A comparative analysis of auditory perception in humans and songbirds: A modular approach

Ronald Weisman a, *, Marisa Hoeschele b, Christopher B. Sturdy c

a Department of Psychology, Queen’s University, Kingston, ON, Canada K7L 3N6
b Department of Cognitive Biology, University of Vienna, Vienna, Austria
c Department of Psychology, University of Alberta, Edmonton, AB T6G 2E9, Canada

ABSTRACT

We propose that a relatively small number of perceptual skills underlie human perception of music and speech. Humans and songbirds share a number of features in the development of their auditory communication systems. These similarities invite comparisons between species in their auditory perceptual skills. Here, we summarized our experimental comparisons between humans (and other mammals) and songbirds (and other birds) in their use of pitch height and pitch chroma perception and discuss similarities and differences in other auditory perceptual abilities of these species. Specifically, we introduced a functional modular view, using pitch chroma and pitch height perception as examples, as a theoretical framework for the comparative study of auditory perception and perhaps all of the study of comparative cognition. We also contrasted phylogeny and adaptation as causal mechanisms in comparative cognition using examples from auditory perception.

Keywords:
Comparative cognition
Auditory perception
Pitch height perception
Pitch chroma perception
Octave equivalence
Double dissociation

At the core of the study of comparative cognition is the question of how best to study similarities and differences in cognitive abilities among species. Researchers have noted similarities worth exploring between human speech and music on the one hand and birdsong on the other (Doupe and Kuhl, 1999).

Both humans and songbirds learn their vocalizations early in life, exhibit a strong dependence on hearing the adults they will imitate, as well as hearing themselves as they practice, and exhibit a diminished dependence on these models as they mature. Innate predispositions for perceiving and learning the correct sounds exist in both groups. Humans also share with songbirds an early phase of learning that is primarily perceptual, which then serves to guide later vocal production. Both humans and songbirds have evolved a complex hierarchy of specialized forebrain areas in which motor and auditory centers interact closely. (Paraphrased from Doupe and Kuhl, 1999).

Here, we use a modular approach to compare some of the most important auditory skills in the perceptual toolkits humans and songbirds use to perceive and produce their communication systems. We ask how animals use their auditory skills to produce and perceive their vocalizations. By this we mean that a species auditory toolkit is a critical determinant of its auditory communication. For example as a first approximation, the toolkit for humans must include neural machinery for perception of relative pitch, pitch height, pitch chroma, temporal duration, rhythm, and tempo (some of the definitions follow and Burns (1999) provides further definitions).

The goal is to use the present work to provide examples of a modular approach to the study of comparative cognition. Some would insist on an extensive a priori definition of modules (e.g., Fodor, 1983). We have a different view; functional modules are products of evolution and obey no a priori rule except natural selection (Barrett and Kurzhan, 2006). We are aware of the controversy over modules and think it unhelpful, especially in the case of perceptual modules. Readers who wish to consider the controversy further are referred to Barrett and Kurzhan’s (2006) extensive discussion.

We describe comparisons between humans and songbirds in two common and powerful auditory perceptual abilities: perception of pitch height and of octave equivalence. Pitch height is a log-linear scale of pitch where the greater the difference between two frequencies, the greater the difference in perceived pitch. When a doubling or halving of frequency separates two successive sounds, humans can perceive this unique relationship as octave equivalence. In musical terms, notes played at the same chroma (relative position) in different octaves can be perceived as similar, that is, equivalent in chroma. Pitch chroma perception is always a cognitive skill because it requires categorization of pitches based
Pitch height perception can require only the decision that one pitch is higher than another, but as a cognitive skill requires categorization of contiguous pitches as sharing a common outcome, as in categorizing voices as adult male or female. Here to avoid long windy definitions, we simply refer to these pitch categorizations as pitch height and pitch chroma perception.

Pitch height and pitch chroma can be seen as opposing perceptions. To a listener attending to pitch chroma, pitches with a frequency relationship of 1:2 seem more similar than those that share a relationship of 3:4; for a listener attending to pitch height, the opposite would be true. In humans, both abilities are important in speech and music (Burns, 1999). Later, we will briefly touch on relative pitch and note timing, which are also important to human and songbird communication.

In trying to make sense of published findings about similarities and differences among species in their auditory skills, we have developed a modular theory of comparative auditory cognition, which might apply quite generally. We freely admit being influenced by current research in neuroscience (Wagner et al., 2007) and evolutionary psychology (Barrett and Kurzban, 2006) in our thinking about a modular view of the mind and the brain. Our working assumption is that cognition consists mainly of the operation of fairly simple and independent modules, which by their operation or lack thereof account for differences in the cognitive abilities of various species. We will have more to say about the theory as we describe our findings about the comparative cognition of auditory perception. Most important, we will introduce our adaptation of some basic neuroscience tools to study the comparative science of cognition.

1. Laboratory protocols for the study of auditory perception in songbirds and humans

We used operant discrimination tasks to evaluate perceptual functioning across species. We show the songbird testing apparatus in Fig. 1. The main features of the songbird testing apparatus are: (i) a monitored perch the birds use to trigger the presentation of acoustic stimuli, (ii) a monitored, motorized feeder, and (iii) a high-quality speaker (Sturdy and Weisman, 2006). The human testing apparatus was a laptop computer and a pair of high-quality over-the-ear headphones (Hoeschele et al., 2012). In research with rats (Rattus norvegicus; Weisman et al., 2004) and pigeons (Columba livia; Weisman et al., 2010b), standard operant chambers were used with high-quality speakers. In experiments with songbirds, pigeons, and rats, the rewards were foods appropriate to the species; in experiments with humans, positive feedback and competitive rewards were standard.

Laboratory protocols are important to the extent that they illuminate how animals use their cognitive skills in their daily lives; the assumption is that expert abilities shown in the laboratory translate to highly adaptive behavior in nature. However, field research must always be used to support laboratory findings.

2. Discriminations based on pitch height perception

We began our studies of pitch height perception with a well-known Australian songbird, the zebra finch (Taeniopygia guttata; Weisman et al., 1998) and we expanded to studies of many other birds and mammals (see Weisman et al., 2004). In two experiments, we compared songbirds and humans in their ability to sort frequencies into ranges. In one experiment (see Weisman et al., 1998, Experiment 1), we divided 27 tones, beginning at either 359 Hz or 1,200 Hz and increasing 6% from one tone to the next into three ranges of nine tones each. Tones in lower and upper ranges were not rewarded (go or S-tones) and tones in the middle range were rewarded (go or S+ tones). Training continued for 5,000–7,000 trials in both species.

A second experiment (see Weisman et al., 1998, Experiment 2) grouped pitches in a rough approximation of the ranges of
frequencies found among individual notes in the songs and calls of conspecific songbirds. In this more difficult, eight-range, discrimination, 40 sinewave tones increased in frequency from one to the next by 120 Hz. The lowest adjacent five successive tones were go tones, responses to these S+ tones were rewarded. Responses to the next five lowest adjacent (no-go, S−) tones were not rewarded. Reward and non-reward stimuli alternated across the eight (see Table 1). In both species, the assignment of reward to tones was counterbalanced between groups. Training continued for 5,000-6,000 trials in both species.

It is important to understand that although reward was presented on the basis of membership in particular ranges of frequencies, the tones were presented at random without replacement in multiple blocks of trials each equal to the number of tones in the discrimination. In other words, there were never sequential cues to the pitches of subsequent tones.

The principal results of the three-range discrimination in zebra finches and humans are shown in Fig. 2. These are discrete trial operant discriminations so it is convenient to report the percent responses to specific tones and ranges of tones. To summarize the results shown in Fig. 2, within-species differences in the final level of performance varied from non-significant to inconsequential from one spectral region to another. In the cross-species comparison with humans, zebra finches appeared to have superior ability in

<table>
<thead>
<tr>
<th>Frequency Range</th>
<th>S− First Group</th>
<th>S+ First Group</th>
<th>Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S−</td>
<td>S+</td>
<td>980</td>
</tr>
<tr>
<td>2</td>
<td>S+</td>
<td>S+</td>
<td>1,580</td>
</tr>
<tr>
<td>3</td>
<td>S+</td>
<td>S+</td>
<td>2,180</td>
</tr>
<tr>
<td>4</td>
<td>S−</td>
<td>S+</td>
<td>2,780</td>
</tr>
<tr>
<td>5</td>
<td>S−</td>
<td>S+</td>
<td>3,380</td>
</tr>
<tr>
<td>6</td>
<td>S+</td>
<td>S+</td>
<td>3,980</td>
</tr>
<tr>
<td>7</td>
<td>S−</td>
<td>S+</td>
<td>4,580</td>
</tr>
<tr>
<td>8</td>
<td>S+</td>
<td>S+</td>
<td>5,180</td>
</tr>
</tbody>
</table>

**Fig. 3.** Mean percentage of responses to eight ranges of five sinewave tones each for zebra finches and humans during the final discrimination session, in the S+ first and S− first frequency-range discrimination counterbalancing groups. Discrimination by zebra finches was precise and accurate, approximating that of a perfect sorter. Discrimination by humans was non-existent, except that humans anchored their discriminations on the first and last ranges of the task, with responding falling off from the first S+ range to the last S− range in the S+ first condition and rising from the first S− range to the last S+ range in S− first condition. Reproduced from Fig. 3 in Weisman et al. (1998).
classifying the frequencies in both the lower and higher frequency regions, despite the fact that the lower frequency region tones are more common in music and the higher frequency tones are more common in birdsongs.

The principal results of the eight-range discrimination in zebra finches and humans are shown in Fig. 3. The most notable result was that zebra finches approximated the performance of a perfect sorter, responding to alternations between S+ and S− ranges of tones with remarkable accuracy (see Fig. 3). The birds did respond slightly more to some S+ tones adjacent to the S+ ranges and slightly less to some S+ tones adjacent to the S− ranges in the eight-range task suggesting some spread of excitation and inhibition across range boundaries. Also, zebra finches showed weak anchor effects in the eight-range tasks, that is, they responded more in the lowest S+ range and less in the highest S− range when the lowest range was S+, whereas responding had the opposite trend when the lowest range was S−. Humans were sensitive to the overall relationship between pitch height and reward, that is, they showed strong anchor effects, but were unable to sort the frequencies into S+ and S− ranges into distinct ranges. In both species, differences between the counterbalanced groups were entirely predictable from the reversal of go and no-go tones.

Several other songbird species have discriminated among these same ranges with precision (Weisman et al., 2010b). In contrast, the eight-range discriminations of humans and rats were little influenced by the alternation of S+ and S− ranges. Instead, anchor effects of the first and last ranges dominated the discrimination (Weisman et al., 2004). In summary, as the number of tones and numbers of ranges that required sorting increased, humans went from 85% accuracy to 53% accuracy, whereas songbirds fell only slightly from 95% to 87% accuracy. With increases in task difficulty humans’ discriminations became very much inferior to those of songbirds.

3. Discriminations based on pitch chroma perception

Music theory depends on the equivalence of pitch chroma across octaves, but as Burns (1999) pointed out, the experimental evidence is confusing. For example, rats appear to show octave equivalence (Blackwell and Schlosberg, 1943) but humans untrained in music appear to show little evidence of it (Allen, 1967). We extended our use of operant discrimination tasks from testing for pitch height to testing for octave equivalence perception; this had the principal benefit of allowing us to use similar tasks in humans (Hoeschele et al., 2012) and songbirds (Hoeschele et al., 2013).

The songbird study species was black-capped chickadees (Poecile atricapillus). Black-capped chickadees reside year round in Canada and in the northern two-thirds of the United States (Smith, 1991). The testing apparatus was the same in these experiments as in the work on pitch height perception.

Multi-stage discrimination and reversal training were needed to test for octave equivalence based on the similarity between pitches at the same chroma played in different octaves. In the first stage, songbirds and humans learned a three-range discrimination that divided Octave 4 into three ranges of four notes each and designated the first and last four notes as S− notes and the middle four notes as S+ notes (see Fig. 4 for the procedure). Humans were trained for 480 trials and chickadees for 4,000 trials. Each species were stable before the end of training.

After acquisition of the three-range discrimination in Octave 4, pitch chroma generalization (octave equivalence) was tested in Octave 5 simultaneously with maintenance of the discrimination in Octave 4 (see Fig. 5 for the procedure). In each species, one group was trained in the original discrimination in Octave 4 and in Octave 5 (see Fig. 5). A second group was trained in original discrimination in Octave 4 and a reversal in Octave 5. In the reversal group, responses to C, C♯, D, or D♯ and the upper notes: G♯, A, A♯, or B (the S+ notes in Octave 5) were rewarded and responses to the middle four notes in Octave 5: E, F, F♯, or G (the S− notes in Octave 5) were not rewarded (see Fig. 5). The logic was straightforward: if responding based on chroma generalized to Octave 5, then acquisition of the original discrimination should have been facilitated and acquisition of the reversal retarded. We used the same training procedures for transfer as for training (Fig. 6).

The exact frequency of each note can be found using the formula, frequencyχ = 2^(1/12) × frequencyχ−1, where C4 = 262 Hz. During original discrimination training, responses to the middle four notes in Octave 4: E, F, F♯, or G (the S+ notes) were rewarded. Responses to the lower notes: C, C♯, D, or D♯ and the upper notes: G♯, A, A♯, or B (the S− notes) were not rewarded. Notice that this discrimination is very similar to those we tested in studies of three-range pitch height perception.

Humans transferred accurately to the original discrimination in Octave 5 but continued to show the same pattern after extensive

![Fig. 4. A summary of the procedure for original discrimination training in Octave 4, which provided reward for responding to a range of S+ notes in the middle of Octave 4 but withheld reward for responding to S− notes in upper and lower note ranges within the same octave. Reproduced from Fig. 1 in Hoeschele et al. (2013).](image)

![Fig. 5. Summary of the transfer procedures: (A) In the original discrimination transfer group, discrimination training provided reward for responding to a range of S+ notes in the middle of Octaves 4 and 5, but withheld reward for responding to other (S−) notes in higher and lower note ranges in both octaves. (B) In the reversed discrimination transfer group, in Octave 4 discrimination was as in A, but in Octave 5, the middle range was S− notes, and the upper and lower ranges were S+ notes. Reproduced from Fig. 2 in Hoeschele et al. (2013).](image)
Fig. 6. The average percent response to each note range during discrimination training in Octave 4. Results are shown for humans on the left and for chickadees on the right. Error bars represent standard error of the mean but are not shown when they are narrower than the symbol for the mean. Reproduced from Fig. 3 in Hoeschele et al. (2013).

Fig. 7. The average percent response to each note range during transfer training in Octaves 4 and 5 separated by each transfer of training group. Both groups had the same reinforcement contingencies as during discrimination training in Octave 4 (i.e., S−, S+, S−). The original discrimination transfer group (top panel) had the same contingencies in Octave 5 as in Octave 4 (i.e., S−, S+, S−). The reversed discrimination transfer group (bottom panel) had the opposite contingencies in Octave 5 (i.e., S+, S−, S+). Results are shown for humans on the left and for chickadees on the right. The dotted line separates the training and testing octaves. Error bars represent standard error of the mean. Reproduced from Fig. 3 in Hoeschele et al. (2013).
training with the reversed discrimination. That is, generalization of the three-range discrimination learned in Octave 4 to Octave 5 interfered with acquisition of the reversed discrimination. The results for chickadees were opposite to expectation from pitch chroma generalization (see Fig. 7); in Octave 5 transfer of the original discrimination was retarded and we observed positive transfer to the reversed discrimination.

Based on results during discrimination training in Octave 4, one might suppose that humans and chickadees would show similar transfer to stimuli in Octave 5, but this is not what we observed. Performance diverged dramatically between species during transfer (see Fig. 7 for the comparison). Humans showed positive transfer to training in the same pattern of reward in Octave 5 as in Octave 4. In contrast, chickadees showed positive transfer to a reversal of the pattern of reward in Octave 4 during training in Octave 5 (see Fig. 7). These results suggest that humans and chickadees do not acquire similar representations and expectations during the original discrimination in Octave 4.

The evidence supported the hypothesis that pitch height predominates over octave equivalence in chickadees. Recall that songbirds learn easily that go (S+) and no-go (S−) ranges alternate in discriminations of up to eight pitch ranges (Weisman et al., 1998). Our hypothesis is that during original training in Octave 4, chickadees learned that an S+ range follows an S− range, and an S− range follows an S+ range, so that when they were presented with novel notes in Octave 5, the birds expected a continuation of the sequence of alternating ranges: an S+ range of notes followed by an S− range, and an S− range followed by an S+ range. This is the opposite pattern to that taken as evidence for octave generalization.

We show only performance averaged over four-note ranges here. Performance for individual notes was presented in the original articles (Hoeschele et al., 2012; Hoeschele et al., 2013) and revealed no surprises.

These studies of pitch chroma perception provide important findings; they suggest very different schema for transposing patterns of expectations across pitches in humans and songbirds. We need to consider alternative explanations before accepting the remarkable hypothesis that the three note-range discrimination taught chickadees only to expect alternating ranges of notes to have different consequences. Among the more likely alternatives is the qualifying hypothesis that songbirds perceive both pitch height and pitch chroma but in some circumstances, including the present ones, height predominates over chroma. Evidence from nature would tilt our conclusions toward the predominance hypothesis: for example, if songbird species could be found that sings notes at the same pitch chroma in different octaves.

Humans tutored in music did better in octave transfer to the original task and worse in transfer to the reversal task, suggesting that music training increased the perception of octave equivalence. Chickadees undergo their own form of auditory training as they learn to produce songs with remarkably constant frequency ratios (Weisman et al., 1990; Lohr, 2008), but they did not give evidence of octave equivalence here.

4. Positive and negative controls in comparative cognition

In cognitive research, comparisons among closely related species are often preferred to those among distantly related species. Closely related species are preferred mainly because their differences are likely to be easier to relate to the species recent evolutionary histories. In terms of methodology, too, comparisons between closely related species are often easier: for example, the procedures for observing the species are likely to be more similar. Also, more closely related species tend to differ less in their perceptual sweet spots.

Because we humans tend to make our species the standard for all cognitive skills, comparisons between humans and species distantly related to humans form a significant portion of the comparative literature. These comparisons present opportunities for egregious error. For example, experiments comparing perception in humans and songbirds might seem to demonstrate consistently better auditory perception in humans than songbirds because of some undocumented advantage that favored humans, such as an advantage in following human designed instructions or in using computers to solve discriminations. Conversely the undocumented advantage might favor songbirds, such as our better control of their motivational states: songbirds but not humans were required to work for food and songbirds were trained in much longer sessions that favored more intense learning. Most important here, the sweet spots for best hearing and resolution of complex auditory events on the frequency spectrum may differ in species as distantly related as humans and songbirds.

Whatever the source of undocumented advantage, careful comparative research across auditory perceptual tasks can reveal its presence. The comparative method can yield especially useful results when distantly related species differ in the perceptual abilities necessary to a cognitive skill. That is whether by chance or design, when one species has some of the advantages and the other species some different advantages, we can eliminate egregious methodological errors and we can better describe the pattern of strengths possessed by both species.

On the methodological side, when humans and songbirds are tested at the same pitches, it is possible that one species has an undocumented advantage. For example, given the region of the auditory spectrum presented in the eight-range discriminations, the increased pitch height performance in songbirds could be due to increased relevance of the test frequencies for bird calls and songs. Similarly, given the region of the auditory spectrum presented in the three-range discrimination that formed the basis of our tests of octave generalization, increased pitch chroma performance in humans could simply be due to our choices of test frequencies similar to those in music.

These hypotheses appear at first blush correct on their face without further evidence. But converging evidence from several studies that provide positive controls for one another disprove explanations based on natural sweet spots for perception in songbirds and humans. Consider the evidence: (a) In the three-range discrimination in Octave 4 reported here (see Hoeschele et al., 2013), songbirds discriminated note ranges more precisely than humans, yet humans showed octave generalization and songbirds did not. The frequencies in Octaves 4 and 5 are common in music but not in birdsongs, so why did chickadees discriminate among the ranges better than humans? (b) In the eight-range task reported here (and Weisman et al., 1998), songbirds discriminated accurately and normal humans did not. The frequencies in the eight-range discriminations are more common in birdsongs than music, so why did we (Weisman et al., 2010a) find humans with absolute pitch (AP) both accurately named the notes represented by the 40 frequencies in the eight-range discrimination and accurately solve the go/no-go discrimination itself (Fig. 8)? The answer is that AP possessors use pitch chroma (which of course depends on pitch perception) to solve the discrimination. Interestingly, AP possessors made a class of errors made by neither non-AP humans or by zebra finches: when a pitch chroma was represented in two or more octaves, sometimes as an S+ sometimes an S−. AP possessors made anomalous errors (see the dark short vertical lines in Fig. 8 and Weisman et al., 2010a). (c) In research not reported in detail here, non-AP humans discriminated the durations of 3,000 Hz tones about equally to songbirds (Weisman et al., 1999), and non-AP humans discriminated the pitch intervals of tone pairs in the 2,400-5,100 Hz frequency region more quickly than songbirds and
to a slightly higher standard than songbirds (Weisman et al., 1994). Why were AP possessors and even ordinary humans able to use frequencies not found in music to solve discriminations based on fundamental musical skills? The evidence we have presented shows that neither humans nor songbirds are confined to using their considerable pitch classification skills in the narrow region of the acoustic spectrum in which they vocalize or play music.

The studies we discuss here demonstrate the power of positive and negative controls that arise out of careful and extensive use of simple experimental protocols. We did not set out to prove frequency was less important than the perceptual categorization of frequencies and durations, but the conclusion seems inescapable. We did set out to develop protocols that permitted us to directly compare both closely and distantly related species in their fundamental auditory abilities, and as much by good fortune as good design we appear to have accomplished this goal. And achieving this goal has led to more than better protocols. For example, our results indicate that over the very wide region of frequencies that humans and songbirds hear and classify, their accuracy depends much on the perceptual ability required by the test and hardly at all on the frequencies tested.

5. A comparative cognition theory based on modularity

We have presented the results of our studies of pitch height and pitch chroma perception in humans and songbirds without seeking to place them in the context of a theoretical approach to comparative cognition: a task we undertake here. The evidence presented so far suggests that it may be possible to perceive pitch height without accurately perceiving pitch chroma and to perceive pitch chroma without accurately perceiving pitch height.

Over the past 50 plus years, neuroscientists, beginning with Hans-Lukas Teuber (1955), have proposed a simple set of experimental designs to infer independence of function between brain regions. These are the single and double dissociation designs. If lesions in one location lead to difficulty on task A, but not task B the result is termed a single dissociation. If lesions in a different location lead to the opposite pattern of performance, then tasks A and B likely rely on independent neural mechanisms: the result is termed a double dissociation. Implicit in the use of disassociations is the assumption of modularity: that modular processes account for most cognitive functions, particularly those in sensation and perception. This sort of thinking is fundamental to neuroscience, but it is not free of criticism (see Van Ordena and Pennington, 2001).

Here, we present an adaptation of disassociation experiments to comparisons between cognitive functions among species. The inferences in comparative studies follow similar logic to those in studies within a single species: if species A is accurate at task X and species B is accurate at task Y but not the opposite (a double disassociation), then our version of comparative theory proposes that species A and B possess independent functional modular cognitive processes that account for differences in their performance on Tasks A and B. Some conditions apply when evaluating these inferences: differences between the species must be strategic or substantial, that is, of broad significance or of considerable size. Significance refers to biological significance, a feature critical to important broadly distributed cognitive skills. Substantial refers to the size of the difference between species in measures of competence in use of the two skills. We can roughly divide competence into five levels, for example, absent or negligible, slight but functional, mediocre, accurate, and expert. We view these levels as substantially different: for example, a species with negligible auditory ability is at a substantial disadvantage relative to a species with even slight but functional auditory ability.

Consider the application of the logic of double disassociations to comparative cognition, in particular, differences between humans and songbirds in pitch height and pitch chroma perception. Table 2 summarizes the experimental findings discussed earlier in this article: Humans show mediocre and songbirds show expert pitch height perception. This could mean that only humans are at a disadvantage under the protocol used in the test. The test for pitch chroma perception shows that humans are capable of at least accurate pitch chroma perception and songbirds show negligible pitch chroma perception under the same protocol applied in the pitch height experiments. Assessing the results together allows the
Table 3 summarizes further comparisons between humans and songbirds in their auditory abilities. Specifically, we compared the species in relative pitch (i.e., pitch ratio; Weisman et al., 1990; 1994) and note timing (Weisman et al., 1999) abilities and found them both accurate in humans and songbirds. These findings provide further positive controls for the disassociation between pitch height and chroma. More importantly, the findings demonstrate remarkable functional similarities between humans and songbirds in relative pitch and in pitch timing. Functional similarities do not demonstrate neural similarities although they suggest that timing and relative pitch perception might have been conserved across mammals and birds.

Not everyone is sanguine about the use of dissociations to demonstrate modular brain functions (see Van Ordena and Pennington, 2001) and some will be unimpressed by our attempt to apply the logic of modular function across species. Those who oppose modular units of brain function prefer the conclusion that neural functions are relatively independent of the brain’s structural elements (Lashley, 1930; Van Ordena and Pennington, 2001) and appear to reason mainly from hypothetical evidence against modularity. For example, claims that modularity fails to converge on a fixed set of cases and that it perpetuates unending fractionation into more modules are unsupportable. Experiments regularly confirm some disassociations and fail to confirm others and there is no infinite regress of disassociations (Wagner et al., 2007). Most important, here we have shown that disassociations are useful concepts in the study of comparative cognition. Finally, one can find modules useful in neuroscience and in comparative cognition without claiming that products of functional modules are never integrated elsewhere in the brain.

6. Parsing environmental and evolutionary causation from phylogenetic causation

It is possible to adopt an entirely environmental view of the differences between songbirds and humans we report here: Perhaps if humans were never exposed to music and songbirds were intensely exposed, their respective perceptions of pitch chroma and pitch height might be reversed. Of course, these suggested experiments beg questions of why humans compose, produce, and sing music and why songbirds adapt songs from conspecifics.

We recommend a different hypothesis: the interaction between experience and genotype, if the experience is suitable, can cause the genes to be transcribed into proteins that change behavior. For example, one consequence of exposure to music in humans and conspecific song in songbirds is the release of the reward coding neurotransmitter dopamine (Salimpoor et al., 2011; Heimovics and Ritters, 2008). No one has yet thought to reverse the conditions and play birdsong to humans and music to songbirds and test for dopamine release - a successful reversal seems highly improbable.

Tinbergen (1963) famously parsed causation in behavioral biology into two evolutionary and two proximal questions. Here, we discuss confusions between the evolutionary causes: adaptation and phylogeny. We are all familiar with adaptation as the engine of evolutionary change. Behavioral clines are such an evolutionary change: for example, when species and subspecies of deer mice, *Peromyscus*, began to occupy niches in North America, they adapted their nesting behavior to those niches, such that species that live in colder climates build bigger nests than those living in warmer climates and species living in intermediate climates build nests that are intermediate in size, independent of the weight of individuals of each species, and independent of individual species members' experience (King et al., 1963). Biologists and evolutionary psychologists fill fine journals with work on biological adaptations and no reasonable person can doubt the power of natural selection.

However, and this is the point, not all differences between species are the result of selection for specific adaptations in individual species within their evolutionary histories. Perhaps the best example of phylogenetic causation is the evolution of fur and feathers. Mammals are protected by hair (i.e., fur) and feathers protect birds. That fur and feathers are adaptations is indisputable, but it is absurd to explain the protective coats of seals and penguins as adaptations to any aspect of these species' current environments. Instead, fur and feathers can be traced back to species alive very early in the history of mammals and birds and both adaptations function to protect animals from the sun and extreme temperatures.

Not all phylogenetic causation is as clearly bifurcated as fur and feathers; other instances, like variants of color vision have more values, for example, birds see four different colors whereas humans see three (Eaton, 2005). In the study of auditory perception, pitch height is one example: birds generally have vastly more accurate pitch height perception than mammals, but if the classification task is simple enough (three ranges), mammals can discriminate almost as well as birds (Weisman et al., 2004). As the task becomes more difficult, avian species show their expert performance and humans and rats fail dismaly.

Some suppose that special tutorials might help humans to excel at pitch height discriminations. When we undertook extensive training in pitch height, our normal human (non-AP) participants failed to improve in 4,000 or even 7,000 trials, but human AP possessors did learn the eight–range discrimination, using their prodigious note naming abilities (Weisman et al., 2010a).

So far, our evidence supports the hypothesis that songbirds lack the perception of pitch chroma for successive sounds. Conversely, many species of songbirds (black-capped, boreal, and mountain chickadees, white-throated sparrows and zebra finches) have similar, accurate pitch height perceptual abilities (Weisman et al., 2010b). For example, even avian species as distantly related to songbirds as pigeons and budgerigars can learn frequency-range discriminations. Also, we have not found any songbird species that sings the same note in more than one octave. These facts lead us to reason in favor of the generality of chickadees’ pitch height and poor pitch chroma perception. We need to qualify this generalization given that the species of songbirds number approximately 5,000.

That humans have good octave equivalence of pitch chroma is supported by experimental evidence (Hoeschele et al., 2012) and of course humans perform and listen to music. In summary, we are suggesting that the disassociations between height and chroma perception likely have their origins hundreds of millions
of years ago in the formation of the mammalian and avian classes of species.

The charm of our hypotheses about phylogenetic causation is that, with few exceptions, they are easy to disprove. We have shown that diverse species groups of birds excel in pitch height categorization and that humans excel over songbirds in pitch chroma categorization. Clearly more cross-species comparisons would be useful. If a primate, a mouse, or a dolphin should fail to show classification by pitch chroma perception, or a songbird or any other bird species not subject to very special isolate adaptation fail to show high-quality classification by pitch height perception, our phylogenetic hypothesis would be falsified. We do hedge our bet slightly: although the pitch chroma and pitch height sorting should be conserved in mammals and birds respectively, some small number of species will be exceptions where because of surprising selection pressures these traits fail to be conserved as with other phylogenetic traits: for example, flightless birds and naked mammals are rare but they do exist.

Acknowledgements

All animal studies were conducted in accordance with the Canadian Council on Animal Care Guidelines and Policies with approval from the Animal Care and Use Committee for Biosciences for the University of Alberta and the University of Calgary Life and Environmental Sciences Animal Care Committee. Chickadees were captured and research was carried out under an Environment Canada Canadian Wildlife Service Scientific permit, Alberta Fish and Wildlife Capture and Research permits, and City of Edmonton Parks Permit. Human studies were approved by the Human Research Ethics Board of the University of Alberta of either Queen’s University. This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and Discovery Accelerator Supplement (DAS), an Alberta Ingenuity Fund (AIF) New Faculty Grant, a Canada Foundation for Innovation Infrastructure Operating Fund (JOF) and CFI New Opportunities Grant along with start-up funding and CFI partner funding from the University of Alberta to CBS. MH was supported by a NSERC postgraduate scholarship, an AIF graduate student scholarship, and an Izaak Walton Killam Memorial Scholarship (IWKMS) at the University of Alberta during the collection of the data summarized in this review and is currently funded by a European Research Council advanced grant [No. 230604 “SOMACCA”] awarded to W. Tecumseh Fitch at the University of Vienna.

References