), 198–207.
- Watkins, M. J., & Schiano, D. J. (1982). Chromatic imaging: And effect of mental coloring on recognition memory. *Canadian Journal of Psychology*, 36(2), 291–299.
- Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2009). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, 22(4), 775–789.