

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

COGNITION

Cognition xx (xxxx) 1–11

[www.elsevier.com/locate/COGNIT](http://www.elsevier.com/locate/COGNIT)

Brief article

# Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate

Josh McDermott<sup>a,\*</sup>, Marc Hauser<sup>b</sup><sup>a</sup>*Perceptual Science Group, Department of Brain and Cognitive Sciences, MIT NE20-444, Cambridge, MA, USA*<sup>b</sup>*Department of Psychology, Harvard University, Cambridge, MA, USA*

Received 27 November 2003; accepted 29 April 2004

## Abstract

Humans find some sounds more pleasing than others; such preferences may underlie our enjoyment of music. To gain insight into the evolutionary origins of such preferences, we explored whether they are present in other animals. We designed a novel method to measure the spontaneous sound preferences of cotton-top tamarins, a species that has been extensively tested for other perceptual abilities. Animals were placed in a V-shaped maze, and their position within the maze controlled their auditory environment. One sound was played when they were in one branch of the maze, and a different sound for the opposite branch; no food was delivered during testing. We used the proportion of time spent in each branch as a measure of preference. The first two experiments were designed as tests of our method. In Experiment 1, we used loud and soft white noise as stimuli; all animals spent most of their time on the side with soft noise. In Experiment 2, tamarins spent more time on the side playing species-specific feeding chirps than on the side playing species-specific distress calls. Together, these two experiments suggest that the method is effective, providing a spontaneous measure of preference. In Experiment 3, however, subjects showed no preference for consonant over dissonant intervals. Finally, tamarins showed no preference in Experiment 4 for a screeching sound (comparable to fingernails on a blackboard) over amplitude-matched white noise. In contrast, humans showed clear preferences for the consonant intervals of Experiment 3 and the white noise of Experiment 4 using the same

\* Corresponding author.

*E-mail address:* [jhm@mit.edu](mailto:jhm@mit.edu) (J. McDermott).

0022-2860/\$ - see front matter © 2004 Published by Elsevier B.V.

doi:10.1016/j.cognition.2004.04.004

46 stimuli and a similar method. We conclude that tamarins' preferences differ qualitatively from  
47 those of humans. The preferences that support our capacity for music may, therefore, be unique  
48 among the primates, and could be music-specific adaptations.

49 © 2004 Published by Elsevier B.V.

50 *Keywords:* Music; Consonant; Tamarin

51

52

## 53 1. Introduction

54

55 Music is among the defining features of human culture, playing a central role in every  
56 society known to Western scholars. However, from the standpoint of evolution, music is  
57 also one of the most mysterious of human behaviors, as it serves no obvious function that  
58 might have driven its evolution. Evolutionary theorists since the time of Darwin have  
59 speculated about the adaptive function of music and its evolutionary origins (Darwin,  
60 1871), with little consensus or empirical support. Recently, however, work on infants and  
61 animals (reviewed in Hauser & McDermott, 2003; Trehub, 2003) has begun to illustrate  
62 how empirical evidence might shape theories of music's evolution. In particular, because  
63 animals can be tested in the absence of any exposure to music, parallel perceptual abilities  
64 in nonhuman animals can help establish whether aspects of our music faculty are innate  
65 and therefore candidate products of natural selection. Moreover, as nonhuman animals do  
66 not themselves make music, any perceptual effect found in a nonhuman animal cannot be  
67 part of an adaptation for music. Music-related experiments on animals are thus poised to  
68 play an important role in the debate about the origins of music.

69 One of the striking and mysterious features of how we experience music and other  
70 forms of art is the aesthetic response we often have to what we experience. As is the case  
71 for most aspects of music, the function and origins of aesthetic responses are unclear. As a  
72 first step in investigating these issues, we studied preferences for relatively simple sounds  
73 that lack the complex temporal structure of extended passages of music. Perhaps the best-  
74 known example of such acoustic preferences involves harmonic musical intervals. Some  
75 combinations of notes tend to sound good, at least to Western listeners, and are termed  
76 consonant; others sound bad and are termed dissonant (Dowling & Harwood, 1986;  
77 Krumhansl, 1990; Malmberg, 1918; Terhardt, 1984). Pythagoras was the first to note that  
78 consonance tends to be generated by pairs of tones whose fundamental frequencies are  
79 related by simple integer ratios. Helmholtz later proposed the widely accepted notion that  
80 peripheral auditory effects (namely, beating) distinguish consonance and dissonance, and  
81 both neurophysiological (Fishman et al., 2001; Tramo, Cariani, Delgutte, & Braidà, 2001)  
82 and behavioral (Hulse, Bernard, & Braaten, 1995; Izumi, 2000) studies in birds, cats, and  
83 primates suggest that these peripheral differences are shared across mammals and birds.  
84 These peripheral effects account for the discriminability of consonant and dissonant  
85 intervals, but shed little light on the preferences between the two classes of stimuli that are  
86 arguably the main reason for their importance in music. Where do such preferences come  
87 from? Are they acquired through exposure to music, which perhaps contains more  
88 consonant intervals than dissonant ones? Are they part of an adaptation to music? Or might  
89 they be a byproduct of some general feature of the auditory system? Experiments in  
90 closely related animals, especially primates, can help to clarify these issues.

## 91 2. Experiment 1: intensity

92

93 To test whether humans share any acoustic preferences with other primates, we first  
94 developed a method to measure such preferences in a well-studied nonhuman primate: the  
95 cotton-top tamarin (see [Watanabe & Nemoto, 1998](#) for a related method developed for use in  
96 birds). In Experiment 1, we compared a low amplitude white noise signal to a high amplitude  
97 white noise signal. We expected the animals to find the high amplitude signal unpleasant.  
98

### 99 2.1. Method

100

#### 101 2.1.1. Participants

102 We tested 6 adult cotton-top tamarins (*Saguinus oedipus*), 3 males and 3 females.  
103

#### 104 2.1.2. Apparatus and procedure

105 We tested subjects in a V-shaped maze ([Fig. 1](#)). A concealed speaker (Advent Powered  
106 Partners) was located at the end of each branch of the maze. Each speaker and branch was  
107 paired with a different sound. Subjects were initially moved from their home room cage to  
108 the test room, and then placed at the entrance to the maze. The experimenter then left the  
109 room and raised the door to the maze by means of a pulley system, thereby allowing the  
110 tamarin to enter. When the subject moved into a branch for the first time, an experimenter  
111 started the playback. The stimulus for a particular side played continuously as long as the  
112 animal was on that side, and switched as soon as they switched sides. The animal's  
113 position in the maze thus determined which sound they heard. Testing continued for 5 min,  
114 during which subjects were videotaped. No food reward was given.  
115

#### 116 2.1.3. Stimuli and design

117 The amplitudes of the two white noise signals were 60 and 90 dB, respectively, when  
118 measured with a sound meter at the center point of the maze. The stimuli were randomly  
119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

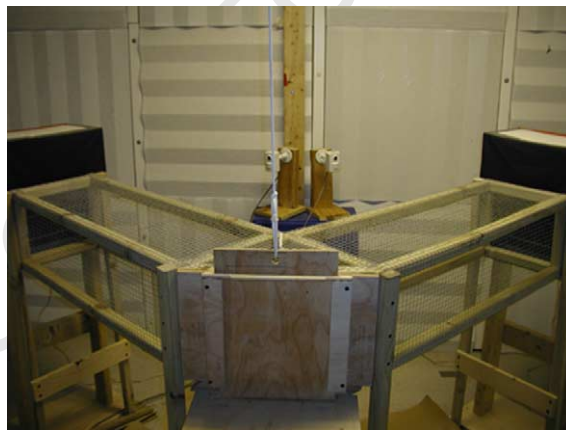


Fig. 1. Photo of the apparatus used in the tamarin experiments. The maze was elevated off the floor. There was a concealed speaker at the end of each branch of the maze.

136 assigned to the sides of the apparatus for each animal. After two sessions of this condition,  
 137 each separated by a full day, the sound-side pairing was reversed, and the animals were run  
 138 for two more sessions, again on separate days. A bias to spend more time in one branch  
 139 than the other was taken as evidence for a preference for one sound over the other.

140 The experimenters and trained assistants coded the video recordings with the sound  
 141 turned off and without knowledge of the side assignment. The video displayed the time of  
 142 recording down to a second's resolution. To code an experimental session the coder noted  
 143 each time at which the animal moved from one side of the apparatus to the other. From these  
 144 times the length of each excursion to one side or the other could be computed, and these were  
 145 then added to yield the total amount of time spent on each side during an experimental  
 146 session. Inter-observer reliability was high; over 10 sessions coded by two observers, the  
 147 correlation coefficient for the switch times noted by two different coders was 0.99.

148

149

150

## 2.2. Results and discussion

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

Fig. 2 plots the time spent on each side of the maze, averaged across the 6 monkeys, in each of 4 successive experimental sessions. Even in the first session there is a pronounced tendency to spend more time on the side playing the low amplitude white noise, a tendency that increases during the second session. After the first two sessions, the sound-side assignments were swapped for each animal, and on average the animals spent equal amounts of time on each side, suggesting that they had learned an association between one side and the low amplitude noise. By the next session they regained the tendency to spend more time on the side with lower amplitude noise. Across all four sessions the animals averaged 70% of the time on the soft side, which was highly significant ( $t[23]=5.5$ ,  $p<0.00001$ ). In a second experiment, we modified the noise amplitudes so that there was

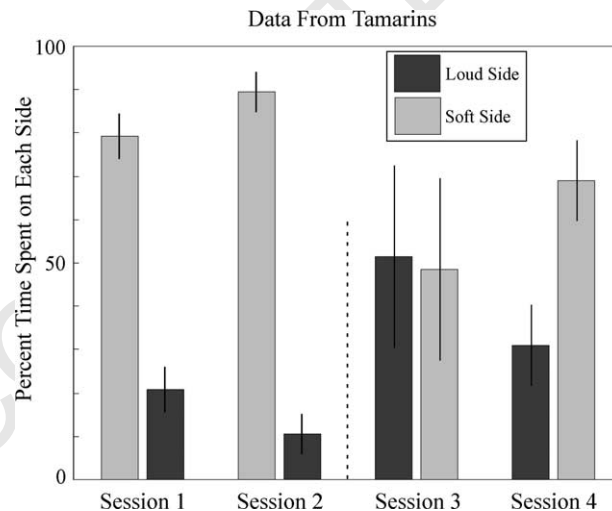


Fig. 2. Results of Experiment 1, comparing high and low amplitude white noise. Each bar plots the average data from 6 subjects, as a proportion of the total time spent in the apparatus. Error bars here and elsewhere denote standard errors. The dashed line denotes reversal of the side assignment that occurred after the second session.

181 only a 10 dB difference between the two sides (75 and 85 dB, respectively). All six  
182 animals again spent more time on the side with the lower amplitude noise (68%,  
183 SE=4.5%) over 2 sessions ( $t[11]=4.09$ ,  $p<0.001$ ).

184 These results, together with those of Watanabe and Nemoto (1998) suggest that our  
185 method provides one way to assess spontaneous acoustic preferences in animals,  
186 especially for stimuli other than their species-specific vocalizations (for a related  
187 technique used to study such vocalizations, see Gerhardt, 1987; Miller, Dibble, et al.,  
188 2001; Miller, Miller, et al., 2001; Ryan, 1980; Wilczynski, Rand, & Ryan, 1995).

189

190

### 191 **3. Experiment 2: distress calls**

192

193 To provide a second verification of our method and extend its ecological validity, we  
194 ran the tamarins on an experiment contrasting two species-specific vocalizations—  
195 screams given during distress and chirps made during feeding. Given the negative  
196 associations of the screams and the positive associations of the food-related chirps, we  
197 predicted that the tamarins would spend more time on the side with food chirps than on the  
198 side with distress screams.

199

#### 200 *3.1. Method*

201

##### 202 *3.1.1. Participants*

203 We tested 5 of the 6 cotton-top tamarins used in Experiment 1.

204

##### 205 *3.1.2. Apparatus and procedure*

206 The apparatus and procedure of Experiment 1 were used again.

207

##### 208 *3.1.3. Stimuli and design*

209 The distress calls were screams produced by animals being held by our veterinary staff  
210 during routine checkups. The food chirps were produced by individuals while eating food  
211 or just as food was presented. Files were created with six exemplars of either the screams  
212 or the chirps separated by brief periods of silence (a variable period between 1 and 1.5 s).  
213 The vocalizations were recorded from three different animals that were not run in the  
214 experiment. Two screams and two chirps from each of the three animals were used for the  
215 stimuli. The six screams or chirps looped continuously during playback.

216 This experiment was run approximately 3 months after the conclusion of Experiments 1,  
217 2 and 4. Each animal was run in 3–4 sessions with a particular side assignment  
218 (determined at random), followed by 2–4 sessions with the side assignment reversed. The  
219 video recordings were coded as in Experiment 1.

220

#### 221 *3.2. Results and discussion*

222

223 Over the course of several sessions ( $N=41$  total across the 5 animals), subjects showed  
224 a statistically significant ( $t[40]=2.53$ ;  $p<0.01$ ; Fig. 3) preference for the side with food  
225 chirps over the side with screams.

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

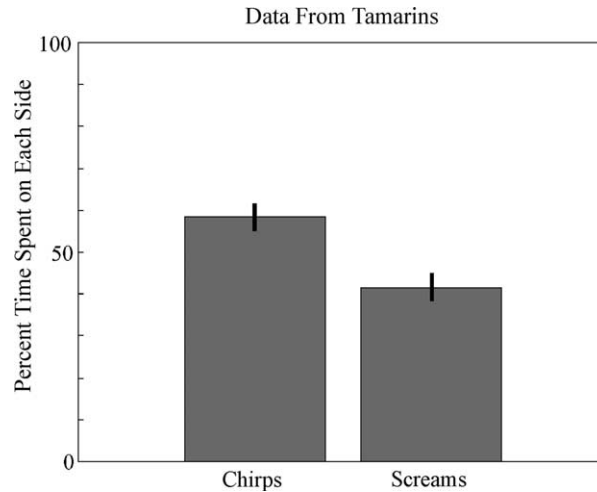


Fig. 3. Results from Experiment 2, comparing food chirps with distress screams.

This result provides further evidence that our method is appropriate for assaying spontaneous sound preferences in tamarins, and shows that these animals can have preferences for a range of stimuli, be they species-specific or artificial.

#### 4. Experiment 3: consonance

As our primary interests are centered on the origins of musical preferences, we began by testing tamarins for preferences for consonant stimuli over dissonant stimuli. Although humans, at least in Western cultures, tend to show a preference for consonant sounds, we ran adult humans on an analogue of the tamarin experiment to ascertain whether the method would translate to another species known to show the preference.

##### 4.1. Method

###### 4.1.1. Participants

The participants were the 5 cotton-top tamarins used in Experiment 2, along with 5 Harvard undergraduates (18–21 years old; 1 male, 4 female). All 5 human subjects had some degree of musical training, ranging from one to many years of music lessons.

###### 4.1.2. Apparatus and procedure

The apparatus and procedure of Experiment 1 were used again with the tamarin subjects. The human subjects were placed in a room divided in half by a stripe taped to the floor. The front wall of the room concealed two speakers, one on each side of the dividing line. Each speaker played a particular sound when the subject was in the corresponding half of the room, thereby mimicking the tamarin setup.

271 The human subjects were told only that they had to stay within the confines of the room  
272 for the designated period of 5 min. No other instructions were given. All the human  
273 subjects were naïve as to the purpose of the experiment.

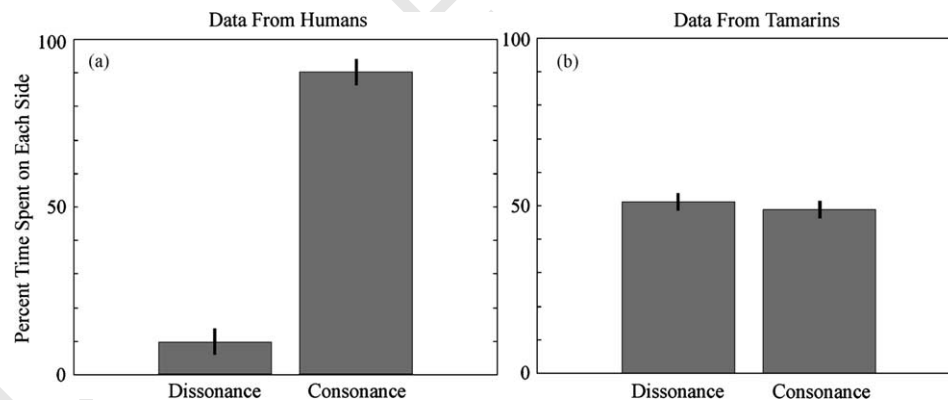
#### 275 4.1.3. Stimuli and design

276 The consonant stimulus consisted of a sequence of two-note chords chosen randomly  
277 from the set of the octave, the fifth, and the fourth, subject to the constraint that no  
278 particular interval repeated more than once in a row. The dissonant stimulus was a  
279 similarly constructed sequence of minor seconds, tritones, and minor ninths. Each note  
280 composing the interval was a synthesized complex tone with ten harmonics. The bass note  
281 of each interval was middle C. Each interval was 1.5 s in duration and was ramped on and  
282 off over 100 ms. There was no gap between successive intervals, and the sequence of  
283 intervals played continuously as long as the subject was on the corresponding side,  
284 switching when they switched sides. The consonant and dissonant stimuli had equal  
285 amplitudes, which were set such that the sound level measured at the center of the  
286 apparatus/room was 80 dB.

287 As in Experiment 1, subjects were left in the apparatus for 5 min, during which they  
288 were free to move within its confines. The human subjects were run in a single session,  
289 while the tamarins were run repeatedly, up to 10 sessions in a row in some cases, as we  
290 wanted to maximize the chances of revealing an effect. The video recordings were coded  
291 as in Experiment 1.

#### 293 4.2. Results and discussion

295 One of the human subjects stood in the same place for the entire experiment, and his  
296 data were thrown out. The average results for the other four human subjects are plotted in  
297 Fig. 4a. Human subjects spent most of their time on the consonant side of the room  
298 ( $t[3] = 10.26$ ;  $p < 0.001$ ); this pattern was consistent across subjects (Mann–Whitney test,  
299  $U = 2.31$ ,  $p < 0.02$ ). In contrast, the tamarins showed no preference, spending



314 Fig. 4. Results of Experiment 3, comparing consonant and dissonant musical intervals. (a) Results for human  
315 subjects. (b) Results for tamarin subjects.

316 approximately equal amounts of time on each side of the maze (Fig. 4b;  $t[30] = 0.47$ ;  $p =$   
317 0.32). The lack of preference is not due to habituation to our test apparatus, as all five  
318 animals showed robust preferences for low over high amplitude white noise when tested  
319 again at the conclusion of the experiment (1 session per animal, mean of 70% of time spent  
320 on low amplitude side). We conclude that under these particular test conditions, tamarins  
321 do not show a spontaneous preference for consonance over dissonance, differing notably  
322 from human adults tested with a similar paradigm.

323

324

325

## 5. Experiment 4: screeching

326

327

328

329

330

331

332

333

334

335

336

337

### 5.1. Method

338

339

340

#### 5.1.1. Participants

341

342

We tested 5 new adult cotton-top tamarins and 4 of the 5 Harvard undergraduates used in Experiment 3.

343

344

#### 5.1.2. Apparatus and procedure

345

The apparatus and procedure of Experiment 3 were used again.

346

347

#### 5.1.3. Stimuli and design

348

349

350

351

352

353

354

355

356

357

358

359

360

To facilitate stimulus generation, we used a variant of the fingernails-on-a-blackboard sound produced by scraping a three-pronged metal garden tool down a pane of glass (Halpern et al., 1986). Informal tests showed that stimuli produced the desired response in humans, suggesting it would produce a pronounced preference for a suitable comparison stimulus. The acoustic structure of our screeches was similar to that previously reported; there were typically several prominent harmonics overlaid with broadband noise. The experimental stimulus consisted of several concatenated recordings of individual screeches. As a comparison stimulus we generated white noise with the amplitude envelope of the screech stimulus. Both files looped continuously as long as a subject remained on the corresponding side of the apparatus. Subjects were again left to move freely in the apparatus for 5 min, during which they were videotaped. The amplitude of both stimuli was set to 80 dB as measured at the center point of the apparatus. The video recordings were coded as in Experiment 1.

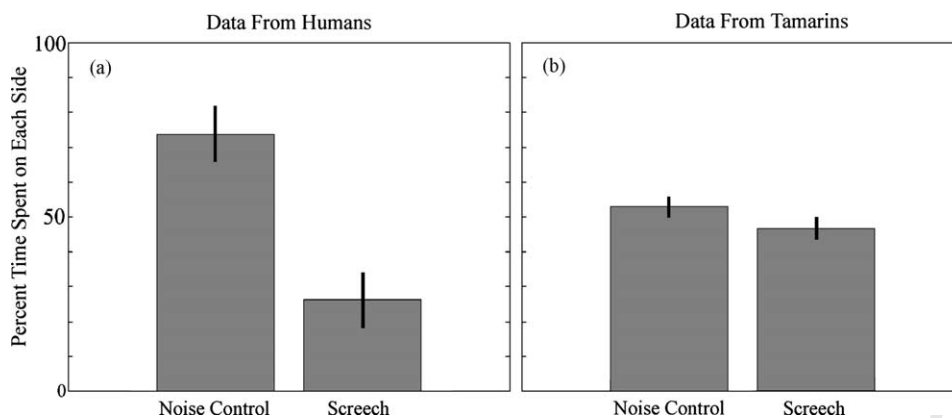


Fig. 5. Results of Experiment 4, comparing a screeching sound with an amplitude-matched noise control stimulus. (a) Results for human subjects. (b) Results for tamarin subjects.

## 5.2. Results and discussion

Fig. 5a shows the proportion of time humans spent on each side of the test room. As expected, there was a pronounced preference for the white noise ( $t[3]=2.94$ ,  $p<0.05$ ); a Mann–Whitney test revealed that all subjects followed this pattern ( $U=2.31$ ,  $p<0.02$ ). Because the tamarins used in this experiment had not been run in Experiments 1 and 2, we first ran all 5 tamarins in a replication of Experiment 1. All of the tamarins spent more time on the side of the maze with the low amplitude noise, and this tendency reversed itself when the side assignments were reversed, as expected (66% of time on soft side;  $SE=2.17\%$ ;  $t[37]=7.24$ ;  $p<0.00001$ ). When tested on the screech and control stimuli, however, the tamarins showed no evidence of a preference. We ran the tamarins for several consecutive sessions ( $N=37$  sessions) to see if a preference would emerge over time. As shown in Fig. 5b, there was no preference ( $t[36]=0.89$ ;  $p=0.15$ ). In contrast with humans, who show a pronounced preference for white noise over the screeching sound, tamarins do not exhibit a preference.

## 6. Conclusions

Preferences for consonance over dissonance are widespread in human adults (Dowling & Harwood, 1986) and have also been demonstrated in human infants (Trainor & Heinmiller, 1998; Zentner & Kagan, 1996, 1998). Our results suggest that although such preferences may be innate in humans, they likely have evolved after the divergence point with our primate cousins. It is of course possible that another primate species, more closely related to humans (e.g. chimpanzees), might exhibit more similar acoustic preferences, or that tamarins tested with a different procedure would show a preference. It is also worth noting that Watanabe and Nemoto (1998) recently found that certain Java sparrows showed preferences for some types of music over others. This preference could

406 conceivably be related to the singing behavior of this species, and it would be interesting to  
407 test them with the consonant and dissonant stimuli that we used. Given the present results,  
408 however, we conclude that if humans and nonhuman primates share acoustic preferences  
409 for sounds, this capacity evolved more recently than the divergence with New World  
410 monkeys such as the cotton-top tamarin (i.e. some 40 million years ago). This conclusion  
411 stands in contrast to the many perceptual mechanisms shared between humans and  
412 tamarins (and other species as well), particularly with respect to speech perception (Miller,  
413 Dibble, et al., 2001; Miller, Miller, et al., 2001; Newport, Hauser, Spaepen, & Aslin, 2004;  
414 Ramus, Hauser, Miller, Morris, & Mehler, 2000) and presumably also to the  
415 discriminability of consonance and dissonance. This contrast raises the possibility that  
416 some of the acoustic preferences observed in humans evolved as a specific adaptation for  
417 music.

418  
419

## 420 Acknowledgements

421

422 We thank Altay Guvench and Matt Kamen for help building the apparatus, and Altay  
423 Guvench, Matt Kamen, Fernando Vera, Adam Pearson, Tory Wobber, Matthew Sussman,  
424 and Alex Rosati for help running the experiments. MDH acknowledges the NIMH and the  
425 McDonnell foundation for funding this project.

426  
427

## 428 References

429

- 430 Dowling, J., & Harwood, D. L. (1986). *Music cognition*. San Diego: Academic Press.
- 431 Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., Howard, M. A., &  
432 Steinschneider, M. (2001). Consonance and dissonance of musical chords: Neuronal in auditory cortex of  
433 monkeys and humans. *Journal of Neurophysiology*, *86*, 271–278.
- 434 Gerhardt, H. C. (1987). Evolutionary and neurobiological implications of selective phonotaxis in the green  
435 treefrog (*Hyla cinerea*). *Animal Behaviour*, *35*, 1479–1489.
- 436 Halpern, D. L., Blake, R., & Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Perception &*  
437 *Psychophysics*, *39*, 77–80.
- 438 Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: Comparative perspectives. *Nature*  
439 *Neuroscience*, *6*, 663–668.
- 440 Hulse, S. H., Bernard, D. J., & Braaten, R. F. (1995). Auditory discrimination of chord-based spectral structures  
441 by European starlings (*Sturnus vulgaris*). *Journal of Experimental Psychology: General*, *124*, 409–423.
- 442 Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of*  
443 *America*, *108*, 3073–3078.
- 444 Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- 445 Malmberg, C. F. (1918). The perception of consonance and dissonance. *Psychological Monographs*, *25*, 93–133.
- 446 Miller, C. T., Dibble, E., & Hauser, M. D. (2001). Amodal completion of acoustic signals by a nonhuman primate.  
447 *Nature Neuroscience*, *4*, 783–784.
- 448 Miller, C. T., Miller, J., Gil de Costa, R., & Hauser, M. D. (2001). Selective phonotaxis by cotton-top tamarins.  
449 *Behaviour*, *138*, 811–826.
- 450 Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance. II. Statistical learning  
of non-adjacent dependencies in a nonhuman primate. *Cognitive Psychology* (in press).
- 451 Ramus, F., Hauser, M. D., Miller, C. T., Morris, D., & Mehler, J. (2000). Language discrimination by human  
newborns and cotton-top tamarin monkeys. *Science*, *288*, 349–351.
- 452 Ryan, M. J. (1980). Female mate choice in a Neotropical frog. *Science*, *209*, 523–525.

- 451 Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music*  
452 *Perception*, 1, 276–295.
- 453 Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to  
454 listen to consonance over dissonance. *Infant Behavior and Development*, 21, 77–88.
- 454 Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of  
455 harmony in western tonal music. *Annals of the New York Academy of Sciences*, 930, 92–116.
- 456 Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, 6, 669–673.
- 457 Watanabe, S., & Nemoto, M. (1998). Reinforcing properties of music in Java sparrows. *Behavioral Processes*, 43,  
458 211–218.
- 458 Wilczynski, W., Rand, A. S., & Ryan, M. J. (1995). The processing of spectral cues by the call analysis system of  
459 the tungara frog, *Physalaemus pustulosus*. *Animal Behavior*, 49, 911–929.
- 460 Zentner, M. R., & Kagan, J. (1996). Perception of music by infants. *Nature*, 383, 29.
- 461 Zentner, M. R., & Kagan, J. (1998). Infants perception of consonance and dissonance in music. *Infant Behavior*  
462 *and Development*, 21, 483–492.
- 463
- 464
- 465
- 466
- 467
- 468
- 469
- 470
- 471
- 472
- 473
- 474
- 475
- 476
- 477
- 478
- 479
- 480
- 481
- 482
- 483
- 484
- 485
- 486
- 487
- 488
- 489
- 490
- 491
- 492
- 493
- 494
- 495