



HAL
open science

Crossmodal pitch-luminance correspondence in human and non-human primates: a failed replication of Ludwig et al. (2011) in baboons (*Papio papio*)

Konstantina Margiotoudi, Joel Fagot, Adrien Meguerditchian, Isabelle Dautriche

► To cite this version:

Konstantina Margiotoudi, Joel Fagot, Adrien Meguerditchian, Isabelle Dautriche. Crossmodal pitch-luminance correspondence in human and non-human primates: a failed replication of Ludwig et al. (2011) in baboons (*Papio papio*). 2023. hal-04310843

HAL Id: hal-04310843

<https://hal.science/hal-04310843v1>

Preprint submitted on 27 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Crossmodal pitch-luminance correspondence in human and non-human primates: a failed replication of Ludwig et al. (2011) in baboons (*Papio papio*)

Konstantina Margiotoudi^{1,2}, Joel Fagot^{1,2}, Adrien Meguerditchian^{1,2} and Isabelle Dautriche^{1*}

¹ Laboratoire de Psychologie Cognitive, UMR7290, LPC, CNRS, Aix-Marseille Université, Marseille, France

² Station de Primatologie-Celphedia UAR846, CNRS, Rousset, France

* corresponding author: isabelle.dautriche@cns.fr

35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68

Abstract

Humans spontaneously and consistently map information coming from different sensory modalities. Surprisingly, the phylogenetic origin of such cross-modal correspondences has been under-investigated. A notable exception is the study of Ludwig et al (2011) which reports that both humans and chimpanzees spontaneously map high-pitched sounds to bright objects and low-pitched sounds to dark objects. Our pre-registered study aimed to directly replicate this research on both humans and baboons (*Papio papio*), an old world monkey which is more phylogenetically distant from humans than chimpanzees. Following Ludwig et al., participants were presented with a visual classification task where they had to sort black and white square (low and high luminance), while background sounds (low or high-pitched tones) were playing. Whereas we replicated the finding that humans' performance on the visual task was affected by congruency between sound and luminance of the target, we did not find any of those effects on baboons' performance. These results question the presence of a shared iconic pitch-luminance mapping in other non-human primates.

Keywords: crossmodal correspondence; comparative cognition; iconicity

69 Crossmodal pitch-luminance correspondence in human and non-human primates: a
70 failed replication of Ludwig et al. (2011) in baboons (*Papio papio*)

71

72 Humans tend to form automatically cross-modal correspondences, i.e., associations
73 between modality-specific features [1]. For instance, they associate certain sound
74 frequency to shape properties: a high-pitched tone will be preferably associated to
75 objects that are small, spiky and bright, whereas a low-pitched tone would be generally
76 associated to objects that are big, round, and dark [2]. Such cross-modal biases have
77 been demonstrated experimentally across a wide range of modalities: between odors
78 and visual stimuli [3], sound and taste [4], and taste and touch [5], among many others
79 [6].

80

81 Cross-modal biases have been proposed as a key mechanism underlying
82 multisensory integration and language evolution. In perception, cross-modal
83 correspondences may help to constrain the cross-modal binding problem (i.e., the
84 problem of knowing which of many perceptual information should be bound together)
85 [7]. In language: cross-modal biases to associate sounds with other sensory
86 modalities may have triggered the evolution of spoken language by facilitating the
87 expression of sensory experiences through a unique auditory channel [8,9].
88 Surprisingly though, despite the potential evolutionary importance of cross-modal
89 correspondences, the investigation of their phylogenetic origins remains very limited.

90

91 One possibility is that cross-modal correspondences may be limited in non-human
92 animals because they are rooted in linguistic experience. For instance, the pitch-
93 luminance mapping [10] –where higher frequencies are associated to bright stimuli –
94 might derive from the usage in English of the same adjectives to refer to luminance
95 and pitch, namely ‘high’ and ‘low’. Alternatively, because non-human species would
96 similarly benefit from cross-modal correspondences to facilitate the integration of
97 multisensory inputs coming from their environment [11], these proclivities may be the
98 result of a natural, structural, feature of our perceptual system that may be shared
99 across species. For example, for the pitch-luminance mapping, both visual high
100 luminance and acoustic high frequency have a shared increase in stimulus intensity
101 which is suggested to be encoded by the same brain structure dedicated to magnitude
102 [12].

103

104 Comparative studies provide an immediate way to tease apart these different
105 possibilities. To our knowledge, only few studies tested the existence of these
106 mappings in other animals, and report evidence for a pitch-elevation mapping [13] –
107 where higher frequencies are associated to high spatial positions-- , and a pitch-size
108 mapping in domestic dogs [14], a space-luminance mapping [15] –where right
109 positions are associated to brighter stimuli-- but not a pitch-luminance mapping in
110 chicks [16], and finally a pitch-luminance mapping in chimpanzee [10]. Some of these
111 correspondences could be learn from their co-occurring regularities in the environment.
112 The pitch-elevation correspondence might be explained by environmental statistics of
113 auditory scenes as higher frequency sounds originate from elevated sources [5].
114 Similarly the pitch-size correspondence might arise from the observation that large
115 objects tend to produce low frequency sounds (e.g., when dropped) while small
116 objects produce high-pitched sounds [17]. Yet, evidence from the pitch-luminance and
117 the space-luminance mappings in non-human animals would suggest that some of
118 these mappings have a non-linguistic, structural origin.

119

120 Of special interest is pitch-luminance correspondence, which is well established in
121 humans [18-20], that triggered different findings in chicks and chimpanzees. Critically,
122 this is the only study reporting the existence of a cross-modal correspondence that
123 cannot be extracted statistically from the environment, in a non-human primate
124 species. In that study, 6 chimpanzees and 33 human participants were trained in a
125 visual speeded classification task on a touch screen, where they had to match a
126 sample square (black or white) to the matching stimulus (i.e., identical black or white
127 square). During testing, both species performed the same task while passively
128 listening to background sounds (i.e., high and low-pitched tones), which were
129 crossmodally either congruent or incongruent to the luminance of the sample square.
130 Both species showed better performance when the frequency of the sound matched
131 the luminance of the sample square – white square congruent with the high-pitched
132 tone and black square with the low-pitched tone.

133

134 Other attempts on testing sound-shape mappings (i.e., mapping the ‘round’ vs ‘sharp’
135 sounding nonwords, such as ‘bouba’ and ‘kiki’, to a curved and ‘sharp’ shapes,
136 respectively [21, 22]) in great apes (i.e., gorillas, chimpanzees and bonobos) failed to

137 show any cross-modal effects in these species. Based on the weak experimental
138 evidence on cross-modal mapping in non-human primates, in the current pre-
139 registered study, we aimed to carry out a direct replication of Ludwig et al. (2011)'s
140 study on the pitch-luminance correspondence, by testing a group of humans and
141 baboons (*Papio papio*) following the original methods and analyses as closely as
142 possible. Baboons and humans were trained on a visual speeded classification task,
143 where they had to classify squares based on their luminance. During the testing phase,
144 both species repeated the same speeded classification task, while background sounds
145 (e.g., low vs high pitched tones) were presented. We examined whether sound
146 congruency, following a pitch-luminance mapping, would affect the performance of
147 both species on the visual classification task.

148

149 Methods

150

151 The methods, design and analysis were pre-registered at: <https://osf.io/vsnp4>.

152 All the data and stimulus material are openly available at:
153 https://osf.io/ng6y3/?view_only=b3797de197744bb79a1bcc544f0cf69a.

154 All methods for the human study were approved by the Ethics Committee of Aix-
155 Marseille University and for the baboon study received approval from the French
156 Ministère de l'Education Nationale (approval no. APAFIS-2717-2015111708173794-
157 V3).

158

159 Human testing: Twenty-seven, right-handed human adults (21 females, age $M=22.70$
160 ± 5.55) participated in the study. The task was a visual speeded classification task
161 involving the matching of a sample square (black or white) to one of the two
162 alternatives which is identical to the sample (identity matching-to-sample procedure).
163 During the test trials, a tone, either high or low-pitched, was played while participants
164 were performing the same task. The design was identical to the one described in
165 Ludwig et al. (2011) [10], with the sole difference on the feedback screens (see
166 supplementary Figure S1.b). In the present study, we used visual feedback, instead
167 of audio feedback as in Ludwig et al., 2011, since that group of baboons was already
168 used to be trained with visual feedback when performing touch-screen tasks [22]. After
169 a correct response, a black screen appeared for 3s combined with food reward only

170 for the baboons, whereas after an incorrect response a green screen appeared for 3
171 s plus a black screen of 3s.

172

173 Before starting with the task, participants received written instructions and were asked
174 to be as fast and accurate as possible in performing the visual speeded classification
175 task. The experiment was divided into two phases, a training and a testing phase.
176 During the training, participants performed 16 matching-to-sample trials without any
177 background sound (for details see supplementary material). By the end of this phase,
178 participants had to reach a performance of 100%, in order to continue to the testing
179 phase. These 16 trials were repeated once in case of mistakes. Two participants did
180 not reach a 100% performance at the first training and repeated the procedure. Once
181 the training was completed, the testing phase followed. Participants received written
182 instructions to focus on the visual speeded classification task, even though sounds
183 would be played in the background. The testing phase started with 4 trials including
184 sounds, which were excluded from the analysis. Then, two blocks of 80 trials each
185 (160 trials in total) followed, with a break in-between. Humans were tested in one
186 single session of 30 minutes.

187

188 Baboon testing: Fifteen Guinea baboons (*Papio papio*), (12 females, age $M=11.4 \pm$
189 5.64) completed the study. Baboons were tested at the CNRS Primatology center in
190 Rousset-sur-Arc, France where they live in two social groups and had *ad libitum*
191 access to 13 ALDMs (Automated Computer Learning devices) [23]. Each device has
192 a food dispenser, a touchscreen and two speakers. The speakers are facing down
193 inside each test chamber. A radio-frequency identification of the subjects by the
194 system facilitated testing, as it doesn't require capturing of the subjects [24]. The
195 baboons have voluntary access to all the devices. Baboons performed the same visual
196 speeded classification task as humans, and were exposed to the same training
197 procedure (for details on the training see supplementary material). All baboons
198 completed the testing in three consecutive days. Each day comprised a testing
199 session. During a session, the monkeys completed a block of 16 trials without any
200 background noise, where they had to achieve a performance of at least 80% in both
201 categories (i.e., black/white) (White: $M=0.80\%$; $SE=0.02\%$, Black: $M=0.88\%$;
202 $SE=0.02$). The block was repeated until the criterion was reached (number of
203 averaged training blocks $=2.71$; $min=1$; $max=11$; $SE=0.31$). After that, baboons

204 performed 4 practice trials with background sounds. Then, a block of 160 testing trials
205 followed. In total, baboons completed 480 testing trials. All trials and conditions
206 between samples and sounds were equally counterbalanced, following the description
207 by [10] (for details on the training see supplementary material).

208

209 Results

210

211 We conducted the analysis described by Ludwig et al., (2011) [10]. We first looked at
212 the mean median RT and the error rates for congruent and incongruent trials in both
213 species. Our results confirm the findings of Ludwig et al. (2011) [10] on humans but
214 not on non-human primates. Specifically, in the original study, chimpanzees made
215 more errors in incongruent than in congruent trials, while, in baboons, no differences
216 were observed on the mean error rates between congruent ($M = 10.58\%$, $SD =$
217 3.16%) and incongruent trials ($M = 10.58\%$, $SD = 4.16\%$). Yet, similarly to the original
218 study, there was also no difference on the mean median RTs between congruent (M
219 $= 653\text{ms}$, $SD = 97.6\text{ms}$) and incongruent trials ($M = 657\text{ms}$, $SD = 102.4\text{ms}$; $t(14) =$
220 -1.31 , $p = 0.20$, $r = 0.33$).

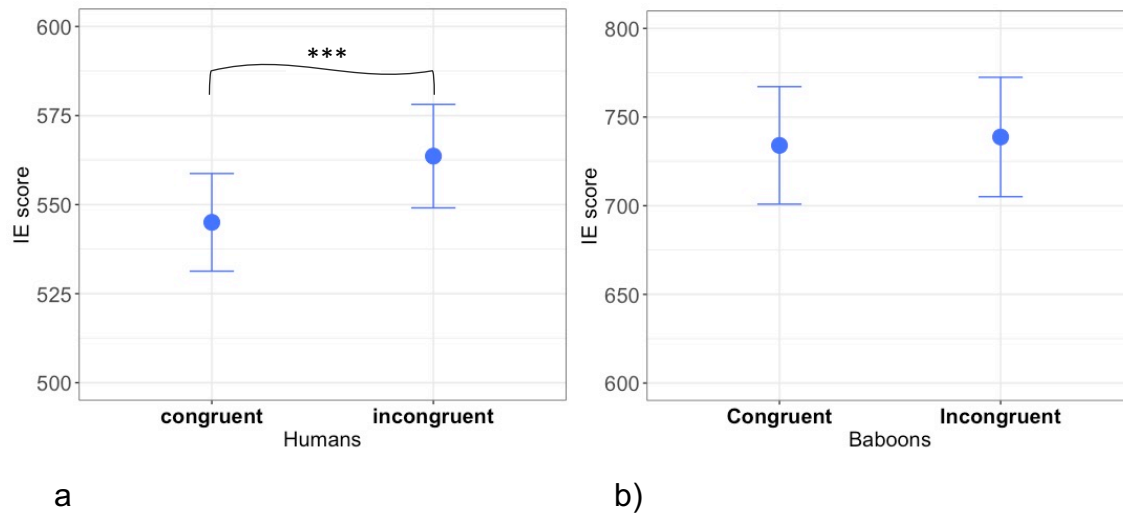
221 Humans had no significant difference in the mean error rates between congruent (M
222 $= 0.18\%$, $SD = 0.04\%$) and incongruent trials ($M = 0.46\%$, $SD = 0.06\%$; $t(26) =$
223 -0.94 , $p = 0.35$, $r = 0.18$). Yet, they were faster for congruent trials ($M = 544\text{ms}$; $SD = 71.8\text{ms}$)
224 than for incongruent trials ($M = 560\text{ms}$, $SD = 69.8\text{ms}$; $t(26) = -4.88$, $p < 0.001$, $r = 0.69$).
225 These findings parallel the results reported in humans by Ludwig et al. (2011).

226

227 Following Ludwig et al., we then calculated the speed-accuracy trade-off for congruent
228 and incongruent trials in both species, by calculating the Inverse Efficiency score (IE)
229 [25]. To calculate the IE scores, for each participant we divided the median RT by the
230 proportion correct, for the two conditions. A paired sample t-test showed that humans
231 had a significantly lower IE score (i.e., indicating better performance) ($t(26) = -4.15$,
232 $p < 0.001$; $r = 0.69$) for congruent (IE = 545, $SD = 71.2$) than for incongruent trials (IE =
233 564, $SD = 75.5$;). In contrast to the findings of Ludwig et al. (2011) on chimpanzees,
234 the results on the baboons revealed no significant difference ($t(14) = -0.53$, $p > 0.10$) on
235 the IE score between congruent (mean IE: 734, $SD = 128$) and incongruent trials
236 (mean IE: 738, $SD = 130$) (see Fig.1). Based on these findings, we replicated the
237 findings of Ludwig et al. (2011) [10] in humans on pitch-luminance congruency

238 crossmodal effects, but not in baboons (for details see results in supplementary
239 material).

240



241

242

243

244 Figure 1. Mean IE scores for congruent and incongruent trials in a) humans and b)
245 baboons. Blue circles and errors bars represent the means and standard errors for
246 each condition (SE). Asterisks indicate significant differences between the two
247 conditions ($p < 0.001$).

248

249 Discussion

250

251 Following the protocol of Ludwig et al. (2011) [10], we tested the pitch-luminance
252 crossmodal correspondence in a group of humans and baboons. In the original study,
253 humans showed longer reaction times when the sound was incongruent to the
254 luminance of the sample stimulus, whereas in chimpanzees an interference effect was
255 observed only on the proportion of error rates. Chimpanzees did more errors for
256 incongruent than congruent trials. In the present replication study, although we
257 replicated the human results, our results did not show any evidence of sound
258 interference effects on the visual speeded classification task in baboons.

259

260 How to account for the difference between chimpanzees [10] and the current baboons
261 result? Certainly, primate species differences might suggest that the mapping of pitch-
262 to-luminance is unique to chimpanzees and humans, but does not extend to more
263 phylogenetically distant relatives such as the Old World monkeys, including the

264 baboons. According to this hypothesis, such a cross-modal correspondence might
265 thus have emerged and be inherited from the last common ancestor of the Hominini
266 tribe. Yet, another possibility is that the absence of any systemic pitch-to-luminance
267 correspondence in baboons, questions the existence of this mapping not only in this
268 species but also in any other nonhuman species. There are several reasons to doubt
269 on the robustness of the chimpanzee's result. First, while the study of Ludwig et al.
270 (2011) showed a pitch interference effect in a visual speeded classification task in
271 chimpanzees, this effect was based on a 2.7% difference between mean error rates
272 for congruent and incongruent trials but was not visible on median response times,
273 contrary to human participants. Second, the size of the chimpanzee sample was
274 relatively low (N=6) compared to the present baboon study (N=15). Third, whether the
275 chimpanzees tested in the Ludwig et al., (2011) [10] had previous experiences on
276 auditory discrimination between high- and low-pitched tones, is not specified in the
277 original study. Any such previous experience could have increased sensitivity to the
278 auditory stimuli of the present paradigm. This is of particular importance as testing
279 non-human primates on auditory paradigms is not a trivial task [26, 27].

280

281 Our results are consistent with previous results showing no evidence of crossmodal
282 mappings in non-human primates [21,22] or in other animals such as chicks [16]. This
283 suggests that some crossmodal correspondences might not be shared with other
284 nonhuman species. As we noted, the absence of any crossmodal effects in these
285 species could have plausibly resulted from attention or discriminatory troubles on
286 auditory stimuli. In addition, factors such as previous habituation to visual tasks,
287 training history, protocol design, or age of the individuals appear to play an important
288 role in testing the performance of nonhuman primates on tasks including auditory
289 stimuli [28].

290

291 To conclude, evidence of the presence of a pitch-luminance mapping in other
292 nonhuman species appears to be weak, suggesting that language, culture or human-
293 specific neuro-anatomical features may be fundamental for the emergence of this non-
294 statistically learned mapping. Future studies will be necessary to investigate the
295 reliability of this hypothesis.

296

297 Fundings

298 This work was supported by Fyssen Foundation, by the European Research Council
299 under the European Union's Horizon 2020 research and innovation program grant
300 agreement No 716931 - GESTIMAGE - ERC-2016-STG, by the "Agence Nationale de
301 le Recherche" ANR-16-CONV-0002 (ILCB) and the Excellence Initiative of Aix-
302 Marseille University (A*MIDEX).

303

304 References

305

306 [1]. Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention,*
307 *Perception, & Psychophysics*, 73(4), 971-995.

308

309 [2]. Spence, C. (2019). On the relative nature of (pitch-based) crossmodal
310 correspondences. *Multisensory Research*, 32(3), 235-265.

311

312 [3]. Hanson-Vaux, G., Crisinel, A. S., & Spence, C. (2013). Smelling shapes:
313 Crossmodal correspondences between odors and shapes. *Chemical senses*, 38(2),
314 161-166.

315

316 [4] K. M. Kösterle and C. Spence, "Crossmodal correspondences between sounds
317 and tastes," *Psychon. Bull. Rev.*, 19, 992–1006 (2012).

318

319 [5] Christensen, Carol M. 1980. Effects of taste quality and intensity on oral
320 perception viscosity. *Perception and Psychophysics* 28:315–320.

321

322 [6] Calvert, G., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of*
323 *multisensory processes*. MIT press.

324

325 [7] Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and
326 touch. *Journal of vision*, 7(5), 7-7.

327

328 [8] Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into
329 perception, thought and language. *Journal of consciousness studies*, 8(12), 3-34.

330

331 [9]. Cuskley, C., & Kirby, S. (2013). Synesthesia, cross-modality, and language
332 evolution. *Oxford handbook of synaesthesia*, 20, 869-907.

333

334 [10] Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings
335 between high luminance and high pitch are shared by chimpanzees (*Pan*
336 *troglodytes*) and humans. *Proceedings of the National Academy of*
337 *Sciences*, 108(51), 20661-20665.

338

339 [11]. Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the
340 temporal ventriloquism effect. *Neuroscience Letters*, 442(3), 257-261.

341

342 [12]. Walsh, V. (2003). A theory of magnitude: common cortical metrics of time,
343 space and quantity. *Trends in cognitive sciences*, 7(11), 483-488.

344

- 345 [13]. Korzeniowska, A. T., Root-Gutteridge, H., Simner, J., & Reby, D. (2019).
346 Audio–visual crossmodal correspondences in domestic dogs (*Canis*
347 *familiaris*). *Biology letters*, *15*(11), 20190564.
- 348 [14]. Korzeniowska, A. T., Simner, J., Root-Gutteridge, H., & Reby, D. (2022). High-
349 pitch sounds small for domestic dogs: abstract crossmodal correspondences
350 between auditory pitch and visual size. *Royal Society Open Science*, *9*(2), 211647.
351
- 352 [15]. Loconsole, M., Pasculli, M. S., & Regolin, L. (2021). Space-luminance
353 crossmodal correspondences in domestic chicks. *Vision Research*, *188*, 26-31.
354
- 355 [16]. Loconsole, M., Gasparini, A., & Regolin, L. (2022). Pitch–Luminance
356 Crossmodal Correspondence in the Baby Chick: An Investigation on Predisposed
357 and Learned Processes. *Vision*, *6*(2), 24.
358
- 359 [17] Gaver, W. W. (1993). What in the world do we hear?: An ecological approach to
360 auditory event perception. *Ecological psychology*, *5*(1), 1-29.
361
- 362 [18]. Marks, L. E. (1974). On associations of light and sound: The mediation of
363 brightness, pitch, and loudness. *The American journal of psychology*, 173-188.
364
- 365 [19]. Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object
366 correspondences in young children. *Cognitive, Affective, & Behavioral*
367 *Neuroscience*, *4*(2), 133-136.
368
- 369 [20]. Haryu, E., & Kajikawa, S. (2012). Are higher-frequency sounds brighter in color
370 and smaller in size? Auditory–visual correspondences in 10-month-old infants. *Infant*
371 *Behavior and Development*, *35*(4), 727-732.
372
- 373 [21]. Margiotoudi, K., Allritz, M., Bohn, M., & Pulvermüller, F. (2019). Sound symbolic
374 congruency detection in humans but not in great apes. *Scientific reports*, *9*(1), 1-12.
375
- 376 [22]. Margiotoudi, K., Bohn, M., Schwob, N., Taglialatela, J., Pulvermüller, F.,
377 Epping, A., ... & Allritz, M. (2022). Bo-NO-bouba-kiki: picture-word mapping but no
378 spontaneous sound symbolic speech-shape mapping in a language trained
379 bonobo. *Proceedings of the Royal Society B*, *289*(1968), 20211717.
380
- 381 [23]. Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in
382 monkeys: Use of a battery of computerized test systems by a troop of semi-free-
383 ranging baboons (*Papio papio*). *Behavior research methods*, *42*(2), 507-516.
384
- 385 [24]. Fagot, J., Gullstrand, J., Kemp, C., Defilles, C., & Mekaouche, M. (2014).
386 Effects of freely accessible computerized test systems on the spontaneous
387 behaviors and stress level of Guinea baboons (*Papio papio*). *American Journal of*
388 *Primatology*, *76*(1), 56-64.
389
- 390 [25]. Townsend JT, Ashby FG (1983) Stochastic Modelling of Elementary
391 Psychological Processes (Cambridge Univ Press, New York).
392

- 393 [26]. Hashiya, K., & Kojima, S. (2001). Acquisition of auditory–visual intermodal
394 matching-to-sample by a chimpanzee (pan troglodytes): comparison with visual—
395 visual intramodal matching. *Animal Cognition*, 4(3), 231-239.
396
- 397 [27]. Martinez, L., & Matsuzawa, T. (2009). Visual and auditory conditional position
398 discrimination in chimpanzees (Pan troglodytes). *Behavioural Processes*, 82(1), 90-
399 94.
400
- 401 [28]. Ennaji, F. (2022). Voice perception in non-human primates : A behavioral study
402 of categorization of vocal and non-vocal sounds in Guinea baboon (*Papio papio*),
403 PhD dissertation, Aix-Marseille University.
404
- 405 [29] Team, R. C. (2013). R: A language and environment for statistical computing.
406
- 407 [30] Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-
408 effects models using lme4. *Journal of Statistical Software* 67: 1–48.
409